**Лабораторная работа №8**

Выполнила: студент группы 4309

Канькина Ольга Александровна

Семантический анализатор: он использует синтаксическое дерево и таблицу символов, чтобы проверить, соответствует ли данная программа семантически определению языка. Он собирает информацию о типе и сохраняет ее либо в синтаксическом дереве, либо в таблице символов. Эта информация о типе впоследствии используется компилятором при генерации промежуточного кода.

Следующий шаг анализа текста программы – семантический, существенно отличается от синтаксического. И дело не столько в том, что фаза семантического анализа реализуется не формальными, а содержательными методами (т. е. на данный момент нет универсальных математических моделей и формальных средств описания «смысла» программы). Лексический и синтаксический анализ имеют дело со структурными, т. е. внешними, текстовыми конструкциями языка. Семантика же, ориентированная на содержательную интерпретацию, имеет дело с внутренним представлением «смысла» объектов, описанных в программе. Для любого, имеющего опыт практического программирования, ясно, что формальные конструкции языка дают описание свойств и действий над внутренними объектами, с которыми имеет дело программа. Для начала перечислим все, что их касается и лежит на поверхности:

· большинство объектов являются именованными. Имя объекта позволяет его идентифицировать, существуют различные области действия имен, соглашения об именах, различные умолчания и т. п.. Все это относится к семантике;

· виды, сложность и набор характеристик объектов различаются в разных языках программирования и сильно зависят от области приложения языка (в этом смысле семантика языков программирования более разнообразна, нежели синтаксис и лексика). Например, классический Си, ориентированный на максимальное приближение к архитектуре компьютера, работает с такими объектами, как типы данных, переменные, функции. Все они имеют различные свойства и характеристики. Например, переменная характеризуется именем, типом данных, размерностью, областью действия, временем жизни, текущим значением;

· объекты связаны между собой (ссылаются друг на друга). В том же Си переменная ссылается на описание того типа данных, к которому она относится, далее производный тип данных ссылается на базовый и т. п.. Можно сказать, что семантика программы во внутреннем представлении выглядит как система взаимосвязанных объектов;

· внутреннее представление семантики программы не совсем удачно называется семантическими таблицами. На самом деле структура данных, соответствующая представлению семантики, может быть любой. Термин «таблицы» говорит о том, что имеются множества объектов различных типов, для каждого из которых заведена отдельная таблица, но нельзя забывать, что элементы различных таблиц связаны между собой. Наиболее близкий термин для описания подобной системы – база данных.

Семантика программы – внутренняя модель (база данных) множества именованных объектов, с которыми работает программа, с описанием их свойств, характеристик и связей.

Семантика программы тоже не обладает структурной целостностью и представлена фрагментарно, но при этом связана с синтаксисом следующим образом:

· один и тот же семантический объект (например, переменная) может встречаться в различных, синтаксически несвязанных частях программы;

· синтаксические конструкции описаний, определений и объявлений являются источником семантики объектов программы, они «заявляют» о существовании объектов и задают их свойства;

· синтаксические конструкции, связанные с действиями, выполняемыми над объектами, являются потребителями семантики, их интерпретация, корректность, «смысл» зависят от семантических свойств объекта. Забегая вперед, можно заметить, что заключительная фаза трансляции (генерация кода, интерпретация) может рассматриваться как особые семантические действия, производимые над объектами;

· первичным источником семантики является лексический анализ. Некоторые из лексем (например, идентификаторы и константы) наряду с классом лексемы (обозначение выходной единицы лексического анализа), т. е. символом (обозначение той же единицы на входе синтаксического анализатора) имеют значение. Значением лексемы является сама распознанная цепочка литер, она и представляет семантическую составляющую лексемы, которая попадает в семантические таблицы;

· лексемы, или то же самое, что терминальные символы входной строки (в терминах синтаксического анализа), ссылаются в семантические таблицах на свою семантику. В дальнейшем каждый промежуточный нетерминал также ссылается на собственную семантику. При этом любое правило преобразует семантику терминалов и нетерминалов правой части в семантику нетерминала левой части при помощи назначенной правилу семантической процедуры. Таким образом, формирование семантической составляющей связано с движением снизу вверх по синтаксическому дереву, от вершин – потомков к предкам.

В процессе семантического анализа проверяется наличие семантических ошибок в исходной программе и накапливается информация о типах для следующей стадии – генерации кода. При семантическом анализе используются иерархические структуры, полученные во время синтаксического анализа для идентификации операторов и операндов выражений и инструкций.

Важным аспектом семантического анализа является проверка типов, когда компилятор проверяет, что каждый оператор имеет операнды допустимого спецификациями языка типа. Например, определение многих языков программирования требует, чтобы при использовании действительного числа в качестве индекса массива генерировалось сообщение об ошибке. В то же время спецификация языка может позволить определенное преобразование типов, например, когда бинарный арифметический оператор применяется к операндам целого и действительного типов. В этом случае компилятору может потребоваться преобразование целого числа в действительное.

Все типы известны во время компиляции, и это относится к типам выражений, идентификаторам и литералам. При этом неважно, насколько сложным является выражение: его тип может определяться во время компиляции за определенное количество шагов исходя из типов его составляющих. Фактически, это позволяет производить контроль типов во время компиляции и находить в процессе компиляции многие программные ошибки.

В процессе компиляции анализатору требуются таблицы:

– таблица символов;

– таблица типов;

– таблица функций;

– таблица меток.

Основная задача таблицы символов – установление соответствия между переменной и ее типом. Операции, связанные с таблицей символов:

– операция, соответствующая определяющему вхождению переменной. В таблицу символов помещаются имя переменной и ее тип;

– операция, соответствующая применимому вхождению переменной. Исследуется таблица символов для нахождения типа переменной.

Как известно, один и тот же синтаксис может быть представлен различными грамматиками. Они эквивалентны между собой, поскольку распознают одно и то же множество цепочек (Как известно, проблема эквивалентности формальных грамматик алгоритмически неразрешима, т. е. невозможно разработать алгоритм, устанавливающий эквивалентность двух произвольных грамматик). Мы уже видели, что требования к грамматикам восходящего и нисходящего разбора также различны. Различны и технологически подходы к реализации одних и тех же элементов синтаксиса (повторений, необязательных элементов синтаксиса и т. п.). Поэтому получается, что в различных грамматиках при разборе одного и того же предложении будут строиться различные деревья, но результат трансляции (например, последовательность выполняемых операций в выражении) должен быть если не одинаковым, то эквивалентным(разобрать методы и описать свой, задание написать).

Ошибки, распознанные семантическим анализатором, заключаются в следующем:

1. Несоответствие типов
2. Необъявленные переменные
3. Неправильное использование идентификатора

В данной лабораторной работе был использован третий пункт, я прописала условия для идентификатора, при неправильном использовании которых выдавало сообщение об ошибке

Функции семантического анализа:

1. Проверка типов – Гарантирует, что типы данных используются в соответствии их определением.
2. Проверка меток – Программа должна содержать ссылки на ярлыки.
3. Проверка управления потоком – Проверяет, что управляющие структуры используются надлежащим образом.

**Задание**. Создать семантический анализатор.

Код программы:

uSemantAnalyzer.cs:

using System;

using System.Collections.Generic;

using System.Drawing;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Windows.Forms;

namespace WindowsFormsApp26

{

public class uSemantAnalyzer

{

public int i = 0;

public string strIndentifier; // значение идентификатора

public string strDigital; //значение числа

private TreeView tree;

public uSemantAnalyzer()

{

}

public uSemantAnalyzer(TreeView treeView)

{

tree = treeView;

TreeController(tree);

}

public void TreeController(TreeView tree)

{

foreach (TreeNode node in tree.Nodes) //Для каждого узла TreeNode в дереве вызывается функция

{

TreeController(node);

}

}

public void TreeController(TreeNode node)

{

if (node.Text == "B") //текст узла равен "B"

{

if (node.Nodes.Count > 1) //количество подузлов больше 1

{

strIndentifier = node.Nodes[0].Text.ToString(); //из первого подузла извлекается значение идентификатора

strDigital = node.Nodes[1].Text.ToString(); //из третьего- значение числа

Check(strIndentifier, strDigital, node);

}

}

foreach (TreeNode childNode in node.Nodes) //проверкf корректности значений идентификатора и числа

{

TreeController(childNode);

}

}

private void Check(string ident, string digit, TreeNode node)

{

int countI = 0;

for (int i = ident.Length - 1; i >= 0; --i)

{

if (ident[i] == 'a')

{

countI++;

}

else

{

break; // Прерываем цикл, если не встречаем 'a'

}

}

int countA = digit.Length / 5 + 1 ;

Console.WriteLine(countI);

Console.WriteLine(countA);

if (countA != countI)

{

tree.SelectedNode = node;

tree.SelectedNode.BackColor = Color.Red;

throw new Exception("Error: count a > count dop digit");

}

}

}

}

Form1.cs:

using nsLex;

using nsLexMainForm;

using nsSynt;

using System;

using System.Collections.Generic;

using System.ComponentModel;

using System.Data;

using System.Drawing;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Windows.Forms;

using WindowsFormsApp26;

namespace CHashTableList

{

public partial class Form1 : Form

{

Dictionary<int, List<string>> hashTableIdentifier = new Dictionary<int, List<string>>();

Dictionary<int, List<string>> hashTableDigital = new Dictionary<int, List<string>>();

Dictionary<int, List<string>> hashTableRezerv = new Dictionary<int, List<string>>();

public MyHashFunction hashFunction = new MyHashFunction();

public CHashTableList htl = new CHashTableList(2);

public Form1()

{

InitializeComponent();

tbFSource.AppendText("100\_a\_10010\_a\_" + "\r\n");

int n = tbFSource.Lines.Length;

}

private void Form1\_Load(object sender, EventArgs e)

{

}

public void TablesToMemo(object sender, System.EventArgs e)

{

List<string> listTable = new List<string>();

listBox1.Items.Clear();

listBox2.Items.Clear();

listBox3.Items.Clear();

htl.TableToStringList(0, listTable);

//for (int i = 0; i < listTable.Count; i++)

// listBox1.Items.Add(listTable[i]);

foreach (var entry in hashTableIdentifier)

{

listBox1.Items.Add(string.Join(", ", entry.Value));

}

listTable.Clear();

htl.TableToStringList(1, listTable);

//for (int i = 0; i < listTable.Count; i++)

// listBox2.Items.Add(listTable[i]);

foreach (var entry in hashTableDigital)

{

listBox2.Items.Add(string.Join(", ", entry.Value));

}

listTable.Clear();

foreach (var entry in hashTableRezerv)

{

listBox3.Items.Add(string.Join(", ", entry.Value));

}

listTable.Clear();

}

private void btnFStart\_Click\_1(object sender, EventArgs e)

{

tbFMessage.Clear();

uSyntAnalyzer Synt = new uSyntAnalyzer(SyntTree); // Используем существующий экземпляр treeView2

Synt.Lex.strPSource = tbFSource.Lines;

Synt.Lex.strPMessage = tbFMessage.Lines;

Synt.Lex.enumPState = TState.Start;

try

{

Synt.Lex.NextToken();

Synt.S();

SyntTree.ExpandAll();

throw new Exception("Текст верный");

}

catch (Exception exc)

{

tbFMessage.Text += exc.Message;

tbFSource.Select();

tbFSource.SelectionStart = 0;

int n = 0;

for (int i = 0; i < Synt.Lex.intPSourceRowSelection; i++) n += tbFSource.Lines[i].Length + 2;

n += Synt.Lex.intPSourceColSelection;

tbFSource.SelectionLength = n;

}

}

private void btnFRecord\_Click\_1(object sender, EventArgs e)

{

CLex Lex = new CLex();

Lex.strPSource = tbFSource.Lines;

Lex.strPMessage = tbFMessage.Lines;

//Lex.intPSourceColSelection = 0;

//Lex.intPSourceRowSelection = 0;

int x = tbFSource.TextLength;

int y = tbFSource.Lines.Length;

tbFMessage.Text = "";

try

{

while (Lex.enumPState != TState.Finish)

{

Lex.NextToken();

string s1 = "", s = "";

switch (Lex.enumPToken)

{

case TToken.lxmIdentifier:

{

hashFunction.AddWord(hashTableIdentifier, Lex.strPLexicalUnit);

s1 = "id " + Lex.strPLexicalUnit; int b = 0;

if (htl.AddLexicalUnit(Lex.strPLexicalUnit, 0, ref b))

{

TablesToMemo(this, e);

}

break;

}

case TToken.lxmNumber:

{

hashFunction.AddWord(hashTableDigital, Lex.strPLexicalUnit);

s1 = "num " + Lex.strPLexicalUnit; int b = 0;

if (htl.AddLexicalUnit(Lex.strPLexicalUnit, 1, ref b))

{

TablesToMemo(this, e);

}

break;

}

case (TToken.lxmdt):

{

hashFunction.AddWord(hashTableRezerv, ":");

s1 = "rez " + Lex.strPLexicalUnit; int b = 0;

if (htl.AddLexicalUnit(Lex.strPLexicalUnit, 1, ref b))

{

TablesToMemo(this, e);

}

break;

}

case (TToken.lxmComma):

{

hashFunction.AddWord(hashTableRezerv, ",");

s1 = "rez " + Lex.strPLexicalUnit; int b = 0;

if (htl.AddLexicalUnit(Lex.strPLexicalUnit, 1, ref b))

{

TablesToMemo(this, e);

}

break;

}

case (TToken.lxmmul):

{

hashFunction.AddWord(hashTableRezerv, "\_");

s1 = "rez " + Lex.strPLexicalUnit; int b = 0;

if (htl.AddLexicalUnit(Lex.strPLexicalUnit, 1, ref b))

{

TablesToMemo(this, e);

}

break;

}

}

String m = "(" + s + "" + s1 + ")";

tbFMessage.Text += m;

}

}

catch (Exception exc)

{

tbFMessage.Text += exc.Message;

tbFSource.Select();

tbFSource.SelectionStart = 0;

int n = 0;

for (int i = 0; i < Lex.intPSourceRowSelection; i++) n += tbFSource.Lines[i].Length + 2;

n += Lex.intPSourceColSelection;

tbFSource.SelectionLength = n;

}

}

private void searchBtn\_Click\_1(object sender, EventArgs e)

{

if (hashFunction.SearchWord(hashTableIdentifier, listBox1.SelectedItem.ToString()) == 1)

{

searchBtn.BackColor = Color.Green;

}

else

{

searchBtn.BackColor = Color.Red;

}

}

private void addBtn\_Click\_1(object sender, EventArgs e)

{

hashFunction.AddWord(hashTableIdentifier, textBox1.Text.ToString());

}

private void deleteBtn\_Click\_1(object sender, EventArgs e)

{

if (hashFunction.RemoveWord(hashTableIdentifier, listBox1.SelectedItem.ToString()))

{

deleteBtn.BackColor = Color.Green;

}

else

{

deleteBtn.BackColor = Color.Red;

}

}

private void changeBtn\_Click\_1(object sender, EventArgs e)

{

if (hashFunction.RemoveWord(hashTableIdentifier, listBox1.SelectedItem.ToString()))

{

hashFunction.AddWord(hashTableIdentifier, textBox1.Text.ToString());

changeBtn.BackColor = Color.Green;

}

else

{

changeBtn.BackColor = Color.Red;

}

}

private void reloadBtn\_Click\_1(object sender, EventArgs e)

{

listBox1.Items.Clear();

listBox2.Items.Clear();

listBox3.Items.Clear();

foreach (var entry in hashTableIdentifier)

{

listBox1.Items.Add(string.Join(", ", entry.Value));

}

foreach (var entry in hashTableDigital)

{

listBox2.Items.Add(string.Join(", ", entry.Value));

}

foreach (var entry in hashTableRezerv)

{

listBox3.Items.Add(string.Join(", ", entry.Value));

}

}

private void convertBtn\_Click(object sender, EventArgs e)

{

CopyTreeView(SyntTree, SyntTree2);

TreeConverter treeConverter = new TreeConverter();

treeConverter.ConvertBinaryToDecimal(SyntTree2);

}

private void CopyTreeNodes(TreeNodeCollection sourceNodes, TreeNodeCollection targetNodes)

{

foreach (TreeNode sourceNode in sourceNodes)

{

TreeNode newNode = new TreeNode(sourceNode.Text);

targetNodes.Add(newNode);

CopyTreeNodes(sourceNode.Nodes, newNode.Nodes); // Рекурсивно копируем дочерние узлы

}

}

private void CopyTreeView(TreeView sourceTreeView, TreeView targetTreeView)

{

targetTreeView.Nodes.Clear(); // Очищаем существующие узлы в целевом TreeView

CopyTreeNodes(sourceTreeView.Nodes, targetTreeView.Nodes);

}

private void duplicateBtn\_Click(object sender, EventArgs e)

{

try

{

Duplicate duplicate = new Duplicate(SyntTree);

duplicate.ChekingForDuplicate(SyntTree);

throw new Exception("Текст верный");

}

catch (Exception exc)

{

tbFMessage.Text = "";

tbFMessage.Text += exc.Message;

}

}

private void treeView1\_AfterSelect(object sender, TreeViewEventArgs e)

{

}

public void SemantFunction()

{

try

{

uSemantAnalyzer Semant = new uSemantAnalyzer(SyntTree);

}

catch (Exception exc)

{

tbFMessage.Text += exc.Message;

}

}

}

}

Analyzer.cs:

using System;

using System.Windows.Forms;

using nsLex;

using WindowsFormsApp26;

namespace nsSynt

{

internal class uSyntAnalyzer

{

private String[] strFSource;

private String[] strFMessage;

public String[] strPSource { set { strFSource = value; } get { return strFSource; } }

public String[] strPMessage { set { strFMessage = value; } get { return strFMessage; } }

public CLex Lex = new CLex();

private TreeView tree; // Declare TreeView for function calls

public uSyntAnalyzer(TreeView treeView)

{

tree = treeView;

}

public void S()

{

TreeNode parent = new TreeNode("S");

tree.Nodes.Add(parent);

Console.WriteLine("S");

A(parent);

Console.WriteLine(Lex.enumPToken);

while (Lex.enumPToken == TToken.lxmmul || Lex.enumPToken != TToken.lxmEmpty)

{

parent.Nodes.Add(new TreeNode("\_"));

TreeNode highParent = new TreeNode("S");

parent.Nodes.Add(highParent);

Lex.NextToken();

A(highParent);

}

throw new Exception("Текст верный.");

}

public void C(TreeNode highParent)

{

highParent.Nodes.Add(new TreeNode("\_"));

Console.WriteLine("C");

if (Lex.enumPToken == TToken.lxmmul)

{

Lex.NextToken();

B(highParent);

if (Lex.enumPToken == TToken.lxmmul)

{

C(highParent);

}

}

}

public void A(TreeNode highParent)

{

Console.WriteLine("A");

TreeNode parent = new TreeNode("A");

highParent.Nodes.Add(parent);

B(parent);

if (Lex.enumPToken == TToken.lxmmul)

{

//parent.Nodes.Add (new TreeNode(","));

C(parent);

}

if (Lex.enumPToken == TToken.lxmNumber || Lex.enumPToken == TToken.lxmIdentifier)

{

throw new Exception("ERROR");

}

}

public void B(TreeNode highParent)

{

TreeNode parent = new TreeNode("B");

highParent.Nodes.Add(parent);

Console.WriteLine("B");

if (Lex.enumPToken == TToken.lxmNumber)

{

parent.Nodes.Add(new TreeNode(Lex.strPLexicalUnit));

while (Lex.enumPToken == TToken.lxmNumber) Lex.NextToken();

}

else if (Lex.enumPToken == TToken.lxmIdentifier)

{

parent.Nodes.Add(new TreeNode(Lex.strPLexicalUnit));

while (Lex.enumPToken == TToken.lxmIdentifier)

{

Lex.NextToken();

}

if (Lex.enumPToken == TToken.lxmmul)

{

parent.Nodes.Add(new TreeNode("\_"));

Lex.NextToken();

if (Lex.enumPToken == TToken.lxmNumber)

{

parent.Nodes.Add(new TreeNode(Lex.strPLexicalUnit));

uSemantAnalyzer uSemant = new uSemantAnalyzer(tree);

while (Lex.enumPToken == TToken.lxmNumber) Lex.NextToken();

}

else throw new Exception("Текст верный");

}

else throw new Exception("Ожидалось \_");

}

else throw new Exception("Ожидался идентификатор");

}

}

}

Результат работы программы:



