

Лабораторная работа №2. Ручное построение нисходящих синтаксических анализаторов

Шовкопляс Григорий М3339

28 апреля 2016

Задание: Описание переменных в Паскале – блок описания переменных в Паскале. Блок начинается ключевым словом **"var"**, далее идет несколько описаний переменных. Описание переменных содержит имена переменных через запятую, затем двоеточие, затем имя типа. Достаточно рассматривать только примитивные типы (массивы, записи и т. п. не требуется). Используйте один терминал для всех имен переменных и имен типов. Используйте один терминал для ключевого слова **var** (не три **'v'**, **'a'**, **'r'**). Пример: **var a, b: integer; c: integer;**

1) Грамматика

Построим грамматику по описанию задания (a – буква латинского алфавита в любом регистре, d – десятичная цифра)

$S \rightarrow V T$

$V \rightarrow \text{var}$

$T \rightarrow L:N; | L:N;T$

$L \rightarrow N|N,L$

$N \rightarrow a|Na|Nd$

Нетерминал	Описание
S	Описание переменных в Паскале
V	Служебное слово var
T	Список объявлений переменных
L	Перечисление имен переменных
N	Имя переменной

В грамматике есть левая рекурсия – устраним ее

$S \rightarrow V T$

$V \rightarrow \text{var}$

$T \rightarrow L:N; | L:N;T$

$L \rightarrow N|N,L$

$N \rightarrow a|aN'$

$N' \rightarrow a|aN' | d|dN'$

Нетерминал	Описание
S	Описание переменных в Паскале
V	Служебное слово var
T	Список объявлений переменных
L	Перечисление имен переменных
N	Имя переменной
N'	Продолжение имени переменной

В грамматике есть правое ветвление – устраним его

$S \rightarrow V T$

$V \rightarrow \text{var}$

$$\begin{aligned}
T &\rightarrow L:N; \mid L:N;T \\
L &\rightarrow N \mid N, L \\
N &\rightarrow aN'' \\
N'' &\rightarrow N' \mid \varepsilon \\
N' &\rightarrow aN'' \mid dN''
\end{aligned}$$

Нетерминал	Описание
S	Описание переменных в Паскале
V	Служебное слово var
T	Список объявлений переменных
L	Перечисление имен переменных
N	Имя переменной
N''	Продолжение имени переменной
N'	Непустое продолжение имени переменной

2) Построение лексического анализатора

В грамматике 7 терминалов – **var**, а (универсальный терминал для обозначения буквы), d (универсальный терминал для обозначения цифры), ', ', ';', ':', '\$'.

Заведем класс **Token** для хранения терминалов, не забудем про конец строки.

```

public enum Token {
    VAR, CHARACTER, DIGIT, COMMA, SEMICOLON, COLON, END;
}

```

Терминал	Токен
var	VAR
a [a-z A-Z]	CHARACTER
d [0-9]	DIGIT
,	COMMA
;	SEMICOLON
:	COLON
\$	END

3) Построение синтаксического анализатора

Построим множества **FIRST** и **FOLLOW** для нетерминалов грамматики

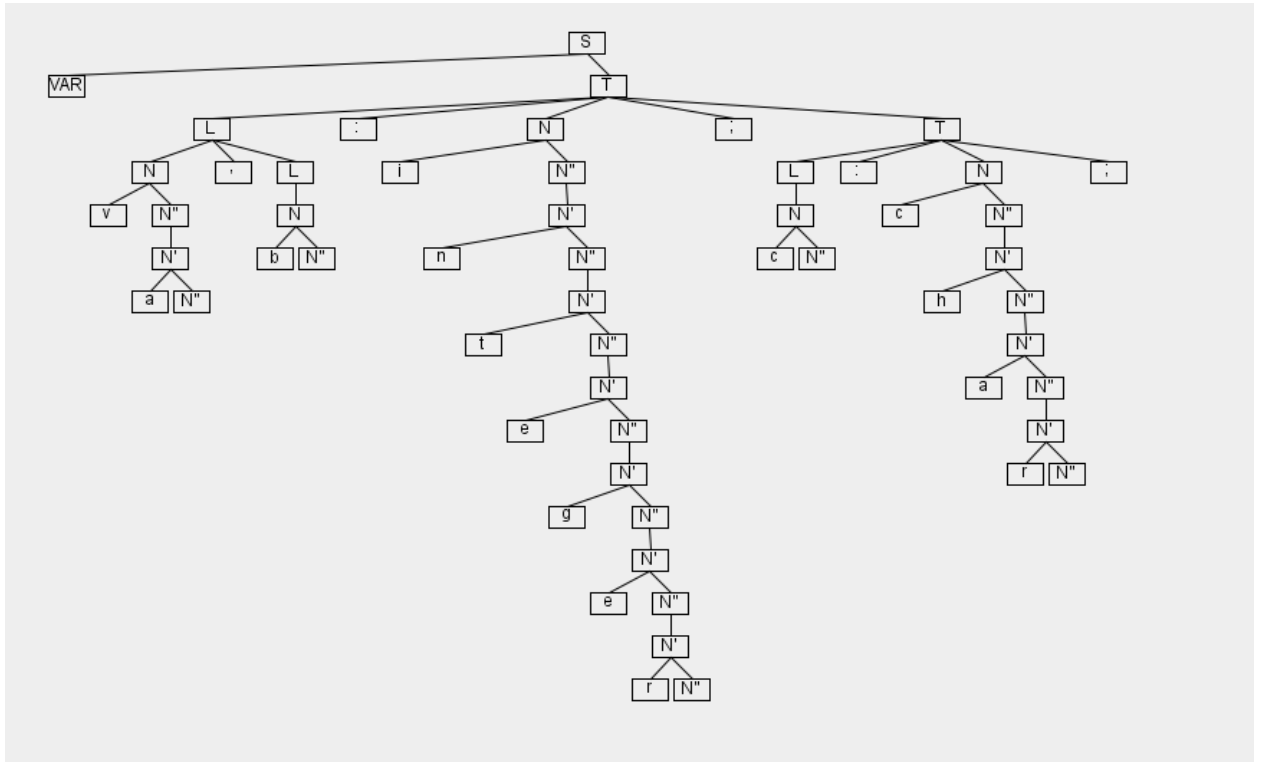
Нетерминал	FIRST	FOLLOW
S	var	\$
V	Var	a
T	a	\$
L	a	':'
N	a	':', \', \', \';'
N''	a, d, ε	':', \', \', \';'
N'	a, d	':', \', \', \';'

4) Визуализатор

Мной были реализованы два варианта визуализатора:

А) Вывод в выходной файл дерева в порядке обхода поиска в глубину с отступами, зависящими от глубины.

Б) При помощи сторонней библиотеки **org.StructureGraphic.v1** Рассмотрим на примере **"var a, b: integer; c: char;"**



5) Тестирование

Программа была вручную протестирована на исчерпывающем наборе тестов.

Тест	Описание
var va: vars;	Проверка обработки переменных имеющих вложенные символы со служебным словом var
Var VeRy: sTrAnGe;	Проверка поддержки разных регистров букв
var a, b: integer; c: char;	Ручной тест для нескольких типов
var a1A2: Int64;	Проверка поддержки цифр
Abacaba	Рандомная ручная строка
var a, integer;	Неправильный разделительный символ в нетерминале T
var 1: integer;	Неправильный первый символ в названии переменной

var b	Отсутствие продолжения
	Пустая строка
var b;	Преждевременный конец строки
var a,,b:int;	Переменная с нулевым именем
var a, b: integer	Отсутствие ; в конце
var a:b:c:d;	Игры с разделителем
var a, b: integer, vars: char;	Игры с разделителем
var a, b: integer: vars: char;	Игры с разделителем

Также программа была проверена на нескольких больших специально сгенерированных тестах.