Министерство образования и науки Российской Федерации

Санкт-Петербургский политехнический университет Петра Великого

—

Институт кибербезопасности и защиты информации

**КУРСОВАЯ РАБОТА**

**«Разработка парсера языка Golang»**

по дисциплине «Формальные грамматики и теория компиляторов»

Выполнили:

студент гр. 4851003/90801 Сороколетов С. А.

подпись   
 студент гр. 4851003/90801 Халилов Р. Н.

подпись

Руководитель

старший преподаватель Семьянов П.В.

подпись

Санкт-Петербург

2022

сОДЕРЖАНИЕ

[ЦЕЛЬ РАБОТЫ 3](#_Toc114212601)

[ЗАДАЧИ 4](#_Toc114212602)

[1 ХОД РАБОТЫ 5](#_Toc114212603)

[1.1 Теоретические сведения 5](#_Toc114212604)

[1.2 Разработка лексического анализатора языка 6](#_Toc114212605)

[1.3 Разработка синтаксического анализатора языка 6](#_Toc114212606)

[1.4 Конфликты грамматики 10](#_Toc114212607)

[1.4.1 Конфликты при обработке указателей 10](#_Toc114212608)

[1.4.2 Конфликт при обработке массивов и идентификаторов 12](#_Toc114212609)

[1.4.3 Конфликт при обработке типов и идентификаторов 12](#_Toc114212610)

[1.5 Трудности при разработке 13](#_Toc114212611)

[1.5.1 Обработка конструкции If-Else 14](#_Toc114212612)

[1.5.2 Обработка объявления переменных 14](#_Toc114212613)

[1.6 Примеры работы парсера 13](#_Toc114212614)

[2 РЕЗУЛЬТАТЫ 15](#_Toc114212615)

[3 ВЫВОД 16](#_Toc114212616)

[Приложение 1 – Лексический анализатор 18](#_Toc114212617)

[Приложение 2 – Синтаскический анализатор 21](#_Toc114212618)

ЦЕЛЬ РАБОТЫ

Получить навыки работы со средствами разработки парсеров языков программирования, а также построения грамматик, позволяющих распознать принадлежность входного кода к определенному языку программирования.

ЗАДАЧИ

Разработать парсер языка Golang используя возможности средств лексического и синтаксического анализа yacc и flex. Составить грамматику, позволяющую определить и разделить синтаксические конструкции языка программирования Go.

# ХОД РАБОТЫ

## Теоретические сведения

Golang - это статически типизированный, скомпилированный язык программирования. Синтаксически он похож на C, но с безопасностью памяти, сборкой мусора, структурной типизацией и параллелизмом в стиле CSP.

Go находится под влиянием C но с акцентом на большую простоту и безопасность. Язык Go имеет следующие характерные особенности синтаксиса:

* Необязательное краткое объявление переменной и инициализация с помощью вывода типа (x := 0 вместо int x = 0; или var x = 0;);
* Отсутствие ООП-структур, возможность определения собственных типов;
* Возможность использования анонимных функций, неопределенного количества параметров, возвращения нескольких значений;
* Система интерфейсов вместо виртуального наследования и встраивание типов вместо невиртуального наследования.

Пример базовой программы на языке Go:

package main

import “fmt”

func main {

var a = 1

var b int = 2

fmt.Println(a, b)

}

## Разработка лексического анализатора языка

При разработке парсера таких комплексных языков программирования как Go, задачу определения конкретных структур кода берет на себя лексический анализатор.

Так, с помощью возможностей flex и применением методов регулярных выражений, было разработано решение, позволяющее определять и обрабатывать ряд структур языка Go. Например, лексический анализатор может определять ключевые конструкции, обозначающие:

* Типы переменных – int8, int16, uint, float32, complex64 и так далее;
* Операторы – if, else, import, package и так далее;
* Идентификаторы – конструкции вида [A-Za-z\_]+[A-Za-z\_0-9]\*;
* Целые числа - конструкции вида [-]?[0-9]+;
* Дробные числа - конструкции вида [-]?[0-9]+[\.][0-9]+;
* Однострочные (//) и многострочные (/\* … \*/) комментарии.

Помимо перечисленных, разработанный лексический анализатор позволяет обрабатывать еще больший ряд конструкций, что позволяет полностью покрыть и обработать синтаксис языка Go, после чего передать полученную информацию уже синтаксическому анализатору. Полный код разработанного решения представлен в Приложении 1.

## Разработка синтаксического анализатора языка

Разработка грамматики основана на документации к языку, полученный с официального сайта языка. Go устроен таким образом, что каждая строка является, по сути, обособленной командой (в отличие от С++, где команды отделяются друг от друга символом «;»). Однако, Go отличается от, например, языка программирования Python тем, что табуляция в нем не имеет значения и блоки операций отделяются друг от друга парой символов «{…}»

Таким образом, при разработке грамматики ключевой и основной единицей была выбрана именно строка. Такая строка, условно называемая «Утверждением» (Statement, рис. 1), может содержать, например, началом оператора if else, вызовом функции, началом switch – конструкции, началом оператора for и так далее. Описание «Утверждения» в грамматике можно увидеть на Рисунке 1.

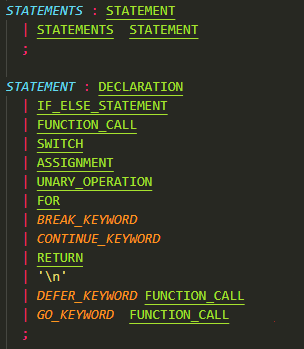


Рисунок 1 – Описание «Утверждения» в грамматике

При дальнейшей разработке грамматики статистического анализатора было установлено, что наиболее комплексной частью грамматики будет, как ни странно, объявление переменных. Это связано с тем, что в языке Go может встречаться несколько различных вариантов объявления переменных (подробнее – в пункте 1.5.2). Разработанное решение в грамматике, обрабатывающее различные возможные варианты объявления переменных представлено на Рисунках 2 и 3. Конкретно на Рисунке 2 показана общая обработка «Определения» (Declaration), через которую может раскрываться описанное выше «Утверждение». В свою очередь на Рисунке 3 представлены «Краткое объявление» (Short\_Defining), «Объявление множественных переменных» (Multiple\_Variable\_Declaration), «Простое объявление переменной» (Variable\_Declaration) и «Объявление переменной с присваиванием ей числового значения» (Variable\_Declaration\_Assignment).

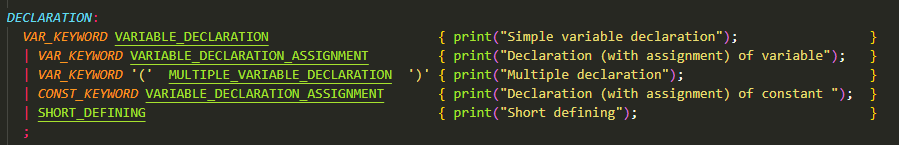


Рисунок 2 – Общее описание объявления переменной

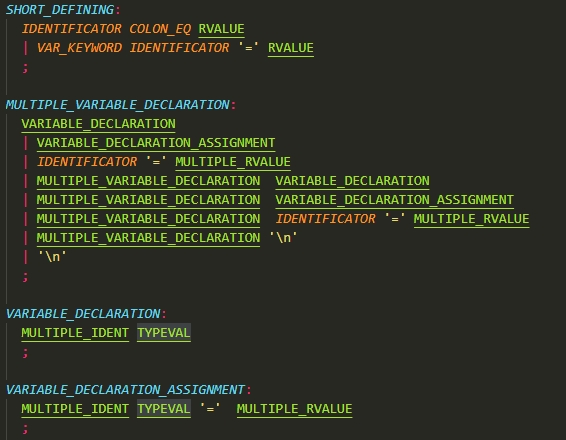


Рисунок 3 – Различные варианты объявления переменной

Помимо вышеупомянутого объявления переменных, разработанная грамматика также обрабатывает большое количество конструкций операций, перечисленных выше в формате «Утверждений». Однако, помимо этого синтаксический анализатор обрабатывает общие и более глобальные конструкции, стоящие отдельно от «Утверждений». Такими конструкциями являются импорты библиотек и объявления используемых пакетов. Обработка подобных конструкций представлена на Рисунке 4.

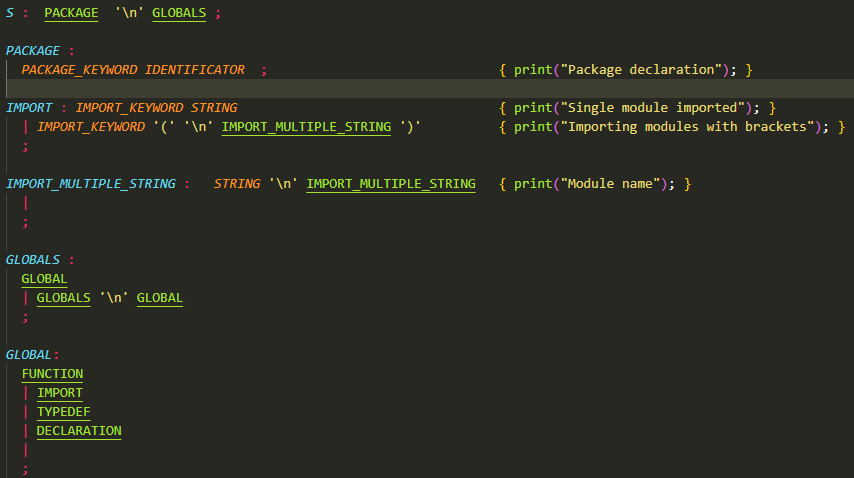


Рисунок 4 – Обработка импортов и объявления пакетов.

Также синтаксический анализатор определяет объявления функций, их вызовы, применения объявленных переменных, логические и арифметические операции и так далее. Разработка грамматики – наиболее трудоемкая и затратная по времени часть данной курсовой работы, основная сложность которой заключается в поддержании полноты и при этом лаконичности грамматики. Подробнее возникшие трудности описаны в пункте 1.5, полный код разработанной грамматики в Приложении 2.

## Конфликты грамматики

### Конфликты при обработке указателей

Во время введения в грамматику понятия указателей и их обработки возникли 3 конфликта вида *сдвиг/свертка*. Определение используемого в конфликтных строках правостороннего значения RVALUE представлено на Рисунке 5, а причиной возникновения конфликтов являются следующие строки грамматики:

POINTER '=' RVALUE { printf("Assignment of pointer by address. \n"); }

и RVALUE '\*' RVALUE

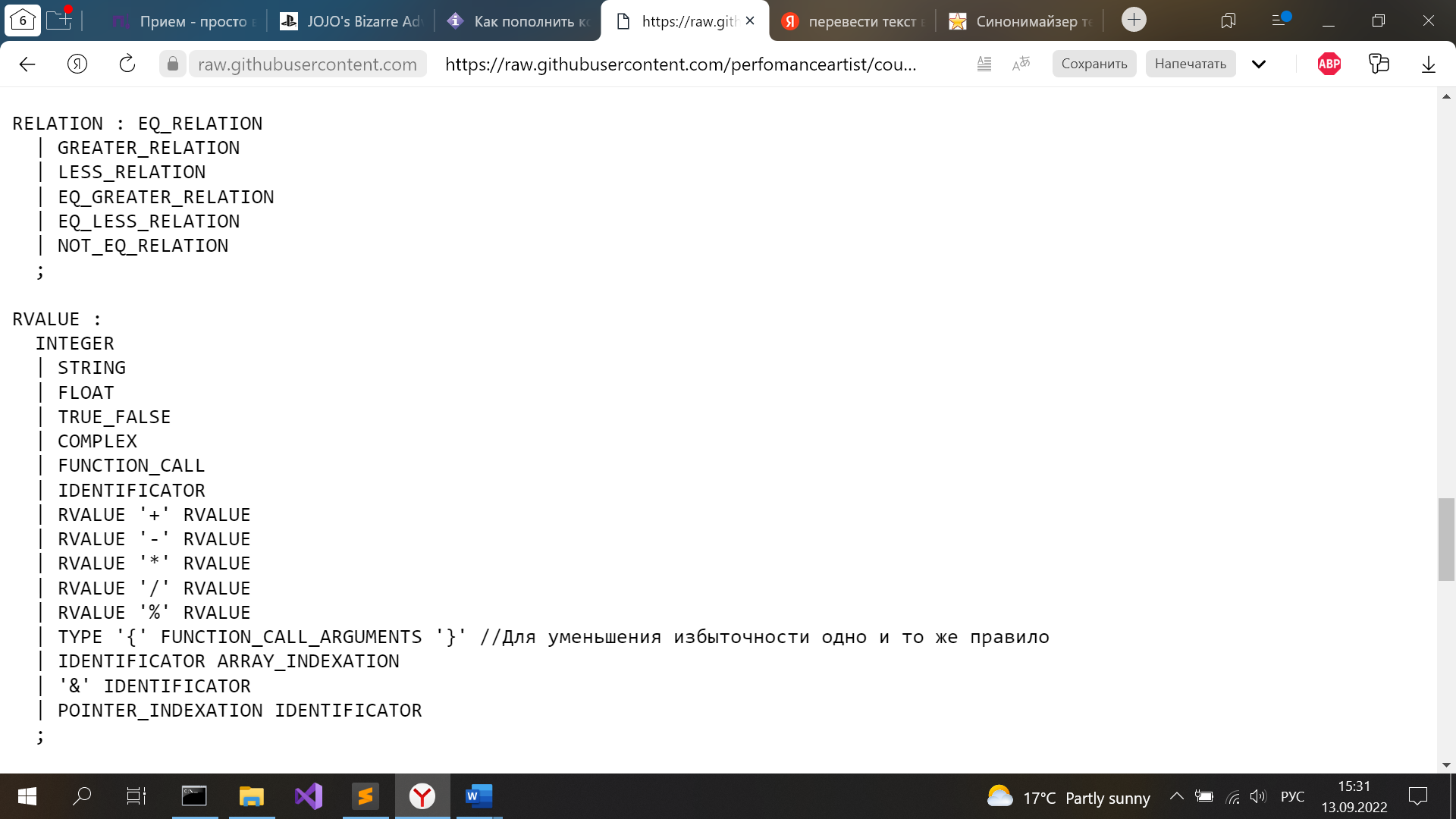


Рисунок 5 – Правостороннее значение

Конфликт возникает так как указатель имеет стандартный вид   
‘\*’ IDENTIFICATOR и при этом может являться правосторонним значением. Однако, правосторонним значением также может являться и конструкция вида RVALUE ‘\*’ RVALUE, где вместо второго RVALUE может стоять указатель, также начинающийся с символа звезды.

Таким образом, анализатор не может определить, какой именно конструкцией является ‘\*’ RVALUE – указателем или частью умножения. При попытках решения данных конфликтов был разработан вариант грамматики, представленный на Рисунке 6 и решающий 2 из 3 конфликтов.

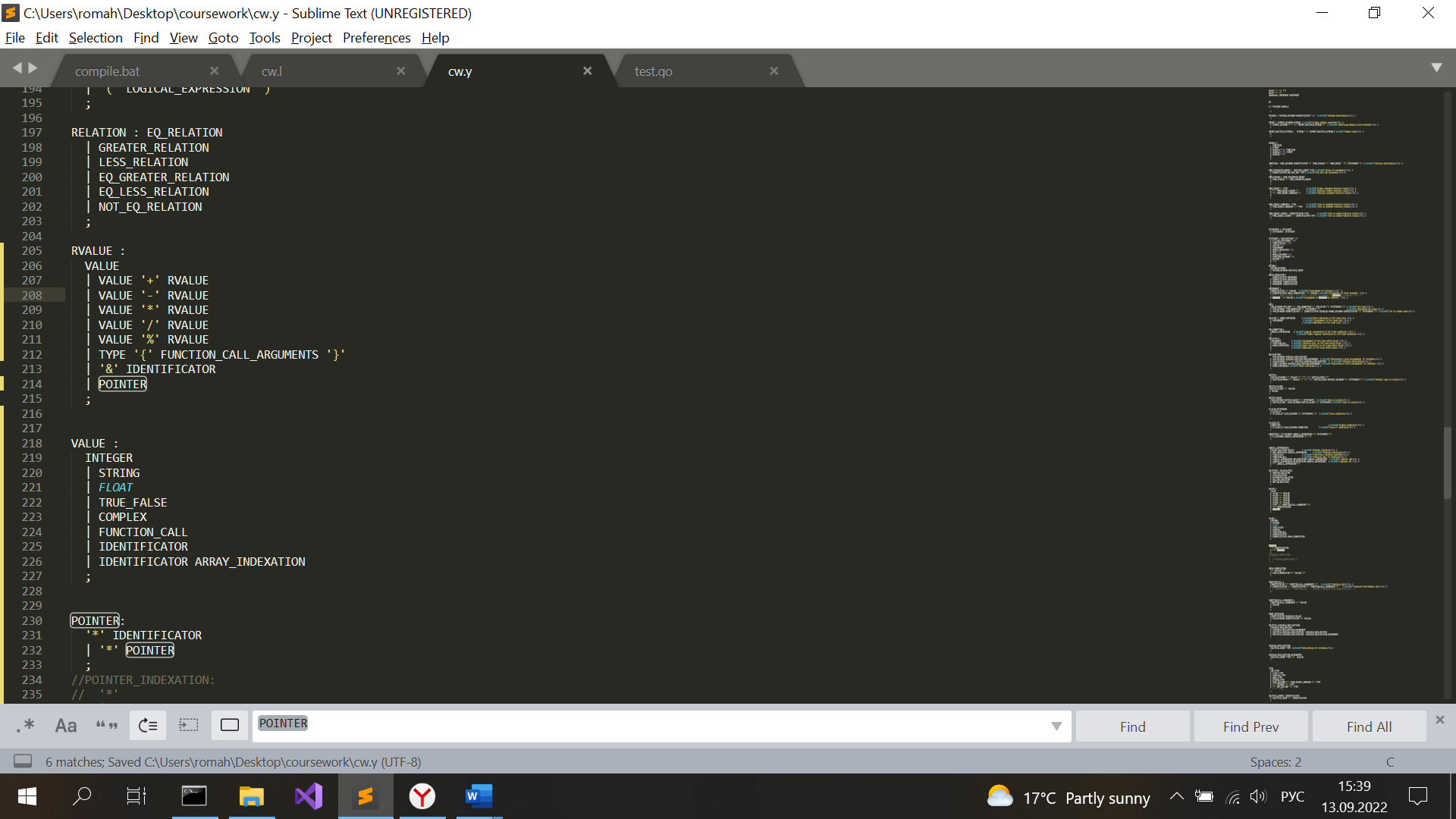


Рисунок 6 – Решение конфликтов.

Смысл решения заключается в ограничении левого значения операции, теперь оно не может быть указателем. Однако, такое решение накладывает ограничения на возможности парсера, так как хоть конструкция вида ЧИСЛО \* УКАЗАТЕЛЬ практически не используется в Go, она не противоречит грамматике языка, следовательно было принято решение вернуться к предыдущей версии грамматики.

### Конфликт при обработке массивов и идентификаторов

Обращение к элементу массива в языке Go проводится аналогично многим другим языкам программирования, и структурно выглядит следующим образом:

ИДЕНТИФИКАТОР[ИНДЕКС\_ЭЛЕМЕНТА]

Однако, при обращении, например, к переменной, также используется идентификатор. Описание данного аспекта языка в грамматике (Рисунок 7) очевидно вызывает конфликт вида *сдвиг/свертка*, так как при появлении идентификатора парсеру необходимо решить между сверткой к RVALUE, либо к сдвигу с ожиданием символа [. Данный конфликт не решается для грамматики из-за особенностей устройства большинства языков программирования, однако так как yacc обращается к дальнейшему контексту при принятии решения, он не мешает корректной работе парсера.

### Конфликт при обработке типов и идентификаторов

Язык программирования Go позволяет пользователям создавать, описывать и использовать в коде собственные типы данных. Чтобы учесть данный аспект языка в грамматике изначально было принято решение ввести понятие IDENTIFICATOR в понятие TYPE.

Однако, при дальнейших разработке и тестировании было обнаружено, что данное решение приводит к большому количеству логических проблем и конфликтов (а именно – 5 конфликтам вида *сдвиг/свертка* и 1 конфликту вида *сдвиг/сдвиг*). Например, при описании пользовательских структур возникали проблемы, так как рядом с идентификатором мог стоять тип данных, в который идентификатор мог сворачиваться.

Все данные конфликты были решены отделением понятий IDENTIFICATOR и TYPE, а также дублированием некоторых правил.

## Трудности при разработке

Язык программирования Go имеет комплексный синтаксис, разработанный под влиянием языков Си и Python. Задачей синтаксиса Golang является дать программистам возможности и удобства в написании кода, позволяя иметь вариативность и свободу при написании кода. Для этого в языке Go многие конструкции имеют несколько вариаций составления.

Однако, такое большое количество возможных вариаций стандартных конструкций усложняет задачу однозначного определения этих конструкций, то есть усложняет процесс составления грамматики, распознающей язык. Трудности, возникшие во время курсовой работы, описаны ниже.

### Обработка конструкции If-Else

При составлении грамматики одной из первых серьезных трудностей стала обработка конструкции If-Else. Проблема возникла, так как в языке Go при составлении конструкции не может быть пустых строк. Это можно увидеть на примере ниже.

package main

import "fmt"

func main() {

    var command = "идти на восток"

    if command == "идти на восток" {

         fmt.Println("Вы направляетесь к горе.")

    } else if command == "зайти внутрь" {

         fmt.Println("Вы заходите в пещеру, где будете жить до конца своей жизни.")

    } else {

         fmt.Println("Пока не совсем понятно.")

    }

}

Как видно на примере (строки 10 и 12), при построении конструкции If-Else после первого набора операций (конструкция if {…}), начало следующего набора операций (конструкция else {…}) должно начинаться с той же строки, на которой оканчивается первая конструкция.

Такая особенность отличает данную конструкцию от других, из-за чего возникают трудности. До этого наличие в коде пустых строк парсером просто игнорировалось, однако в данном конкретном моменте наличие пустой строки не допускается грамматикой Golang.

Таким образом, обработка пустых строк не могла быть одинаковой для любого участка кода. Для решения сложившейся проблемы пришлось изменить возвращаемые значения лексического анализатора и их обработку в синтаксическом анализаторе, а именно ввода обработки подобных пустых строк.

### Обработка объявления переменных

Грамматически задать определение конструкции объявления переменных вызвала некоторые сложности из-за большой вариативности записи данной команды в Golang.

Например, в языке программирования Си объявление переменной строится практически всегда либо в формате единичной переменной:

<type> <name> = <value>;

Либо в формате множественных переменных:

<type> <name> = <value>, <name> = <value>;

Или, для еще более простого в обработке примера, в языке Python все переменные объявляются простой конструкцией вида:

<name> = <val>

В отличие от этих простых примеров, язык Go имеет четыре возможных варианта объявления переменных. Эти варианты:

* Простое объявление – var <name>
* Объявление со значением – var <name> := <value>
* Краткое объявление – <name> := <value>
* Два варианта множественного объявления –
  + - var ( <name> := <value> … <name> := <value>)
    - var <name1>, <name2> = <value1>, <value2>

Таким образом, обработка объявления переменных стала более сложной и комплексной. Решить сложившуюся проблему получилось с помощью разбиения Definition на более мелкие варианты раскрытия, что помогло полностью покрыть все варианты данных конструкций.

## Примеры работы парсера

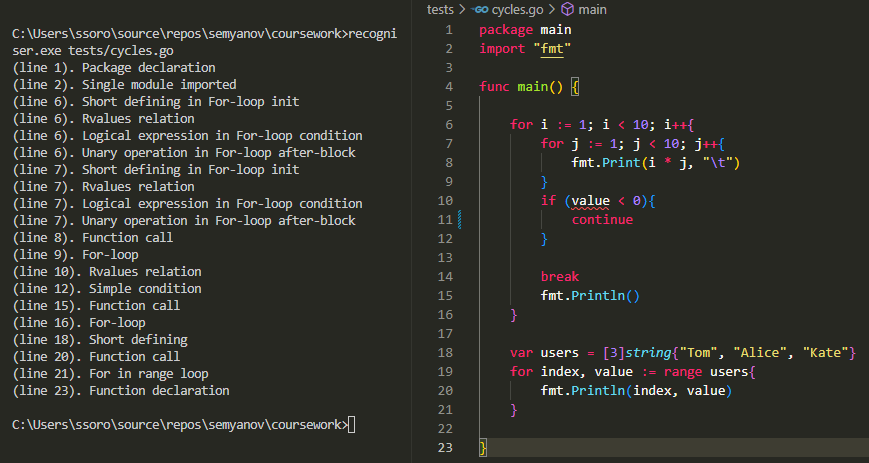


Рисунок 7 – Пример работы с цикловыми конструкциями.

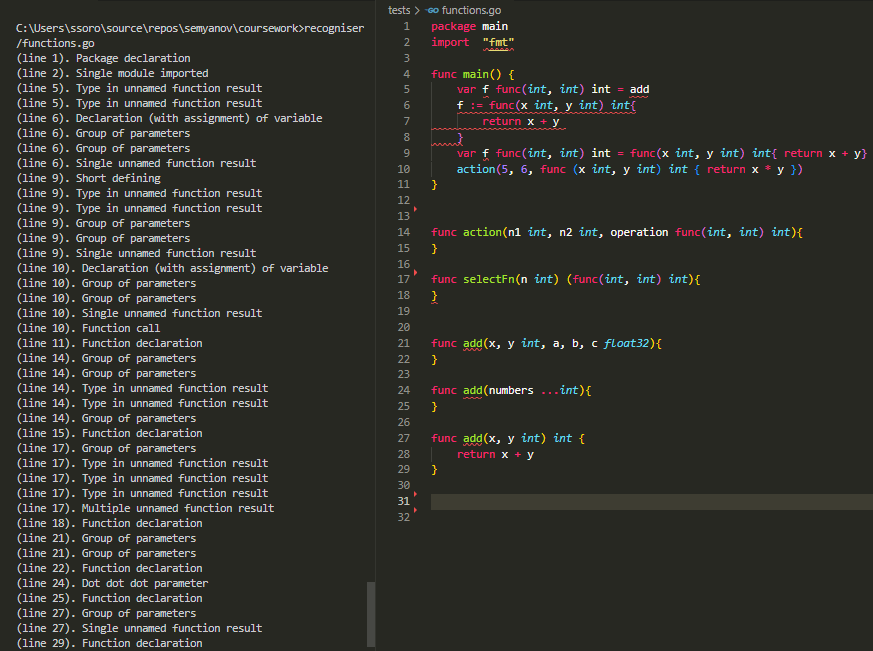


Рисунок 8 – Пример работы с функциональными конструкциями.

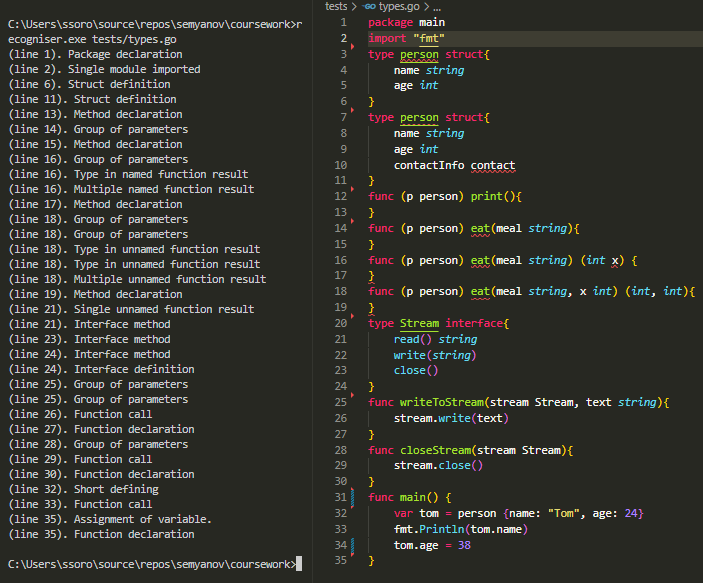


Рисунок 9 – Пример работы с типами, методами, интерфейсами.

# 2 РЕЗУЛЬТАТЫ

В данной курсовой работе были выполнены следующие задачи:

1. Разработан лексический анализатор языка программирования Golang.
2. Разрешено большинство конфликтов типа shift/reduce в разработанной грамматике.
3. С помощью возможностей bison была собрана программа-парсер, распознающая код на языке программирования Golang.

3 ВЫВОД

В ходе выполнения данной курсовой работы были получены практические навык разработки приложения – парсера существующего популярного языка программирования с использованием yacc и flex и написания грамматики для Golang, определяющего корректность написанного и предоставленного на вход программе кода.

Разработка грамматики для определения синтаксиса языка Go проводилась по принципу «от наименьшего наибольшему», то есть основные базовые идентификаторы собираются парсером в более крупные групповые системы, которые уже обрабатывались им. Главной идеей при построении грамматики было разделение элементов на правосторонние и левосторонние значения. Такое решение позволяет логично построить грамматику, построенную на присваивании левым значениям правых и определением их парсером как таковых.

Функционал данного парсера, однако не покрывает возможности обнаружения изменений в сложившейся грамматике Go, возможном использовании фреймворков и так далее. Следовательно, возможности разработанного парсера в дальнейшем может быть расширены с помощью добавления обработки новых элементов в грамматику.

Как и большинство современных языков программирования, Golang имеет косплексную и многоуровневую структуру синтаксиса, основная цель которой - удобство конечного пользователя, то есть программиста. Однако, при создании удобной среды разработки возникают проблемы в грамматике, из-за которых образуются конфликты (в основном вида сдвиг/свертка). Такие конфликты в основном являеются неразрешаемыми, однако парсеры справляются с ними с помощью анализа контекста, а значит они не представляют особых проблем для компиляторов.

Таким образом, исходя из большого количество возникших трудностей и нерешаемых конфликтов при разработке парсера можно сделать вывод о сложности устройства подобных комплексных языков программирования, а также о сложности разработки и структурирования их грамматических и лексических анализаторов.

Приложение 1 – ЛЕКСИЧЕСКИЙ АНАЛИЗАТОР

%{

#include <stdio.h>

#include <stdlib.h>

#include "cw.tab.h"

unsigned line\_number;

extern int yylex();

extern void yyerror(char \* msg);

FILE \*inputStream;

#define yyin inputStream

#define tmp\_scanf(f\_, ...) fscanf(yyin, (f\_), \_\_VA\_ARGS\_\_)

int yywrap()

{

return 1;

}

void print(char\* msg) {

printf("(line %d). %s\n", line\_number, msg);

}

int main(int argc, char\* argv[])

{

line\_number = 1;

inputStream = fopen(argv[1], "r");

if (inputStream == NULL)

{

printf("File wasn't found\n");

return -1;

}

yyparse();

fclose(inputStream);

}

%}

%x C\_COMMENT

%%

"/\*" { BEGIN(C\_COMMENT); }

<C\_COMMENT>"\*/" { BEGIN(INITIAL); }

<C\_COMMENT>\n { line\_number++; }

<C\_COMMENT>. { }

[/][/].\* { ; }

package { return PACKAGE\_KEYWORD; }

import { return IMPORT\_KEYWORD; }

func { return FUNC\_KEYWORD; }

if { return IF\_KEYWORD; }

else { return ELSE\_KEYWORD; }

switch { return SWITCH\_KEYWORD; }

case { return CASE\_KEYWORD; }

default { return DEFAULT\_KEYWORD; }

for { return FOR\_KEYWORD; }

break { return BREAK\_KEYWORD; }

continue { return CONTINUE\_KEYWORD; }

range { return RANGE\_KEYWORD; }

return { return RETURN\_KEYWORD; }

var { return VAR\_KEYWORD; }

int8 { return INT\_TYPE; }

int16 { return INT\_TYPE; }

int32 { return INT\_TYPE; }

int64 { return INT\_TYPE; }

uint8 { return INT\_TYPE; }

uint16 { return INT\_TYPE; }

uint32 { return INT\_TYPE; }

uint64 { return INT\_TYPE; }

byte { return INT\_TYPE; }

rune { return INT\_TYPE; }

int { return INT\_TYPE; }

uint { return INT\_TYPE; }

float32 { return FLOAT\_TYPE; }

float64 { return FLOAT\_TYPE; }

complex64 { return COMPLEX\_TYPE; }

complex128 { return COMPLEX\_TYPE; }

bool { return BOOL\_TYPE; }

string { return STRING\_TYPE; }

const { return CONST\_KEYWORD; }

type { return TYPE\_KEYWORD; }

struct { return STRUCT\_KEYWORD; }

interface { return INTERFACE\_KEYWORD; }

nil { return NIL\_KEYWORD; }

defer { return DEFER\_KEYWORD; }

"go" { return GO\_KEYWORD; }

">>" { return SHIFT\_RIGHT; }

"<<" { return SHIFT\_LEFT; }

[-]?[0-9]+ { return INTEGER; }

[-]?[0-9]+[\.][0-9]+ { return FLOAT; }

true { return TRUE\_FALSE; }

false { return TRUE\_FALSE; }

[-]?[0-9]+[[\+][0-9]+[i]]? { return COMPLEX; }

\"(\\.|[^"\\])\*\" { return STRING; }

:= return COLON\_EQ;

[A-Za-z\_]+[A-Za-z\_0-9]\* { return IDENTIFICATOR; }

"==" { return EQ\_RELATION; }

">" { return GREATER\_RELATION; }

"<" { return LESS\_RELATION; }

">=" { return EQ\_GREATER\_RELATION; }

"<=" { return EQ\_LESS\_RELATION; }

"!=" { return NOT\_EQ\_RELATION; }

"+=" { return OPER\_ASSIGNMENT; }

"-=" { return OPER\_ASSIGNMENT; }

"\*=" { return OPER\_ASSIGNMENT; }

"/=" { return OPER\_ASSIGNMENT; }

"%=" { return OPER\_ASSIGNMENT; }

">>=" { return OPER\_ASSIGNMENT; }

"<<=" { return OPER\_ASSIGNMENT; }

"|=" { return OPER\_ASSIGNMENT; }

"^=" { return OPER\_ASSIGNMENT; }

"++" { return INCREMENT; }

"--" { return DECREMENT; }

"!" { return NOT\_OPERATION; }

"&&" { return AND\_OPERATION; }

"||" { return OR\_OPERATION; }

"..." { return DOT\_DOT\_DOT; }

[-+()=\.,\{\}\:\[\]] { return yytext[0]; }

[\\*]\* {return '\*'; }

[ \t\f\v] { ; }

"\n" { line\_number++; return yytext[0]; }

"&" { return yytext[0]; }

"|" { return yytext[0]; }

"^" { return yytext[0]; }

";" { return yytext[0]; }

%%

Приложение 2 – синтаскический АНАЛИЗАТОР

%{

#include <stdio.h>

#include <stdlib.h>

int yylex();

//extern int yylex();

void yyerror(char \*msg);

extern void print(char \*msg);

%}

%union {

int num;

char letter;

}

%token IDENTIFICATOR COLON\_EQ DOT\_DOT\_DOT NIL\_KEYWORD DEFER\_KEYWORD SHIFT\_LEFT SHIFT\_RIGHT OPER\_ASSIGNMENT

%token STRING INTEGER FLOAT TRUE\_FALSE COMPLEX

%token TYPE\_KEYWORD STRUCT\_KEYWORD INTERFACE\_KEYWORD

%token CONST\_KEYWORD PACKAGE\_KEYWORD IMPORT\_KEYWORD VAR\_KEYWORD FUNC\_KEYWORD RETURN\_KEYWORD

%token IF\_KEYWORD ELSE\_KEYWORD SWITCH\_KEYWORD CASE\_KEYWORD DEFAULT\_KEYWORD

%token EQ\_RELATION GREATER\_RELATION LESS\_RELATION EQ\_GREATER\_RELATION EQ\_LESS\_RELATION NOT\_EQ\_RELATION

%token FOR\_KEYWORD BREAK\_KEYWORD CONTINUE\_KEYWORD RANGE\_KEYWORD GO\_KEYWORD

%token INT\_TYPE FLOAT\_TYPE COMPLEX\_TYPE BOOL\_TYPE STRING\_TYPE

%left OR\_OPERATION

%left AND\_OPERATION

%left NOT\_OPERATION

%left '-' '+' '%'

%left '\*' '/' '|'

%left '^' '&'

%left SHIFT\_LEFT SHIFT\_RIGHT

%nonassoc INCREMENT DECREMENT

%%

S : PACKAGE '\n' GLOBALS ;

PACKAGE :

PACKAGE\_KEYWORD IDENTIFICATOR { print("Package declaration"); } ;

IMPORT : IMPORT\_KEYWORD STRING { print("Single module imported"); }

| IMPORT\_KEYWORD '(' '\n' IMPORT\_MULTIPLE\_STRING ')' { print("Importing modules with brackets"); }

;

IMPORT\_MULTIPLE\_STRING : STRING '\n' IMPORT\_MULTIPLE\_STRING { print("Module name"); }

|

;

GLOBALS :

GLOBAL

| GLOBALS '\n' GLOBAL

;

GLOBAL:

FUNCTION

| IMPORT

| TYPEDEF

| DECLARATION

|

;

TYPEDEF:

TYPE\_KEYWORD TYPEVAL TYPEVAL { print("Extra type definition"); }// QUESTION

| TYPE\_KEYWORD IDENTIFICATOR STRUCT\_KEYWORD '{' STRUCT\_FIELDS '}' { print("Struct definition"); }

| TYPE\_KEYWORD IDENTIFICATOR INTERFACE\_KEYWORD '{' INTERFACE\_FIELDS '}' { print("Interface definition"); }

;

INTERFACE\_FIELDS:

INTERFACE\_FIELDS INTERFACE\_FIELD '\n'

|

;

INTERFACE\_FIELD:

IDENTIFICATOR '(' INTERFACE\_METHOD\_ARGS')' FUNC\_RESULT { print("Interface method "); }

| IDENTIFICATOR { print("Inner interface "); }

|

;

INTERFACE\_METHOD\_ARGS:

TYPEVAL

| INTERFACE\_METHOD\_ARGS ',' TYPEVAL

|

;

STRUCT\_FIELDS:

STRUCT\_FIELDS STRUCT\_FIELD '\n'

|

;

STRUCT\_FIELD:

MULTIPLE\_IDENT TYPEVAL

|

;

FUNCTION : FUNC\_KEYWORD IDENTIFICATOR '(' FUNC\_PARAMS ')' FUNC\_RESULT '{' FUNCTION\_STATEMENTS '}' { print("Function declaration"); }

| FUNC\_KEYWORD '(' TYPEVAL TYPEVAL ')' IDENTIFICATOR '(' FUNC\_PARAMS ')' FUNC\_RESULT '{' FUNCTION\_STATEMENTS '}' { print("Method declaration"); }//QUESTION ?!?!??!?!??!

;

ANON\_FUNCTION:

FUNC\_KEYWORD '(' FUNC\_PARAMS ')' FUNC\_RESULT '{' FUNCTION\_STATEMENTS '}'

;

FUNCTION\_STATEMENTS:

STATEMENTS

|

;

FUNC\_PARAMETER\_GROUP :

MULTIPLE\_IDENT TYPEVAL { print("Group of parameters"); }

| IDENTIFICATOR DOT\_DOT\_DOT TYPEVAL { print("Dot dot dot parameter"); }

;

FUNC\_PARAMS: FUNC\_PARAMETER\_GROUPS

|

;

FUNC\_PARAMETER\_GROUPS : FUNC\_PARAMETER\_GROUP

| FUNC\_PARAMETER\_GROUPS ',' FUNC\_PARAMETER\_GROUP

;

FUNC\_RESULT :

TYPEVAL { print("Single unnamed function result"); }

| '(' FUNC\_RESULT\_NAMED ')' { print("Multiple named function result"); }

| '(' FUNC\_RESULT\_UNNAMED ')' { print("Multiple unnamed function result"); }

|

;

FUNC\_RESULT\_UNNAMED :

TYPEVAL { print("Type in unnamed function result"); }

| FUNC\_RESULT\_UNNAMED ',' TYPEVAL { print("Type in unnamed function result"); }

|

;

FUNC\_RESULT\_NAMED :

TYPEVAL TYPEVAL { print("Type in named function result"); }//QUESTION

| FUNC\_RESULT\_NAMED ',' TYPEVAL TYPEVAL { print("Type in named function result"); }//QUESTION

;

STATEMENTS : STATEMENT

| STATEMENTS STATEMENT

;

STATEMENT : DECLARATION

| IF\_ELSE\_STATEMENT

| FUNCTION\_CALL

| SWITCH

| ASSIGNMENT

| UNARY\_OPERATION

| FOR

| BREAK\_KEYWORD

| CONTINUE\_KEYWORD

| RETURN

| '\n'

| DEFER\_KEYWORD FUNCTION\_CALL { print("Defer function call");}

| GO\_KEYWORD FUNCTION\_CALL { print("Starting goroutine"); }

;

RETURN :

RETURN\_KEYWORD

| RETURN\_KEYWORD MULTIPLE\_RVALUE //REDUCE/REDUCE

;

UNARY\_OPERATION :

IDENTIFICATOR INCREMENT

| IDENTIFICATOR DECREMENT

| INCREMENT IDENTIFICATOR

| DECREMENT IDENTIFICATOR

;

ASSIGNMENT :

FULL\_IDENTIFICATOR '=' RVALUE { print("Assignment of variable."); }

| FULL\_IDENTIFICATOR ARRAY\_INDEXATION '=' RVALUE { print("Assigment of array element. "); }

| '\*' FULL\_IDENTIFICATOR '=' RVALUE { print("Assigment of pointer by address. "); } //

| FULL\_IDENTIFICATOR OPER\_ASSIGNMENT RVALUE { print("Operation + assignment");}

;

FOR :

FOR\_KEYWORD FOR\_INIT ';' FOR\_CONDITION ';' FOR\_AFTER '{' STATEMENTS '}' { print("For-loop"); }

| FOR\_KEYWORD FOR\_CONDITION '{' STATEMENTS '}' { print("Shortened for-loop"); }

| FOR\_KEYWORD IDENTIFICATOR ',' IDENTIFICATOR COLON\_EQ RANGE\_KEYWORD IDENTIFICATOR '{' STATEMENTS '}' { print("For in range loop"); }

;

FOR\_INIT : SHORT\_DEFINING { print("Short defining in For-loop init "); }

| ASSIGNMENT { print("Assigment in For-loop init "); }

| { print("Emptyness in For-loop init "); }

;

FOR\_CONDITION :

LOGICAL\_EXPRESSION { print("Logical expression in For-loop condition "); }

| { print("Empty logical expression in For-loop condition "); }

;

FOR\_AFTER :

ASSIGNMENT { print("Assignment in For-loop after-block "); }

| FUNCTION\_CALL { print("Function call in For-loop after-block "); }

| UNARY\_OPERATION { print("Unary operation in For-loop after-block "); }

| { print("Emptyness in For-loop after-block "); }

;

DECLARATION:

VAR\_KEYWORD VARIABLE\_DECLARATION { print("Simple variable declaration"); }

| VAR\_KEYWORD VARIABLE\_DECLARATION\_ASSIGNMENT { print("Declaration (with assignment) of variable"); }

| VAR\_KEYWORD '(' MULTIPLE\_VARIABLE\_DECLARATION ')' { print("Multiple declaration"); }

| CONST\_KEYWORD VARIABLE\_DECLARATION\_ASSIGNMENT { print("Declaration (with assignment) of constant "); }

| SHORT\_DEFINING { print("Short defining"); }

;

SWITCH :

SWITCH\_KEYWORD '(' RVALUE ')' '{' '\n' SWITCH\_CASES '}'

| SWITCH\_KEYWORD '(' RVALUE ')' '{' '\n' SWITCH\_CASES DEFAULT\_KEYWORD ':' STATEMENTS '}' { print("Default case in switch"); }

;

SWITCH\_VALUES:

SWITCH\_VALUES ',' RVALUE

| RVALUE

;

SWITCH\_CASES:

CASE\_KEYWORD SWITCH\_VALUES ':' STATEMENTS { print("Case in switch"); }

| SWITCH\_CASES CASE\_KEYWORD SWITCH\_VALUES ':' STATEMENTS { print("Case in switch"); }

;

IF\_ELSE\_STATEMENT:

IF\_ELSE\_IF

| IF\_ELSE\_IF ELSE\_KEYWORD '{' STATEMENTS '}' { print("Else condition"); }

;

IF\_ELSE\_IF:

CONDITION { print("Simple condition"); }

| IF\_ELSE\_IF ELSE\_KEYWORD CONDITION { print("Else-if condition"); }

;

CONDITION : IF\_KEYWORD LOGICAL\_EXPRESSION '{' STATEMENTS '}'

| IF\_KEYWORD LOGICAL\_EXPRESSION '{' '}'

;

LOGICAL\_EXPRESSION :

RVALUE RELATION RVALUE { print("Rvalues relation"); }

| NOT\_OPERATION LOGICAL\_EXPRESSION { print("Denying expression"); }

| TRUE\_FALSE { print("True/False relation operand"); }

| FUNCTION\_CALL { print("Function call in relation"); }

| LOGICAL\_EXPRESSION AND\_OPERATION LOGICAL\_EXPRESSION { print("Logical AND"); }

| LOGICAL\_EXPRESSION OR\_OPERATION LOGICAL\_EXPRESSION { print("Logincal OR "); }

| '(' LOGICAL\_EXPRESSION ')'

;

RELATION : EQ\_RELATION

| GREATER\_RELATION

| LESS\_RELATION

| EQ\_GREATER\_RELATION

| EQ\_LESS\_RELATION

| NOT\_EQ\_RELATION

;

RVALUE :

VALUE

| VALUE '+' RVALUE

| VALUE '-' RVALUE

| VALUE '\*' RVALUE

| VALUE '/' RVALUE

| VALUE '%' RVALUE

| VALUE '|' RVALUE

| VALUE '^' RVALUE

| VALUE '&' RVALUE

| VALUE SHIFT\_LEFT RVALUE

| VALUE SHIFT\_RIGHT RVALUE

| TYPE '{' INITIALIZER '}'

| IDENTIFICATOR '{' INITIALIZER '}'

| TYPE '{' FUNCTION\_CALL\_ARGUMENTS '}'

| IDENTIFICATOR '{' FUNCTION\_CALL\_ARGUMENTS '}'

| ANON\_FUNCTION

;

VALUE:

INTEGER

| STRING

| FLOAT

| TRUE\_FALSE

| COMPLEX

| NIL\_KEYWORD

| FUNCTION\_CALL

| FULL\_IDENTIFICATOR

| FULL\_IDENTIFICATOR ARRAY\_INDEXATION

| '&' FULL\_IDENTIFICATOR { print("Getting address of identificator");}

| '\*' FULL\_IDENTIFICATOR

;

INITIALIZER:

IDENTIFICATOR ':' VALUE

| INITIALIZER ',' IDENTIFICATOR ':' VALUE

;

ARRAY\_INDEXATION:

'[' VALUE ']'

| ARRAY\_INDEXATION '[' VALUE ']'

;

FUNCTION\_CALL :

FULL\_IDENTIFICATOR '(' FUNCTION\_CALL\_ARGUMENTS ')' { print("Function call"); }

| ANON\_FUNCTION '(' FUNCTION\_CALL\_ARGUMENTS ')' { print("Anon function call"); }

;

FULL\_IDENTIFICATOR:

IDENTIFICATOR

| FULL\_IDENTIFICATOR '.' IDENTIFICATOR

;

FUNCTION\_CALL\_ARGUMENTS :

FUNCTION\_CALL\_ARGUMENTS ',' RVALUE

| RVALUE

|

;

SHORT\_DEFINING:

IDENTIFICATOR COLON\_EQ RVALUE

| VAR\_KEYWORD IDENTIFICATOR '=' RVALUE

;

MULTIPLE\_VARIABLE\_DECLARATION:

VARIABLE\_DECLARATION

| VARIABLE\_DECLARATION\_ASSIGNMENT

| IDENTIFICATOR '=' MULTIPLE\_RVALUE

| MULTIPLE\_VARIABLE\_DECLARATION VARIABLE\_DECLARATION

| MULTIPLE\_VARIABLE\_DECLARATION VARIABLE\_DECLARATION\_ASSIGNMENT

| MULTIPLE\_VARIABLE\_DECLARATION IDENTIFICATOR '=' MULTIPLE\_RVALUE

| MULTIPLE\_VARIABLE\_DECLARATION '\n'

| '\n'

;

VARIABLE\_DECLARATION:

MULTIPLE\_IDENT TYPEVAL

;

VARIABLE\_DECLARATION\_ASSIGNMENT:

MULTIPLE\_IDENT TYPEVAL '=' MULTIPLE\_RVALUE

;

MULTIPLE\_RVALUE:

MULTIPLE\_RVALUE ',' RVALUE

| RVALUE

;

TYPEVAL:

TYPE

| IDENTIFICATOR

;

TYPE:

INT\_TYPE

| FLOAT\_TYPE

| COMPLEX\_TYPE

| BOOL\_TYPE

| STRING\_TYPE

| FUNC\_KEYWORD '(' FUNC\_RESULT\_UNNAMED ')' TYPE

| '[' INTEGER ']' TYPE

| '[' DOT\_DOT\_DOT ']' TYPE

| '[' ']' TYPE

| '\*' TYPE //Создает конфликт свептка/свертка с ???

;

MULTIPLE\_IDENT:

IDENTIFICATOR

| MULTIPLE\_IDENT ',' IDENTIFICATOR

;

%%

void yyerror(char \* msg) {

fprintf(stderr, "%s", msg);

exit(1);

}