Project 2: Yacc

1 Overview

In this assignment you will use the parser generator yacc to construct an interpreter for a language called Snail containing the set of statements listed below. The interpreter executes the statements in a Snail program, one after the other in the order which they appear in the Snail program.

In Section 2 we describe the various statements. In Section 3 we give the grammar of the Simple language. In Section 4 we describe what your yacc code will do. In Section 5 are instructions to hand-in your assignment.

2 The Snail Programming Language

Snail is a very simple programming language. The body of a Snail program consists of a sequence of statements. There are three kinds of statements: assign, print, and if. A basic component for all kinds of statements is the expression. The expression and the statements are described below.

• expression

An expression is any mathematical expression made from identifiers, integers, parenthesis, the arithmetic operators

+  -  *  /

and the comparison operators

<  >  <=  >=  ==  !=

For example, this is a valid expression:

10 + 20 * (10 < 3)

The value of an expression is obtained by executing all the arithmetic operations in the expression. The result of a comparison operation is 1 if the comparison result is true, and 0 otherwise. For example, the above expression has value 10 (since, 10 < 3 = 0).

The value of an identifier is the last value assigned to it in an assign statement. An identifier which hasn’t been assigned a value before cannot be used inside an expression and in this case you should report an error message.

• assign statement

The assign statement has the form:

identifier = expression;
For example, this is a valid assign statement:

\[ \text{var1} = 20 - 3 \times 2; \]

In the assign statement the identifier gets the value of the expression. As an example, in the above assign statement the new value of variable \text{var1} is 14.

• **print statement**

The print statement print messages on the screen. The print statement has one of following forms:

\[ \text{print} \, \text{string}; \quad \text{//prints the string} \]
\[ \text{print} \, \text{newline}; \quad \text{//prints a newline character} \]
\[ \text{print} \, \text{expression}; \quad \text{//prints the expression value} \]

For example, the execution of the following statements

\[ \text{print} \, \text{'The value of } 10 \times 5 \text{ is '}; \]
\[ \text{print} \, 10 \times 5; \]

produces the output:

The value of 10*5 is 50

• **if statement**

An if statement has two forms:

\[ \text{if} \, \text{expression then } \quad \text{//if-then statement} \]
\[ \text{statement} \]
\[ \text{statement} \]
\[ \text{...} \quad \text{//more statements} \]
\[ \text{endif} \]

\[ \text{if} \, \text{expression then } \quad \text{//if-then-else statement} \]
\[ \text{statement} \]
\[ \text{...} \quad \text{//more statements} \]
\[ \text{else} \]
\[ \text{statement} \]
\[ \text{...} \quad \text{//more statements} \]
\[ \text{endif} \]

The “if-then” statement means that if the expression value is not 0 then the sequence of statements between \text{then} and \text{endif} will be executed. The “if-then-else” statement means that if the expression value is not 0 then the statements between \text{then} and \text{else} will be executed, and otherwise, if the expression value is 0, the statements between \text{else} and \text{endif} will be executed. For example, the following is a valid if statement:
if (x < 10) then
    print ‘‘x is smaller than ten’’;
    x = x - y + 20;
else
    x = 10* y;
endif

A Snail program is a sequence of statements and has the following general form:

statement
statement
...
statement

We can have comments in a Snail program right after “//” (as in a C++ program). An simple example Snail program is the following:

x = 10;
y = 20;
print ‘‘the value of x is ’’; print x; print newline;
print ‘‘the value of y is ’’; print y; print newline;
print ‘‘the sum of x and y is ’’; print x + y; print newline;

if x < y then //test x < y
    print ‘‘x is smaller than y’’; // x is smaller
else
    print ‘‘x is bigger or equal than y’’ // x is not smaller
endif

The output of the program is:

the value of x is 10
the value of y is 20
the sum of x and y is 30;
x is smaller than y;

3 Snail Grammar

All the Snail programs can be described by the context-free grammar given below. The start variable is program, the grammar variables are in small letters, and the terminals in capital letters. Notice that this grammar is ambiguous in the expr variable; however, all the ambiguities can be removed using the precedence rules of yacc.

program -> stmt_list

stmt_list -> stmt_list stmt

expr -> term
expr -> expr + term
expr -> expr - term
expr -> expr * term
expr -> expr / term

term -> factor
term -> term * factor
term -> term / factor

factor -> num
factor -> ( expr )

num -> 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

3
4 Yacc Code

You will write a yacc code which implements the interpreter for Snail programs. The main part of your yacc code will consist of the snail grammar augmented with actions.

- **expression**
  Whenever you find an expression, you need to evaluate the expression. This can be done in a bottom up fashion with actions such as:

  \[
  \text{expr : expr } '+\text{'} \text{ expr} \{\$\$ = \$1 + \$3;\}
  \]

- **assign statement**
  The value of an ID (identifier) node is stored in the symbol table. For an assign node you need to update the value of the variable to be equal to the expr of the right hand side of the assign statement.
• **print statement**

For a `print` node you just print on the output the contents of the STRING stripped from quotes, or the value of the `expr`, or a newline character.

• **if statement**

For the if statement you first evaluate the `expr`. If the value of `expr` is true (not zero), then you execute the if part of the statement, otherwise you execute the else part of the statement. You need to be careful not to execute the part of the if statement you don’t want to execute.

In the following example, when expression is true then only statement1 will be executed, while when expression is false then only statement2 will be executed.

```plaintext
if expression then
  statement1
else
  statement2
endif
```

In order to achieve this, you need to have a stack which will control whether or not to execute particular statements. We push true (1) for the part of the if which will be executed; for the other part we push false (0). The stack is needed in order to handle nested if statements. The code to handle this is as follows:

```plaintext
if_stmt: IF expr
  THEN {top()==1 ? push($2!=0); : push(0);} stmt_list {pop();}
  ELSE {top()==1 ? push($2==0); : push(0);} stmt_list {pop();} ENDIF
```

If the top of the stack contains true then statements can be executed. For example, the actions of the assignment statement:

```plaintext
assign_stmt -> ID = expr ; {if (top() == 1) then {execute the assignment}}
```

In order to handle statements and expressions in the main body, outside from if statements, you need to push true at the beginning of the execution of the interpreter.

If you find a syntax error in the input program then report an error message with the line number where you found the error and abort the program.

Your yacc program will use parts of the lexical analyzer you built in the first project and for this you need to modify appropriately the lex code.

### 5 Hand-In

You should hand-in in paper your lex and yacc code. Also you should hand-in the output of your interpreter for Snail test programs which are given in the course web page.

In the course web page you can also find example yacc programs (together with lex programs) that will help you to get started with your project.