SetlCup Tutorial

Jonas Eilers

January 4, 2016

Contents

Chapter 1

Functionality

The Setlx-addition SetlCup is a LR-Parser-Generator based on JavaCup. The idea is to use a user given scanner- and parser-definition and create an AST out of a given input using the definitions.

In this document the needed syntax of the definitions is examined and the given output is evaluated.

A sample input file is divided into three Sections:

1. Commentpart
2. Scanner-Part
3. Parser-Part

At first the correct call of the program is discussed.

1.1 Using SetlCup

SetlCup has three different variants in which it can be called:

1. `setlx setlcup.stlx -p parser_scanner_file.stlx file_to_be_read.txt`

   With this call there will be no output for the user.

2. `setlx setlcup.stlx -p parser_scanner_file.stlx file_to_be_read.txt -d`

   With this call debugging is possible. It shows the different tables and states and the whole parsing progress. HINT: It is recommended to pipe the output into a file if you are using the "-d" option.

3. `setlx setlcup.stlx -h`

   With this call a little help will be showed, on how to call SetlCup correctly.

1.2 Comment-Part

In the comment-part everything which is written will not be used by the Program itself. It is adviced to comment your idea behind the parser and scanner structure in this section. The section is ended with the "%%%" symbol.

1.3 Scanner-Part

The scanner is responsible for checking whether the input file consists of the defined tokens. It can be written like this:

```
1 INTEGER := 1-9[0-9]*|0;
2 ASTERISK := \*;
3 WHITESPACE := [ ];
4 SKIP := ASTERISK | INTEGER | WHITESPACE;
```

1. In line 1 the Token "INTEGER" is defined. Tokens are in the following way:
   token_name := regex ;
2. Predefined tokens in Regular Expressions like "$*, +, ?, |, \{, \}, (,), \cdots$" need to be escaped.
3. In some contexts tokens like Whitespaces are not needed. They can be skipped by using defining the "SKIP"-Token with the tokens, which shall be skipped. Multiple tokens need to be seperated by a pipe "|".

1.4 Parser-Part

In this part the grammar-rules are defined with the following syntax:

```
1 rule_name := rule_element:id {: action_code :}
2          | rule_element
3          | {: action_code :}
4          | ;
```

rule_name The rule_name is the name of the rule. It is possible to reference defined rules via their rule_name

rule_element The element can consist of multiple Tokens (defined in the scanner) and rule_names. Each can have an id, which is possible to be used in the action_code.

action_code The action_code is an optional part in a rule. It needs to be at the end of the rule it self. Each rule_element can have an action_code. In this action_code Setlx Code can be written. By using the variable "‘result"’ it is possible to pass values between rules. The id of the elements in the respective rule can be referred to by using its name.

| The pipe seperates the rule_elements.

1.5 Example

```
1
2
3  %%%
4
5  SEMICOLON := ; ;
6  TIMES := \* ;
7  MINUS    := - ;
8  DIVIDE    := \\ ;
9  INTEGER      := 0|[1-9][0-9]* ;
10 NEWLINE := \n ;
11 WHITESPACE := [ \t\v\n\r\s] ;
12 MOD := %;
13 PLUS  := \+ ;
14 LPAREN := \( ;
15 RPAREN := \) ;
16 SKIP := WHITESPACE | NEWLINE ;
17
18 %%%
19
20 expr_list ::= expr_list:l expr_part:part {: result := l + [part]; :}
21             | expr_part:epart {: result := [epart]; :}
22             ;
23 expr_part ::= expr:e SEMICOLON {: result := e; :}
24             ;
25 expr ::= expr:e PLUS   prod:p {: result := Plus(e , p); :}
26        | expr:e MINUS  prod:p {: result := Minus(e , p); :}
27        | prod:p               {: result := p;      :}
28        ;
29 prod ::= prod:p TIMES  fact:f {: result := Times(p , f); :}
30        | prod:p DIVIDE fact:f {: result := Div(p , f); :}
31        | prod:p MOD    fact:f {: result := Mod(p , f); :}
32        | fact:f               {: result := f;      :}
33        ;
34 fact ::= LPAREN expr:e_part RPAREN {: result :=  e_part ;   :}
35        | INTEGER:n             {: result := Integer(eval(n)); :}
36        ;
```

```
1 1 + 2 * 3 - 4;
2 1 + 2 + 3 + 4;
3 1 + ( 2 * 3 ) * 5 % 6;
```

```
1 ExprList([Minus(Plus(Integer(1), Times(Integer(2), Integer(3))),
      Integer(4)), Plus(Plus(Plus(Integer(1), Integer(2)), Integer(3)
      ), Integer(4)), Plus(Integer(1), Mod(Times(Times(Integer(2),
      Integer(3)), Integer(5)), Integer(6)))])
```

```
1
2  %%%
3
4  SEMI := ; ;
5  TIMES := \* ;
6  MINUS    := - ;
7  DIV    := \\ ;
8  MOD := %;
9  PLUS  := \+ ;
10 LPAR := \( ;
11 RPAR := \) ;
12 LBRACE := \{ ;
```

```
13 RBRACE := \} ;
14 COMMA := , ;
15 ASSIGN := = ;
16 EQ := == ;
17 NE := != ;
18 LT := < ;
19 GT := > ;
20 LE := <= ;
21 GE := >= ;
22 AND := && ;
23 OR := \|\| ;
24 NOT := ! ;
25 FUNCTION := function ;
26 RETURN := return ;
27 IF := if ;
28 ELSE := else ;
29 WHILE := while ;
30 FOR := for ;
31 PRINT := print ;
32 QUIT := exit ;
33 STRING := \"(?:\\.|[^\"])*\" ;
34 NEWLINE := \n ;
35 COMMENTS := //[^\n]* ;
36 WHITESPACE := [ \t\v\n\r\s] ;
37 SKIP := WHITESPACE | NEWLINE | COMMENTS ;
38 INTEGER := 0|[1-9][0-9]* ;
39 DECIMAL := 0\.[0-9]+|[1-9][0-9]*\.[0-9]+ ;
40 ZID := [a-zA-Z_][a-zA-Z0-9_]* ;
41
42
43 %%%
44 program ::= dfnStmntList:d {: result := Program(d); :};
45
46 dfnStmntList
47     ::= definition:d dfnStmntList:dl {: result := [d] + dl; :}
48       | statement:stmts  dfnStmntList:dsl {: result := [stmts] +
              dsl; :}
49       | {: result := []; :}
50       ;
51
52 definition ::= FUNCTION ZID:function_name LPAR paramList:param_list
       RPAR LBRACE stmntList:statement_list RBRACE
53         {: result := Function(function_name, param_list,
               statement_list);:}
54       ;
55
56 stmntList
57     ::= statement:s stmntList:sl {: result := [s] + sl ; :}
58       | {: result := []; :}
59       ;
60
61 statement
62     ::= assignment:a SEMI {: result := Ass(a); :}
63       | PRINT LPAR printExprList:printexpr_list RPAR SEMI        {:
              result := Print(printexpr_list); :}
64       | IF LPAR boolExpr:b RPAR LBRACE stmntList:st_list1 RBRACE
                    {: result := If(b, st_list1); :}
65       | WHILE LPAR boolExpr:b RPAR LBRACE stmntList:st_list2 RBRACE
                  {: result := While(b, st_list2); :}
66       | FOR LPAR assignment:i_a SEMI boolExpr:b SEMI assignment:e_a
            RPAR LBRACE stmntList:st_list3 RBRACE {: result := For(
           i_a, b, e_a, st_list3);  :}
```

```
67      |  RETURN expr:e SEMI {: result := Return(e); :}
68      |  RETURN SEMI {: result := Return(); :}
69      |  expr:e SEMI {: result := Expr(e); :}
70      |  QUIT SEMI {: result := Exit(); :}
71      ;
72
73  printExprList
74      ::= printExpr:p COMMA nePrintExprList:np {: result := [p] + np
            ; :}
75      |  printExpr:p {: result := [p]; :}
76      |  {: result := []; :}
77      ;
78
79  nePrintExprList
80      ::= printExpr:p {: result := [p]; :}
81      |  printExpr:p COMMA nePrintExprList:np {: result := [p] + np
             ; :}
82      ;
83
84  printExpr
85      ::= STRING:string {: result := PrintString(string); :}
86      |  expr:e  {: result := Expr(e); :}
87      ;
88
89  assignment
90      ::= ZID:id ASSIGN expr:e {: result := Assign(id, e); :}
91      ;
92
93  paramList
94      ::= ZID:id COMMA neIDList:nid {: result := [id] + nid ; :}
95      |  ZID:id {: result := [id] ; :}
96      |  {: result := []; :}
97      ;
98
99  neIDList
100     ::= ZID:id COMMA neIDList:nid {: result := [id] + nid ; :}
101     |  ZID:id  {: result := [id] ; :}
102     ;
103
104
105 boolExpr
106     ::= expr:lhs EQ expr:rhs  {: result := Equation(lhs,rhs); :}
107     |  expr:lhs NE expr:rhs  {: result := Inequation(lhs,rhs); :}
108     |  disjunction:lhs EQ disjunction:rhs  {: result := Equation(
             lhs,rhs); :}
109     |  disjunction:lhs NE disjunction:rhs  {: result := Inequation
             (lhs,rhs); :}
110     |  expr:lhs LE expr:rhs  {: result := LessOrEqual(lhs,rhs); :}
111     |  expr:lhs GE expr:rhs  {: result := GreaterOrEqual(lhs,rhs);
             :}
112     |  expr:lhs LT expr:rhs  {: result := LessThan(lhs,rhs); :}
113     |  expr:lhs GT expr:rhs  {: result := GreaterThan(lhs,rhs); :}
114     |  disjunction:d {: result := d; :}
115     ;
116 disjunction
117     ::= disjunction:d OR conjunction:c {: result := Disjunction(d,c
            ); :}
118     |  conjunction:c {: result := c; :}
119     ;
120 conjunction
121     ::= conjunction:c AND boolFactor:f {:result := Conjunction(c,f)
            ; :}
```

```
122       | boolFactor:f {: result := f; :}
123       ;
124 boolFactor
125     ::= LPAR boolExpr:be_par RPAR {:  result := be_par; :}
126       | NOT boolExpr:e {: result := Negation(e); :}
127       ;
128
129
130 expr ::= expr:e PLUS   prod:p {: result := Sum(e,p); :}
131        | expr:e MINUS  prod:p {: result := Difference(e,p); :}
132        | prod:p                {: result := p;      :}
133        ;
134 prod ::= prod:p TIMES  fact:f {: result := Product(p,f); :}
135        | prod:p DIV fact:f {: result := Quotient(p,f); :}
136        | prod:p MOD    fact:f {: result := Mod(p,f); :}
137        | fact:f                {: result := f;      :}
138        ;
139 fact ::= LPAR expr:e_par RPAR {: result := e_par;   :}
140        | INTEGER:n             {: result := Integer(eval(n));   :}
141        | DECIMAL:d               {: result := Decimal(eval(d)); :}
142        | ZID:id_1 LPAR exprList:el RPAR {: result := FunctionCall(
               id_1,el); :}
143        | ZID:id_2 {: result := Variable(id_2); :}
144        ;
145
146 exprList
147     ::= expr:e COMMA neExprList:el {: result := [e] + el; :}
148       | expr:e {: result := [e]; :}
149       | {: result := []; :}
150       ;
151
152 neExprList
153     ::= expr:e COMMA neExprList:el {: result := [e] + el; :}
154       | expr:e {: result := [e]; :}
155       ;
```

```
1  function factorial(n) {
2      if (n == 0) {
3                  return 1;
4      }
5      return n * factorial(n - 1);
6  }
7
8  print("Berechnung der Fakultät für i = 1 bis 9");
9  for (i = 0; i < 10; i = i + 1) {
10     print(i, "! = ", factorial(i));
11 }
12 print();
```

```
1  Program([Function("factorial", ["n"], [If(Equation(Variable("n"),
       Integer(0)), [Return(Integer(1))]), Return(Product(Variable("n
       "), FunctionCall("factorial", [Difference(Variable("n"),
       Integer(1))]))])]), Print([PrintString("Berechnung der Fakultaet
        fuer i = 1 bis 9")]), For(Assign("i", Integer(0)), LessThan(
       Variable("i"), Integer(10)), Assign("i", Sum(Variable("i"),
       Integer(1))), [Print([Expr(Variable("i")), PrintString("! = "),
        Expr(FunctionCall("factorial", [Variable("i")]))])]), Print
       ([])])
```