CS31003: Compilers: Yacc

Partha Pratim Das

Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

ppd@cse.iitkgp.ernet.in

August 7, 11, & 13 2014

A Simple Calculator Grammar

$$
\begin{array}{rlcl}
1: & S & \rightarrow & E \\
2: & E & \rightarrow & E + T \\
3: & E & \rightarrow & E - T \\
4: & E & \rightarrow & T \\
5: & T & \rightarrow & T * F \\
6: & T & \rightarrow & T / F \\
7: & T & \rightarrow & F \\
8: & F & \rightarrow & (E) \\
9: & F & \rightarrow & - E \\
10: & F & \rightarrow & \textbf{num}
\end{array}
$$

Yacc Specs (calc.y) for Calculator Grammar


```
%{ /* C Declarations and Definitions */
#include <string.h>
#include <iostream>
extern int yylex();
void yyerror(char *s);
%}

%union {
    int intval;
}

%token <intval> NUMBER

%type <intval> expression
%type <intval> term
%type <intval> factor

%%
statement: expression
                { printf("= %d\n", $1); }
         ;
expression: expression '+' term
                 { $$ = $1 + $3; }
          | expression '-' term
                 { $$ = $1 - $3; }
          | term
          ;
```

```
term: term '*' factor
          { $$ = $1 * $3; }
    | term '/' factor
          { if ($3 == 0)
              yyerror("divide by zero");
            else $$ = $1 / $3;
          }
    | factor
    ;
factor: '(' expression ')'
           { $$ = $2; }
      | '-' factor
           { $$ = -$2; }
      | NUMBER
      ;
%%

void yyerror(char *s) {
    std::cout << s << std::endl;
}

int main() {
    yyparse();
}
```

Note on Yacc Specs (calc.y)

- Terminal Symbols
  - Symbolized terminals (like `NUMBER`) are identified by `%token`. Usually, but not necessarily, these are multi-character.
  - Single character tokens (like `'+'`) may be specified in the rules simply with quotes.
- Non-Terminal Symbols
  - Non-Terminal symbols (like `expression`) are identified by `%type`.
  - Any symbol on the left-hand side of a rule is a non-terminal.
- Production Rules
  - Production rules are written with left-hand side non-terminal separated by a colon (`:`) from the right-hand side symbols.
  - Multiple rules are separated by alternate (`|`).
  - ϵ productions are marked by empty right-hand side.
  - Set of rules from a non-terminal is terminated by semicolon (`;`).
- Start Symbol
  - Non-terminal on the left-hand side of the first production rule is taken as the start symbol by default.
  - Start symbol may be explicitly defined by `%start`: `%start statement`.

Note on Yacc Specs (calc.y)

- Attributes
  - Every terminal and non-terminal has an (optional) attribute.
  - Multiple types of attributes are possible. They are bundled in a `C union` by `%union`.
  - An attribute is associated with a terminal by the `%token`: `%token <intval> NUMBER`
  - An attribute is associated with a non-terminal by the `%type`: `%type <intval> term`
- Actions
  - Every production rule has an action (`C` code snippet) at the end of the rule that fires when a reduction by the rule takes place.
  - In an action the attribute of the left-hand side non-terminal is identified as `$$` and the attributes of the symbols on the right-hand side are identified as `$1`, `$2`, `$3`, ... counting from left to right.
  - Missing actions for productions with single right-hand side symbol (like `factor` $\rightarrow$ `NUMBER`) imply a default action of copying the attribute (should be of compatible types) from the right to left: `{ $$ = $1 }` .

Header (y.tab.h) for Calculator


```
/* A Bison parser, made by GNU Bison 2.5.  */
/* Tokens.  */
#ifndef YYTOKENTYPE
# define YYTOKENTYPE
   /* Put the tokens into the symbol table, so that GDB and other debuggers
      know about them.  */
   enum yytokentype {
     NUMBER = 258
   };
#endif
/* Tokens.  */
#define NUMBER 258

#if ! defined YYSTYPE && ! defined YYSTYPE_IS_DECLARED
typedef union YYSTYPE
{
/* Line 2068 of yacc.c  */
#line 8 "calc.y"

int intval;

/* Line 2068 of yacc.c  */
#line 62 "y.tab.h"
} YYSTYPE;
# define YYSTYPE_IS_TRIVIAL 1
# define yystype YYSTYPE /* obsolescent; will be withdrawn */
# define YYSTYPE_IS_DECLARED 1
#endif

extern YYSTYPE yylval;
```

Note on Header (y.tab.h)

- `y.tab.h` is generated by Yacc from `calc.y` to specify the token constants and attribute type.
- `y.tab.h` is automatically included in `y.tab.c` and must be included in `calc.l` so that it can feature in `lex.yy.c`.
- Symbolized tokens are enumerated beyond 256 to avoid clash with ASCII codes returned for single character tokens.
- `%union` has generated a `C union YYSTYPE` .
- Line directives are used for cross references to source files. These help debug messaging. For example:

  ```
  #line 8 "calc.y"
  ```

- `yylval` is a pre-defined global variable of `YYSTYPE` type.

  ```
  extern YYSTYPE yylval;
  ```

  This is used by `lex.yy.c`.

Flex Specs (calc.l) for Calculator Grammar


```
%{
#include "y.tab.h"
#include <math.h>
%}

%%
[1-9]+[0-9]*    {
                    yylval.intval = atoi(yytext);
                    return NUMBER;
                }

[ \t]           ;  /* ignore white space */

"$"             {
                    return 0; /* end of input */
                }

\n|.            return yytext[0];
%%
```

Note on Flex Specs (calc.l)

- `y.tab.h` is automatically included in `y.tab.c` and must be included in `calc.l` so that it can feature in `lex.yy.c`.
- `yylval` is a pre-defined global variable of `YYSTYPE` type. So attributes of terminal symbols should be populated in it as appropriate. So for `NUMBER` we have:

  ```
  yylval.intval = atoi(yytext);
  ```

  Recall, in `calc.y`, we specified:

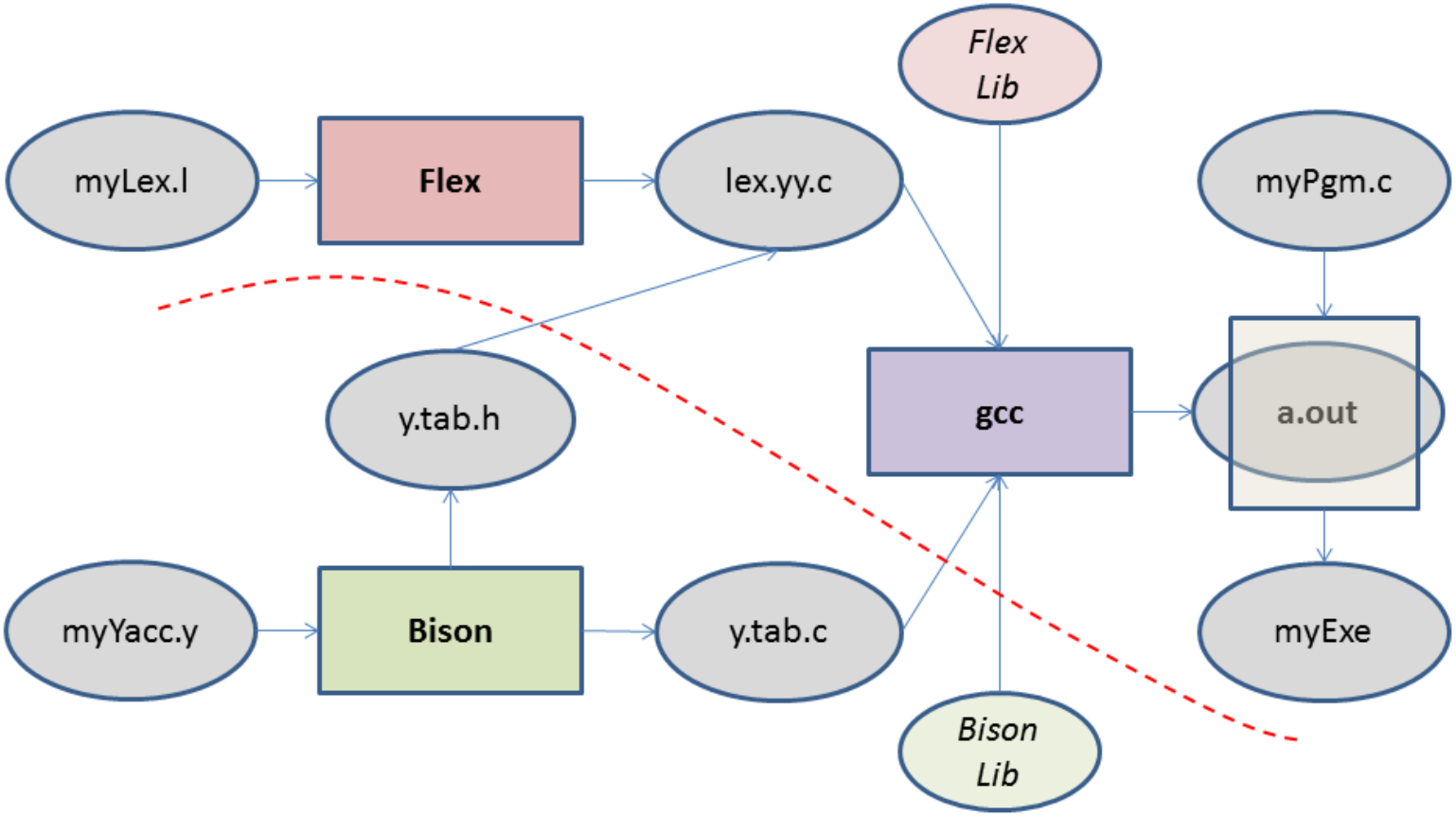
  ```
  %token <intval> NUMBER
  ```

  binding `intval` to `NUMBER`.
- Note how

  ```
  \n|.            return yytext[0];
  ```

  would return single character operators by their ASCII code.
- Newline is not treated as a white space but returned separately so that calc.y can generate error messages on line numbers if needed (not shown in the current example).

Flex-Bison Flow & Build Commands




```
$ flex calc.l
$ yacc -dtv calc.y
$ g++ -c lex.yy.c
$ g++ -c y.tab.c
$ g++ lex.yy.o y.tab.o -lfl
```

Sample Run


```
$ ./a.out
12+8 $
= 20
$ ./a.out
12+2*45/4-23*(7+1) $
= -150
```

Handling of 12+8 $

- In the next slide we show the working of the parser on the input:

  `12 + 8 $`
- We use a pair of stacks – one for the grammar symbols for parsing and the other for keeping the associated attributes.
- We show the snapshot on every reduction (skipping the shifts).

Handling of 12+8 $

	Grammar				Reductions	
1:	S	$\rightarrow$	E		$\underline{\mathbf{num}_{12}} + \mathbf{num}_8$ \$	
2:	E	$\rightarrow$	$E + T$	$\Rightarrow$	$\underline{F} + \mathbf{num}_8$ \$	
3:	E	$\rightarrow$	$E - T$	$\Rightarrow$	$\underline{T} + \mathbf{num}_8$ \$	
4:	E	$\rightarrow$	T	$\Rightarrow$	$E + \underline{\mathbf{num}_8}$ \$	
5:	T	$\rightarrow$	$T * F$	$\Rightarrow$	$E + \underline{F}$ \$	
6:	T	$\rightarrow$	T / F	$\Rightarrow$	$\underline{E + T}$ \$	
7:	T	$\rightarrow$	F	$\Rightarrow$	$\underline{E}$ \$	
8:	F	$\rightarrow$	(E)	$\Rightarrow$	S \$	
9:	F	$\rightarrow$	$- E$			
10:	F	$\rightarrow$	**num**			

Stack

								num	8
								$+$	
num	12	F	12	T	12	E	12	E	12

Stack

F	8	T	8						
$+$		$+$							
E	12	E	12	E	20	S			

Output $= 20$

A Programmable Calculator Grammar

$$
\begin{array}{rlcl}
1: & L & \rightarrow & L\ S\ \backslash\mathrm{n} \\
2: & L & \rightarrow & S\ \backslash\mathrm{n} \\
3: & S & \rightarrow & \mathbf{id} = E \\
4: & S & \rightarrow & E \\
5: & E & \rightarrow & E + T \\
6: & E & \rightarrow & E - T \\
7: & E & \rightarrow & T \\
8: & T & \rightarrow & T * F \\
9: & T & \rightarrow & T\ /\ F \\
10: & T & \rightarrow & F \\
11: & F & \rightarrow & (E) \\
12: & F & \rightarrow & -E \\
13: & F & \rightarrow & \mathbf{num} \\
14: & F & \rightarrow & \mathbf{id}
\end{array}
$$

Yacc Specs (calc.y) for Programmable Calculator Grammar


```
%{
#include <string.h>
#include <iostream>
#include "parser.h"

extern int yylex();
void yyerror(char *s);

#define NSYMS 20 /* max # of symbols */
symboltable symtab[NSYMS];
%}

%union {
    int intval;
    struct symtab *symp;
}

%token <symp> NAME
%token <intval> NUMBER

%type <intval> expression
%type <intval> term
%type <intval> factor

%%
stmt_list: stmt_list statement '\n'
         | statement '\n'
         ;
```

```
statement: NAME '=' expression
                { $1->value = $3; }
         | expression
                { printf("= %d\n", $1); }
         ;
expression: expression '+' term
                { $$ = $1 + $3; }
          | expression '-' term
                { $$ = $1 - $3; }
          | term
          ;
term: term '*' factor
          { $$ = $1 * $3; }
    | term '/' factor
          { if ($3 == 0.0)
                yyerror("divide by zero");
            else
                $$ = $1 / $3;
          }
    | factor
    ;
factor: '(' expression ')'
            { $$ = $2; }
      | '-' factor
            { $$ = -$2; }
      | NUMBER
      | NAME
            { $$ = $1->value; }
      ;
%%
```

Yacc Specs (calc.y) for Programmable Calculator Grammar


```
struct symtab *symlook(char *s) {
    char *p;
    struct symtab *sp;
    for(sp = symtab;
        sp < &symtab[NSYMS]; sp++) {
        /* is it already here? */
        if (sp->name &&
           !strcmp(sp->name, s))
            return sp;
        if (!sp->name) {
        /* is it free */
            sp->name = strdup(s);
            return sp;
        }
        /* otherwise continue to next */
    }
    yyerror("Too many symbols");
    exit(1); /* cannot continue */
} /* symlook */
```

```
void yyerror(char *s) {
    std::cout << s << std::endl;
}

int main() {
    yyparse();
}
```

Header (y.tab.h) for Programmable Calculator


```
/* A Bison parser, made by GNU Bison 2.5.  */
/* Tokens.  */
#ifndef YYTOKENTYPE
# define YYTOKENTYPE
   /* Put the tokens into the symbol table, so that GDB and other debuggers know about them. */
   enum yytokentype {
     NAME = 258,
     NUMBER = 259
   };
#endif
/* Tokens.  */
#define NAME 258
#define NUMBER 259

#if ! defined YYSTYPE && ! defined YYSTYPE_IS_DECLARED
typedef union YYSTYPE {
#line 11 "calc.y" /* Line 2068 of yacc.c  */

    int intval;
    struct symtab *symp;

#line 65 "y.tab.h" /* Line 2068 of yacc.c  */
} YYSTYPE;
# define YYSTYPE_IS_TRIVIAL 1
# define yystype YYSTYPE /* obsolescent; will be withdrawn */
# define YYSTYPE_IS_DECLARED 1
#endif

extern YYSTYPE yylval;
```

Header (parser.h) for Programmable Calculator


```
#ifndef __PARSER_H
#define __PARSER_H

typedef struct symtab {
    char *name;
    int value;
} symboltable;

symboltable *symlook(char *);

#endif // __PARSER_H
```

Flex Specs (calc.l) for Programmable Calculator Grammar


```
%{
#include <math.h>
#include "y.tab.h"
#include "parser.h"
%}

ID          [A-Za-z][A-Za-z0-9]*

%%
[0-9]+      {
              yylval.intval = atoi(yytext);
              return NUMBER;
            }

[ \t]       ;  /* ignore white space */


{ID}        { /* return symbol pointer */
              yylval.symp = symlook(yytext);
              return NAME;
            }

"$"         { return 0; /* end of input */ }

\n|.        return yytext[0];
%%
```

Note on Programmable Calculator

- Symbol Table
  - We have introduced variables (**id**) in the grammar now to support programmability (to store intermediate results).
  - **id**'s are maintained in the (rudimentary) symbol table as a name-value doublet (refer: `parser.h`).

    ```
    struct symtab { char *name; int value; };
    ```

  - Every **id**, as soon as found in the lexer for the first time, is inserted in the symbol table. On every subsequent occurrence the same **id** is referred from the symbol table. The function `struct symtab *symlook(char *);` achieves this.
- `union` Wrapper
  - Tokens `NAME` and `NUMBER` have different attributes `intval` and `symp` respectively.
  - For defining a value-stack in `C`, these are wrapped in a single union:

    ```
    typedef union YYSTYPE {
        int intval;
        struct symtab *symp;
    } YYSTYPE;
    ```

Sample Run

Output

```
$ ./a.out
a = 8 + 9
a + 4
= 21
$
```

Grammar

1:	L	$\rightarrow$	$L\ S\ \backslash n$
2:	L	$\rightarrow$	$S\ \backslash n$
3:	S	$\rightarrow$	$\mathbf{id} = E$
4:	S	$\rightarrow$	E
5:	E	$\rightarrow$	$E + T$
6:	E	$\rightarrow$	$E - T$
7:	E	$\rightarrow$	T
8:	T	$\rightarrow$	$T * F$
9:	T	$\rightarrow$	T / F
10:	T	$\rightarrow$	F
11:	F	$\rightarrow$	(E)
12:	F	$\rightarrow$	$- E$
13:	F	$\rightarrow$	$\mathbf{num}$
14:	F	$\rightarrow$	$\mathbf{id}$

Derivation

$$
\begin{aligned}
L\ \$ &\Rightarrow \underline{L\ S\ \backslash n}\ \$ \\
&\Rightarrow L\ \underline{E}\ \backslash n\ \$ \\
&\Rightarrow \underline{L\ E + T}\ \backslash n\ \$ \\
&\Rightarrow L\ E + \underline{F}\ \backslash n\ \$ \\
&\Rightarrow L\ E + \underline{\mathbf{num}_4}\ \backslash n\ \$ \\
&\Rightarrow L\ \underline{T} + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow L\ \underline{F} + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow L\ \underline{\mathbf{id}_a} + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \underline{S\ \backslash n}\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \underline{\mathbf{id}_a = E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{E + T}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = E + \underline{F}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = E + \underline{\mathbf{num}_9}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{T} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{F} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{\mathbf{num}_8} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$
\end{aligned}
$$

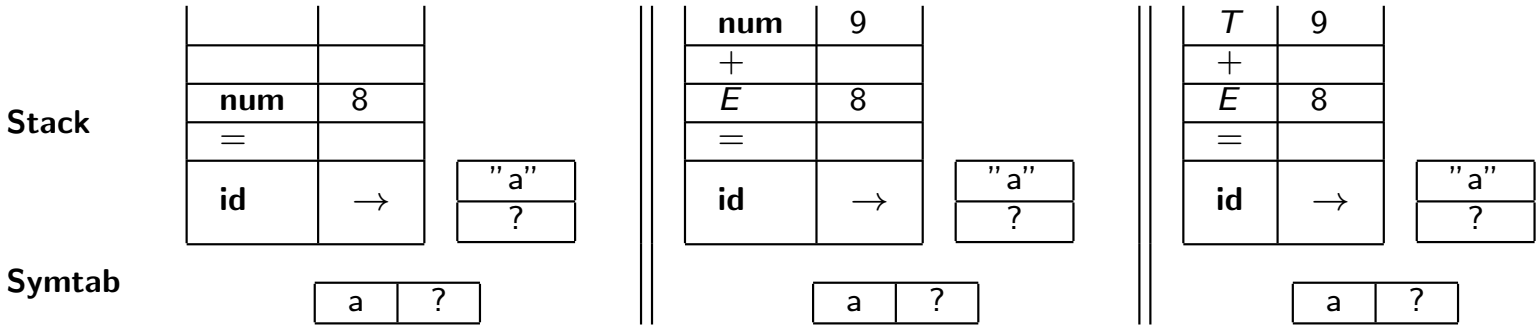
Handling of a = 8 + 9 \n a + 4 \n $

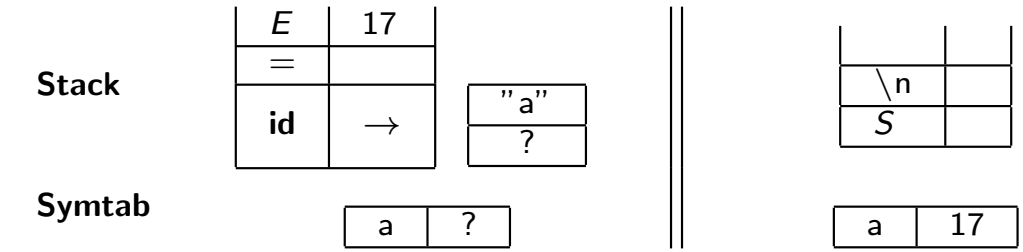
Grammar

L	$\rightarrow$	$L\ S\ \backslash n$	T	$\rightarrow$	$T * F$
L	$\rightarrow$	$S\ \backslash n$	T	$\rightarrow$	$T\ /\ F$
S	$\rightarrow$	$\mathbf{id} = E$	T	$\rightarrow$	F
S	$\rightarrow$	E	F	$\rightarrow$	(E)
E	$\rightarrow$	$E + T$	F	$\rightarrow$	$-E$
E	$\rightarrow$	$E - T$	F	$\rightarrow$	**num**
E	$\rightarrow$	T	F	$\rightarrow$	**id**

Reductions

$$
\begin{aligned}
& \mathbf{id}_a = \underline{\mathbf{num}_8} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \mathbf{id}_a = \underline{F} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \mathbf{id}_a = \underline{T} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \mathbf{id}_a = E + \underline{\mathbf{num}_9}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \mathbf{id}_a = E + \underline{F}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \mathbf{id}_a = \underline{E + T}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \underline{\mathbf{id}_a = E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \underline{S\ \backslash n}\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
\Leftarrow\ & \underline{L\ \mathbf{id}_a} + \mathbf{num}_4\ \backslash n\ \$
\end{aligned}
$$





Handling of a = 8 + 9 \n a + 4 \n $

Grammar

L	$\rightarrow$	$L\ S\ \backslash n$	T	$\rightarrow$	$T * F$		
L	$\rightarrow$	$S\ \backslash n$	T	$\rightarrow$	$T\ /\ F$		
S	$\rightarrow$	**id** $= E$	T	$\rightarrow$	F		
S	$\rightarrow$	E	F	$\rightarrow$	(E)		
E	$\rightarrow$	$E + T$	F	$\rightarrow$	$-\ E$		
E	$\rightarrow$	$E - T$	F	$\rightarrow$	**num**		
E	$\rightarrow$	T	F	$\rightarrow$	**id**		

Reductions

$\Leftarrow$ L $\underline{\textbf{id}_a}$ + **num**$_4$ \n $

$\Leftarrow$ L $\underline{F}$ + **num**$_4$ \n $

$\Leftarrow$ L $\underline{T}$ + **num**$_4$ \n $

$\Leftarrow$ L $\underline{E}$ + $\underline{\textbf{num}_4}$ \n $

$\Leftarrow$ L E + $\underline{F}$ \n $

$\Leftarrow$ L $\underline{E + T}$ \n $

$\Leftarrow$ L $\underline{E}$ \n $

$\Leftarrow$ $\underline{L\ S\ \backslash n}$ $

$\Leftarrow$ L $

Stack

id	$\rightarrow$	"a"
L		17

num	4
+	
E	17
L	

T	4
+	
E	17
L	

Symtab

a	17

a	17

a	17

Stack

E	21
L	

\n	
S	
L	

L	

Symtab

a	17

a	17

a	17

Output

$= 21$

A Programmable Calculator Grammar (with Ambiguous Grammar)

$$
\begin{array}{rlcl}
1: & L & \rightarrow & L\ S\ \backslash\mathrm{n} \\
2: & L & \rightarrow & S\ \backslash\mathrm{n} \\
3: & S & \rightarrow & \mathbf{id} = E \\
4: & S & \rightarrow & E \\
5: & E & \rightarrow & E + E \\
6: & E & \rightarrow & E - E \\
7: & E & \rightarrow & E * E \\
8: & E & \rightarrow & E\ /\ E \\
9: & E & \rightarrow & (E) \\
10: & E & \rightarrow & -\ E \\
11: & E & \rightarrow & \mathbf{num} \\
12: & E & \rightarrow & \mathbf{id}
\end{array}
$$

Yacc Specs (calc.y) for Programmable Calculator Grammar


```
%{
#include <string.h>
#include <iostream>
#include "parser.h"
extern int yylex();
void yyerror(char *s);
#define NSYMS 20 /* max # of symbols */
symboltable symtab[NSYMS];
%}

%union {
    int intval;
    struct symtab *symp;
}

%token <symp> NAME
%token <intval> NUMBER

%left '+' '-'
%left '*' '/'
%nonassoc UMINUS

%type <intval> expression
%%

stmt_list: statement '\n'
         | stmt_list statement '\n'
         ;
```

```
statement: NAME '=' expression
                { $1->value = $3; }
         | expression
                { printf("= %d\n", $1); }
         ;

expression: expression '+' expression
                { $$ = $1 + $3; }
          | expression '-' expression
                { $$ = $1 - $3; }
          | expression '*' expression
                { $$ = $1 * $3; }
          | expression '/' expression
                { if ($3 == 0)
                      yyerror("divide by zero");
                  else
                      $$ = $1 / $3;
                }
          | '(' expression ')'
                { $$ = $2; }
          | '-' expression %prec UMINUS
                { $$ = -$2; }
          | NUMBER
          | NAME
                { $$ = $1->value; }
          ;
%%
```

Yacc Specs (calc.y) for Programmable Calculator Grammar


```
struct symtab *symlook(char *s) {
    char *p;
    struct symtab *sp;
    for(sp = symtab;
        sp < &symtab[NSYMS]; sp++) {
        /* is it already here? */
        if (sp->name &&
           !strcmp(sp->name, s))
            return sp;
        if (!sp->name) {
        /* is it free */
            sp->name = strdup(s);
            return sp;
        }
        /* otherwise continue to next */
    }
    yyerror("Too many symbols");
    exit(1); /* cannot continue */
} /* symlook */
```

```
void yyerror(char *s) {
    std::cout << s << std::endl;
}

int main() {
    yyparse();
}
```

Note on Yacc Specs (calc.y) for Ambiguous Grammar

- Ambiguous Grammars
  - Ease specification of languages - particularly the operator expressions.
  - Offer shorter and more compact representation.
  - Lead to less reduction steps during parsing.
  - Introduce shift / reduce conflicts in the LR parser.
  - Conflict are resolved by precedences and associativities of operators.
- Associativity
  - `%left` is used to specify left-associative operators.
  - `%right` is used to specify right-associative operators.
  - `%nonassoc` is used to specify non-associative operators.
- Precedence
  - Precedence is specified by the order of `%left`, `%right`, or `%nonassoc` definitions. Later in the order, higher the precedence. However, all operators in the same definition have the same precedence.
  - All operators having the same precedence must have the same associativity.

Note on Yacc Specs (calc.y) for Ambiguous Grammar

- Overloaded Operators
  - Operators like '-' are overloaded in unary and binary forms and have different precedences. We use a symbolic name UMINUS for (say) the unary operator while the binary one is marked as '-'.

    ```
    %left '-'
    %nonassoc UMINUS
    ```
  - The rule with the unary minus is bound to this symbolic name using %prec marker.

    ```
    expression: '-' expression %prec UMINUS
              | expression '-' expression
    ```
  - Note that the lexer (calc.l) would continue to return the same '-' token for unary as well as binary instances of the operators. However, Yacc can use the precedence information to resolve between the two.

Header (y.tab.h) for Programmable Calculator


```
/* A Bison parser, made by GNU Bison 2.5.  */
/* Tokens. */
#ifndef YYTOKENTYPE
# define YYTOKENTYPE
   /* Put the tokens into the symbol table, so that GDB and other debuggers know about them. */
   enum yytokentype {
     NAME = 258,
     NUMBER = 259,
     UMINUS = 260
   };
#endif
/* Tokens. */
#define NAME 258
#define NUMBER 259
#define UMINUS 260

#if ! defined YYSTYPE && ! defined YYSTYPE_IS_DECLARED
typedef union YYSTYPE {
#line 11 "calc.y" /* Line 2068 of yacc.c  */

    int intval;
    struct symtab *symp;

#line 67 "y.tab.h" /* Line 2068 of yacc.c  */
} YYSTYPE;
# define YYSTYPE_IS_TRIVIAL 1
# define yystype YYSTYPE /* obsolescent; will be withdrawn */
# define YYSTYPE_IS_DECLARED 1
#endif

extern YYSTYPE yylval;
```

Header (parser.h) for Programmable Calculator


```
#ifndef __PARSER_H
#define __PARSER_H

typedef struct symtab {
    char *name;
    int value;
} symboltable;

symboltable *symlook(char *);

#endif // __PARSER_H
```

Flex Specs (calc.l) for Programmable Calculator Grammar


```
%{
#include <math.h>
#include "y.tab.h"
#include "parser.h"
%}

ID          [A-Za-z][A-Za-z0-9]*

%%
[0-9]+      {
              yylval.intval = atoi(yytext);
              return NUMBER;
            }

[ \t]       ;  /* ignore white space */


{ID}        { /* return symbol pointer */
              yylval.symp = symlook(yytext);
              return NAME;
            }

"$"         { return 0; /* end of input */ }

\n|.        return yytext[0];
%%
```

Sample Run

Output

```
$ ./a.out
a = 8 + 9
a + 4
= 21
$
```

Grammar

1:	L	$\rightarrow$	$L\ S\ \backslash n$
2:	L	$\rightarrow$	$S\ \backslash n$
3:	S	$\rightarrow$	$\mathbf{id} = E$
4:	S	$\rightarrow$	E
5:	E	$\rightarrow$	$E + E$
6:	E	$\rightarrow$	$E - E$
7:	E	$\rightarrow$	$E * E$
8:	E	$\rightarrow$	E / E
9:	E	$\rightarrow$	(E)
10:	E	$\rightarrow$	$- E$
11:	E	$\rightarrow$	$\mathbf{num}$
12:	E	$\rightarrow$	$\mathbf{id}$

Derivation

$$
\begin{aligned}
L\ \$ &\Rightarrow \underline{L\ S\ \backslash n}\ \$ \\
&\Rightarrow L\ \underline{E}\ \backslash n\ \$ \\
&\Rightarrow L\ \underline{E + E}\ \backslash n\ \$ \\
&\Rightarrow L\ E + \underline{E}\ \backslash n\ \$ \\
&\Rightarrow L\ E + \underline{\mathbf{num}_4}\ \backslash n\ \$ \\
&\Rightarrow L\ \underline{\mathbf{id}_a} + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \underline{S\ \backslash n}\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \underline{\mathbf{id}_a = E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{E + E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = E + \underline{\mathbf{num}_9}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$ \\
&\Rightarrow \mathbf{id}_a = \underline{\mathbf{num}_8} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$
\end{aligned}
$$

Handling of a = 8 + 9 \n a + 4 \n $

Grammar

L	$\rightarrow$	$L\ S\ \backslash n$	E	$\rightarrow$	$E * E$		
L	$\rightarrow$	$S\ \backslash n$	E	$\rightarrow$	$E\ /\ E$		
S	$\rightarrow$	$\mathbf{id} = E$	E	$\rightarrow$	(E)		
S	$\rightarrow$	E	E	$\rightarrow$	$-E$		
E	$\rightarrow$	$E + E$	E	$\rightarrow$	$\mathbf{num}$		
E	$\rightarrow$	$E - E$	E	$\rightarrow$	$\mathbf{id}$		

Reductions

$$\mathbf{id}_a = \underline{\mathbf{num}_8} + \mathbf{num}_9\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$$$
$$\Rightarrow \mathbf{id}_a = E + \underline{\mathbf{num}_9}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$$$
$$\Rightarrow \mathbf{id}_a = \underline{E + E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$$$
$$\Rightarrow \underline{\mathbf{id}_a = E}\ \backslash n\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$$$
$$\Rightarrow \underline{S\ \backslash n}\ \mathbf{id}_a + \mathbf{num}_4\ \backslash n\ \$$$
$$\Rightarrow \underline{L\ \mathbf{id}_a} + \mathbf{num}_4\ \backslash n\ \$$$

Stack

				num	9			E	9		
				+				+			
num	8			E	8			E	8		
=				=				=			
id	→	"a"		id	→	"a"		id	→	"a"	
		?				?				?	

Symtab

a	?			a	?			a	?		

Stack

E	17										
=				\n				L			
id	→	"a"		S							
		?									

Symtab

a	?			a	17			a	17		

Handling of a = 8 + 9 \n a + 4 \n $

Grammar

L	$\rightarrow$	$L\ S\ \backslash n$	E	$\rightarrow$	$E * E$	$\Rightarrow$	$L\ \underline{\mathbf{id}_a} + \mathbf{num}_4\ \backslash n\ \$$	
L	$\rightarrow$	$S\ \backslash n$	E	$\rightarrow$	E / E	$\Rightarrow$	$L\ E + \underline{\mathbf{num}_4}\ \backslash n\ \$$	
S	$\rightarrow$	$\mathbf{id} = E$	E	$\rightarrow$	(E)	$\Rightarrow$	$L\ E + \underline{E}\ \backslash n\ \$$	
S	$\rightarrow$	E	E	$\rightarrow$	$- E$	$\Rightarrow$	$L\ \underline{E + E}\ \backslash n\ \$$	
E	$\rightarrow$	$E + E$	E	$\rightarrow$	$\mathbf{num}$	$\Rightarrow$	$L\ \underline{E}\ \backslash n\ \$$	
E	$\rightarrow$	$E - E$	E	$\rightarrow$	$\mathbf{id}$	$\Rightarrow$	$\underline{L\ S\ \backslash n}\ \$$	
						$\Rightarrow$	$L\ \$$	

Reductions

Stack

			num	4
id	$\rightarrow$	"a"	+	
L		17	E	17
			L	

E	4	
+		
E	17	
L		

Symtab

a	17

a	17

a	17

Stack

E	21
L	

$\backslash n$	
S	
L	

L	

Symtab

a	17

a	17

a	17

Output

$= 21$