

# *Introduction to Parsing Using Lex/Flex and Yacc/Bison*

Uday Khedker

([www.cse.iitb.ac.in/~uday](http://www.cse.iitb.ac.in/~uday))

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



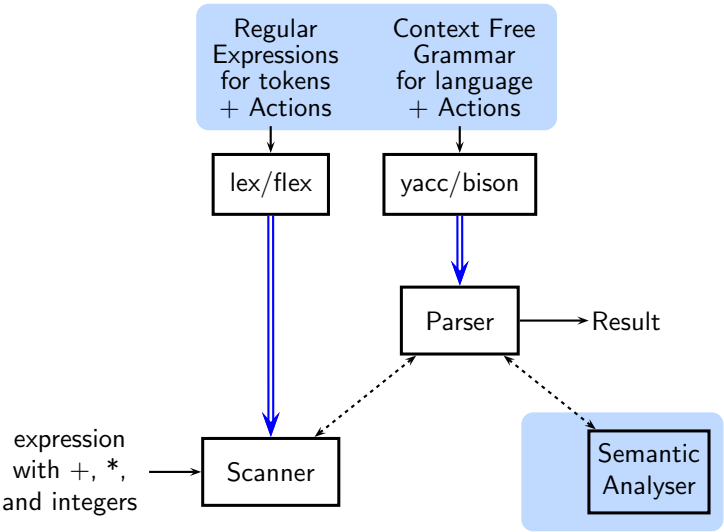
Jan 2024



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

# simCalc: A Simple Calculator





# Introduction to Lex and Yacc

IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

- lex script num.l to show token identification
- exp1.y to show a simple expression. Use the lex script exp.l
- exp2.y to show construction of a PLUS expression  
Interesting input:  $1 + 2 + 3 + 4 + 5 + 6$
- exp3.y fixes the problem by making + left associative
- exp4.y includes + and \*  
Interesting inputs:  $1 + 2 * 3$  and  $1 * 2 + 3$
- exp5.y fixes the above problem

# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

- Grammar

```
Expr  :  Expr '+' Expr
       |  Expr '*' Expr
       |  NUM
```

- Terminals and non-terminals get defined by the grammar
- Scanner identifies the tokens and communicates the details to the parser

Token Name	Token Lexemes	Token Value	Token Code
Number	"10"	10	NUM
	"345"	345	
	"03"	3	
+ operator	"+"		'+'
* operator	"*"		'*'

# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

scanner.l



lex/flex

parser.y



yacc/bison

# The Interaction Between Scanner and Parser

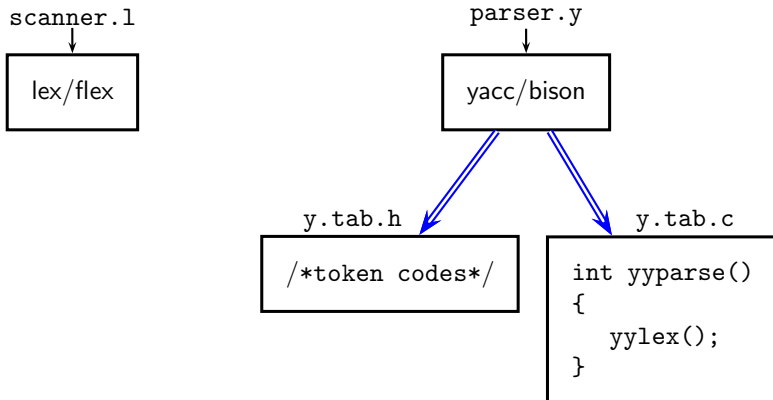


IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

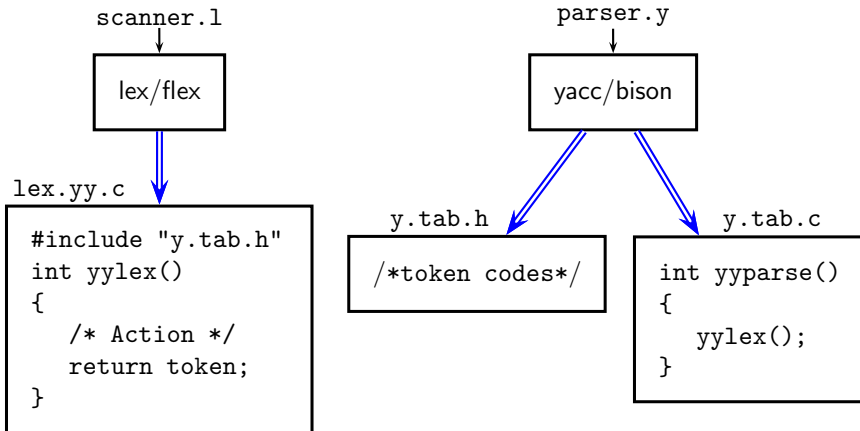


# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

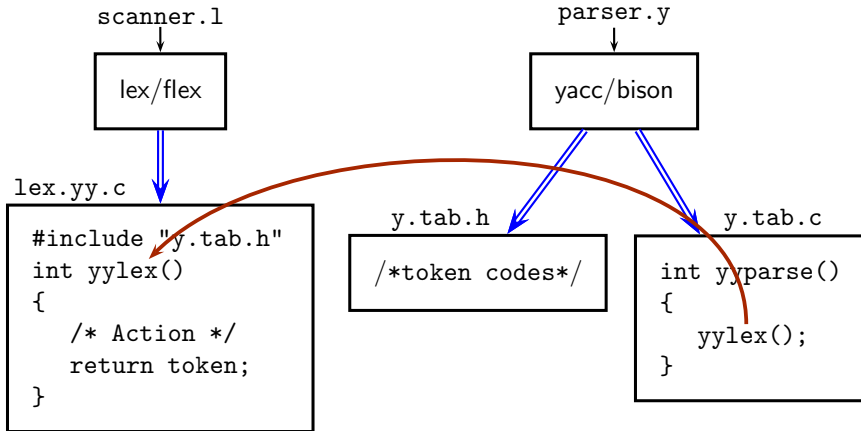


# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:



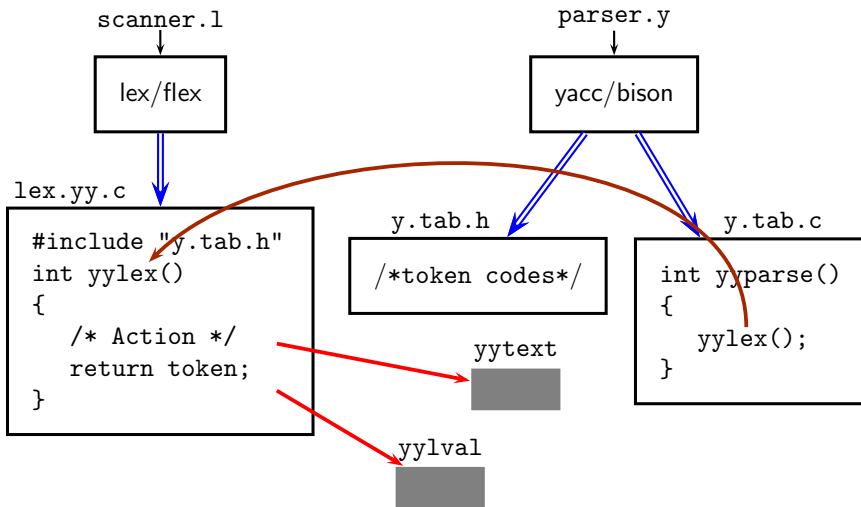


# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

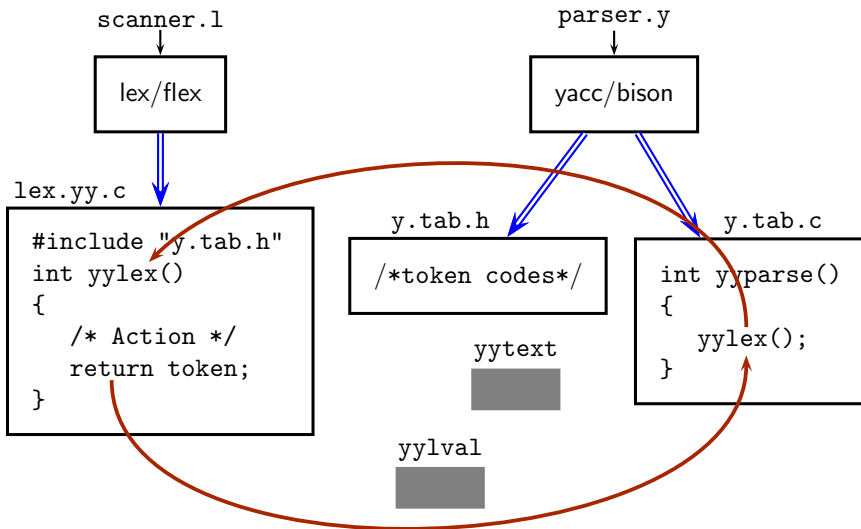


# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

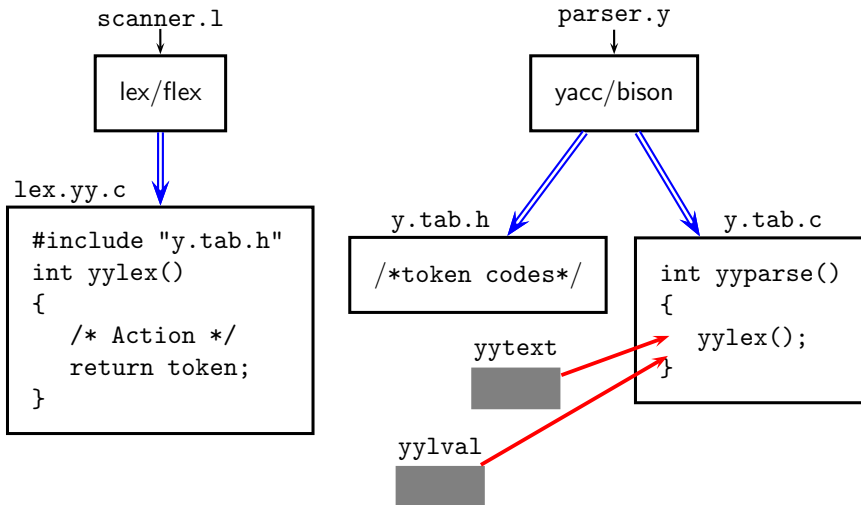


# The Interaction Between Scanner and Parser



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:





# Some Issues in the Simple Calculator

IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

- Syntax: Precedences and associativities
- Semantics: Values of
  - Terminals (from scanner)
  - Non-terminals (in the parser)
- Unit testing:  
Use of `#if SCANTEST` to test the scanner



# Notes

IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

- Sample inputs  
a 20 + b 3
- Use of "."
  - For any pattern, the longest matching lexeme is chosen  
(If a proper prefix of the lexeme matches a pattern, the match is ignored)
  - If a lexeme can match two patterns, the first pattern is chosen.
- The Lex and Yacc Page  
<http://dinosaur.compilertools.net/>



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
       |  Expr '*' Expr
       |  NUM
```

- The process of parsing

10    +    20    \*    3    eof  
NUM

Parsing  
Stack

Action: Shift NUM



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  : Expr '+' Expr
      | Expr '*' Expr
      | NUM
```

- The process of parsing

10    +    20    \*    3    eof

+

top → NUM  
Parsing  
Stack

Action: Reduce by “Expr: NUM”



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
       |  Expr '*' Expr
       |  NUM
```

- The process of parsing

10    +    20    \*    3    eof

+

top → 

Expr

  
Parsing  
Stack      Action: Shift +





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

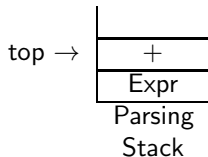
# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  : Expr '+' Expr
      | Expr '*' Expr
      | NUM
```

- The process of parsing

10    +    20    \*    3    eof  
          NUM



Action: Shift NUM



# An Overview of Shift Reduce Parsing

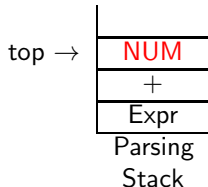
- Grammar

```
Expr  : Expr '+' Expr
      | Expr '*' Expr
      | NUM
```

- The process of parsing

10    +    20    \*    3    eof

\*



Action: Reduce by "Expr: NUM"



# An Overview of Shift Reduce Parsing

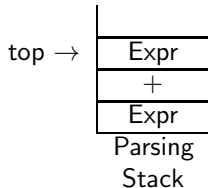
- Grammar

```
Expr  :  Expr '+' Expr
       |  Expr '*' Expr
       |  NUM
```

- The process of parsing

10    +    20    \*    3    eof

\*



Action: Shift \* or Reduce by “Expr: Expr + Expr”?



IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

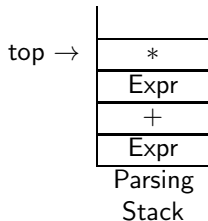
# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
       |  Expr '*' Expr
       |  NUM
```

- The process of parsing

10    +    20    \*    3    eof  
                        NUM



Action: Shift NUM



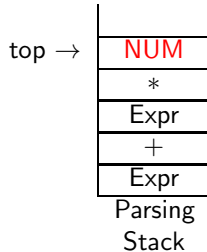
# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
      |  Expr '*' Expr
      |  NUM
```

- The process of parsing

```
10      +      20      *      3      |eof
                                           EOF
```



Action: Reduce by “Expr: NUM”

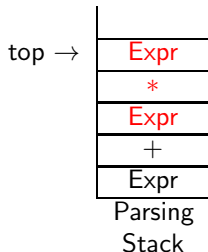
## An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
      |  Expr '*' Expr
      |  NUM
```

- The process of parsing

```
10      +      20      *      3      eof
                                           EOF
```



Action: Reduce by "Expr: Expr \* Expr"

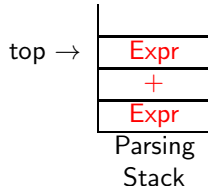
## An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
      |  Expr '*' Expr
      |  NUM
```

- The process of parsing

```
10      +      20      *      3      |eof
                                           EOF
```



Action: Reduce by "Expr: Expr + Expr"



# An Overview of Shift Reduce Parsing

- Grammar

```
Expr  :  Expr '+' Expr
      |  Expr '*' Expr
      |  NUM
```

- The process of parsing

```
10      +      20      *      3      |eof
                                           EOF
```

top → 

Expr

      Action: Accept

Parsing  
Stack





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:

Lex & Yacc

Section:

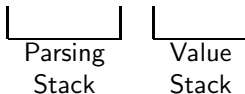
# An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 eof  
NUM





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

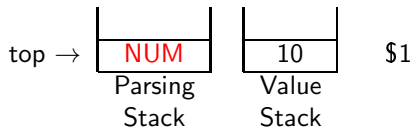
# An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 eof  
     +





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

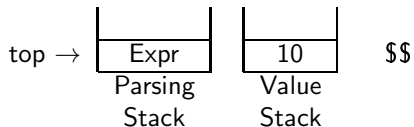
# An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10    +    20    eof  
          +





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

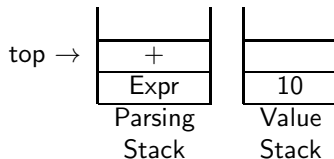
# An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 eof  
          NUM





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

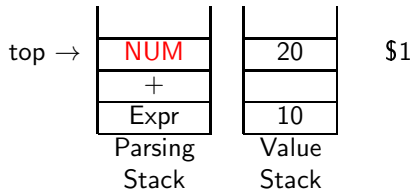
# An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 eof  
EOF





# An Overview of Attribute Evaluation

- Grammar and the associated actions

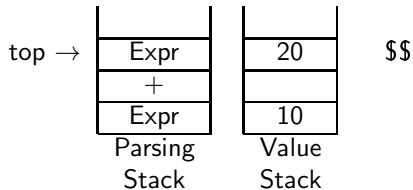
Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 

eof
-----

  
EOF





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

# An Overview of Attribute Evaluation

- Grammar and the associated actions

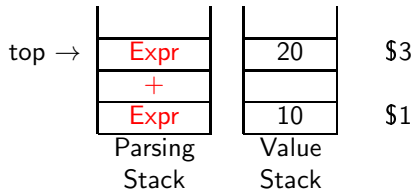
Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 

eof
-----

  
EOF





IIT Bombay  
cs302: Implementation  
of Programming  
Languages

Topic:  
Lex & Yacc  
Section:

## An Overview of Attribute Evaluation

- Grammar and the associated actions

Expr : Expr '+' Expr { \$\$ = \$1 + \$3; }  
     | NUM

- The process of parsing and attribute evaluation

10 + 20 eof  
EOF

