Compiler Design Laboratory (CS 753)

Samit Biswas

samit@cs.iiests.ac.in

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

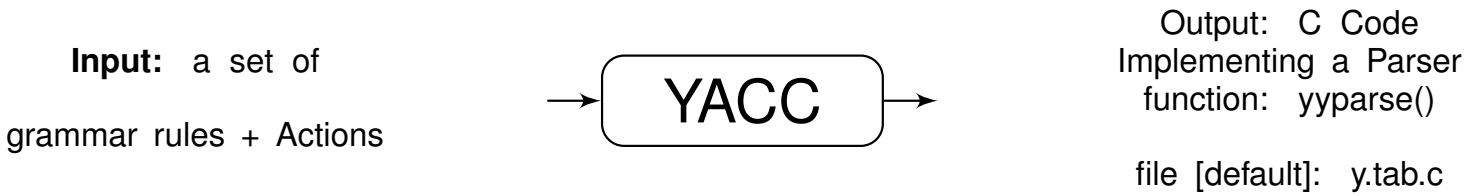
September 20, 2019

YACC

YACC Parser — Overview

Parser generator

- Takes a specification for a context-free grammar.
- Produces code for a parser.

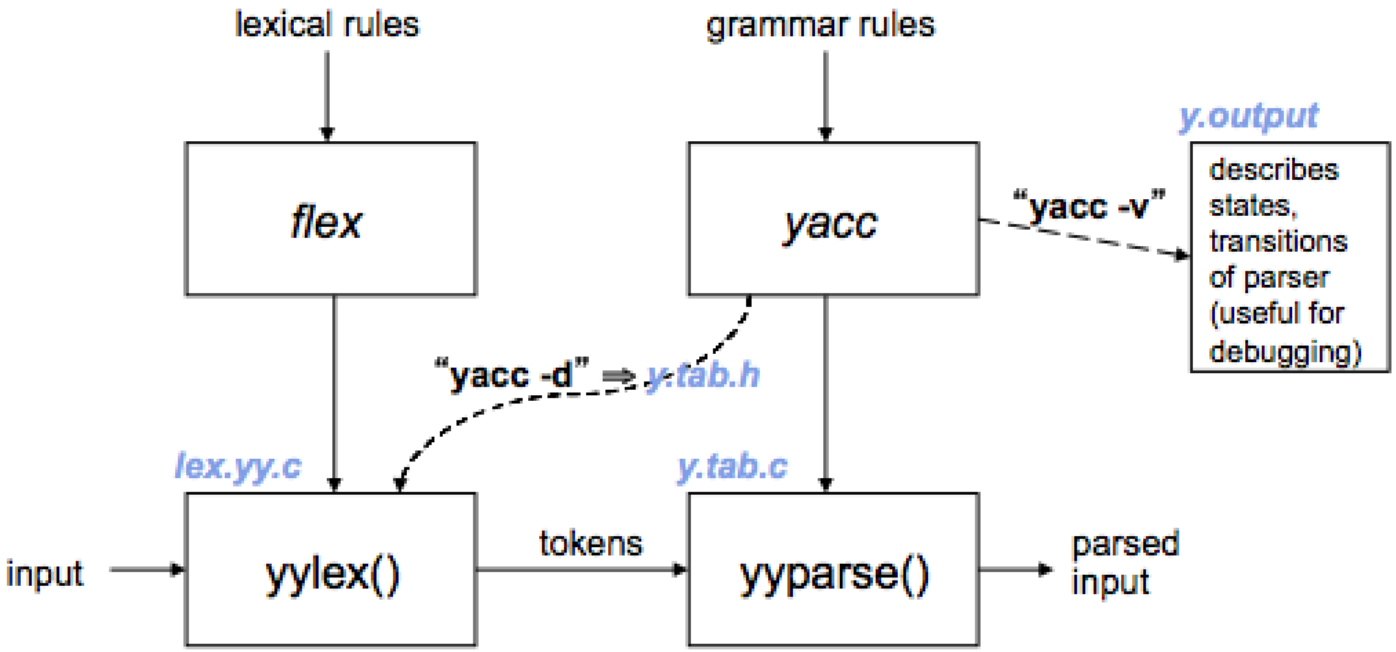


Scanner-Parser interaction

- Parser assumes `int yylex()` implements the scanner.
- Scanner:
  - return value indicates the type of token found;
  - other values communicated using `yytext, yylval`.
- YACC determines integer representations for tokens:
  - Communicated to scanner in file `y.tab.h`

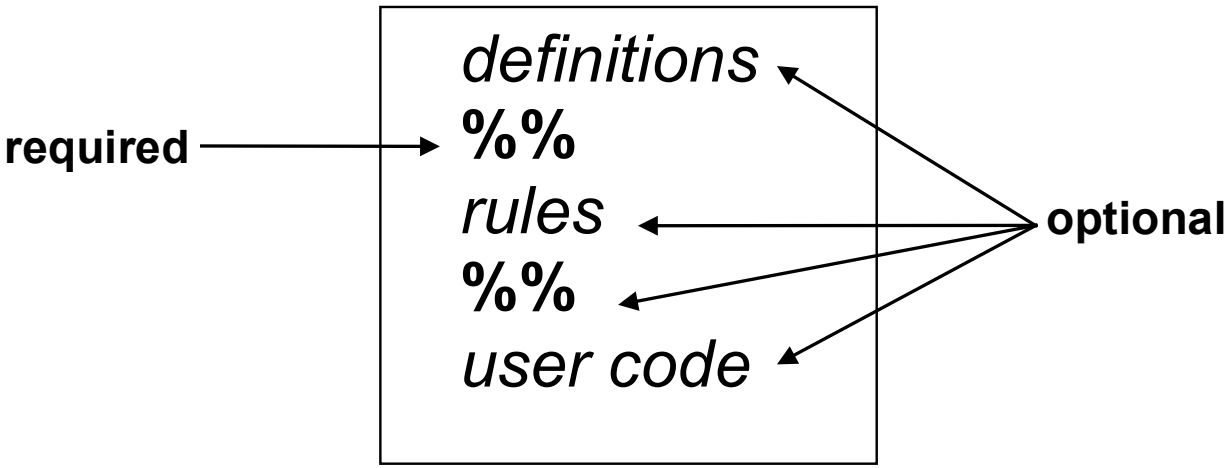
    use `yacc -d` to produce `y.tab.h`
  - Token encodings:
    - "end of file" represented by 0;
    - a character literal: its ASCII value;
    - other tokens: assigned numbers ≥ 257

Scanner-Parser interaction



flex input format

A YACC input file consists of three parts:



- Shortest possible legal YACC input:

```
%%
```

Definition Section

There are three things that can go in the definitions section:

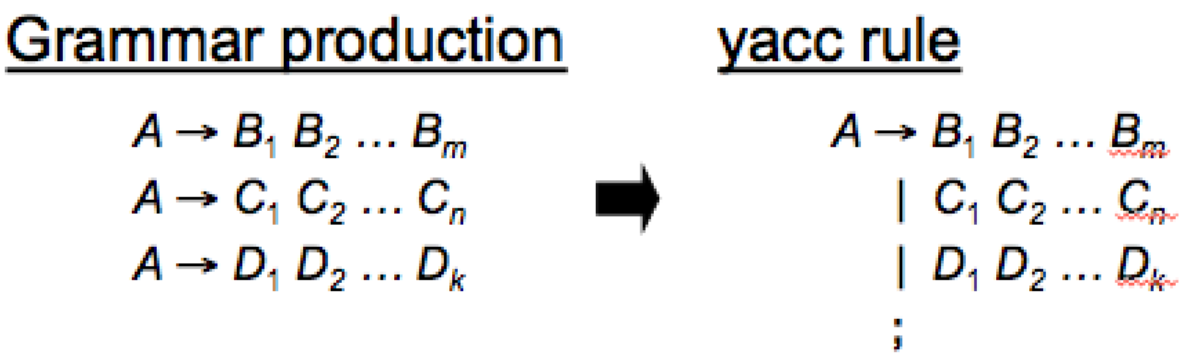
- **C code:** Code between %{ and %} is copied to the C file.
- **Definitions:**
- **Associativity Rules:** These handle associativity and priority of operators.

Information about Tokens

- **token names:**
  - declared using '**%token**'.
  - single-character tokens don't have to be declared.
  - any name not declared as a token is assumed to be a nonterminal.
- start symbol of grammar, using %start [optional]
- operator info:
  - precedence, associativity.

Rules section

The rules section contains the grammar of the language you want to parse. This looks like



- Rule RHS can have arbitrary C code embedded, within { ... }.

Conflicts

- Conflicts arise when there is more than one way to proceed with parsing.
- Two types:
  - shift-reduce [default action: shift]
  - reduce-reduce [default: reduce with the first rule listed]
- Removing conflicts:
  - specify operator precedence, associativity;
  - restructure the grammar

User code section

The minimal main program is:

```
int main ()
 {
      yyparse();
      return 0;
 }
```

Assignment

Write YACC specification to design a parser for a subset of the C language. Consider the following programming language constructs:

- Main function
- Statements
  - Statements like local/global declarations.
  - Assignment Statements.
  - Conditional Statements.
  - Iterative Statements
  - Function Call
- User defined functions

Consider the assumptions for the language previously considered in last lab sessions. Variables of data types allowed in the subset can be defined at the very beginning of any function / block. A variable must be defined before it is used.