

GNU Bison

Kulikov Alexey

ABBYY

Moscow, 2017

```
%x quote  
%%
```

...other rules for dealing with quotes...

```
<quote><<EOF>> {  
    error( "unterminated quote" );  
    yyterminate();  
}  
<<EOF>> {  
    if ( *++filelist )  
        yyin = fopen( *filelist , "r" );  
    else  
        yyterminate();  
}
```

- %option c++, or 'flex -+', or 'flex++'
- 'lex.yy.c' → 'lex.yy.cc', includes 'FlexLexer.h'
- FlexLexer
 - const char* YYText() ↔ yytext
 - int YYLeng() ↔ yyleng
 - int lineno() const ↔ %option yylineno
- yyFlexLexer : FlexLexer
 - yyFlexLexer(istream* yyin, ostream* yyout)
 - virtual int yylex()

What is Bison?

- CF grammar \rightarrow LR parser with LALR(1) tables;
- Upward-compatible with Yacc;
- Is a part of the GNU Project;
- www.gnu.org/software/bison/;
- Used by: Ruby, PHP, GCC (before 2004), Go, Bash, MySQL, LilyPond, ...

- 1 *.y → Bison → yyparse: *.tab.c, *.tab.h
- 2 *.lex → Lex → yylex: *.yy.c
- 3 Input → yylex → Token_ID
- 4 Token_ID → yyparse → Semantic Operation

```
%{  
Prologue  
%}
```

Bison declarations

```
%%  
Grammar rules  
%%
```

Epilogue

- Contains definitions of: macros, functions, etc;
- Inserted into .c file before definition of yyparse;
- **%code** inserts:
 - (unqualified): into default location, *verbatim*;
 - **requires**: into .h and .c before YYL/STYPE defs good for writing YYL/STYPE overrides;
 - **provides**: into .h and .c after YY*, token defs good for other defs and decls;
 - **top**: inserted near the top of .c file;
 - **imports**: Java-specific, for import directives.

- YYLTYPE: Macro for the data type of yylloc;
- YYSTYPE: Macro for the data type of semantic values; int by default.

À bientôt.

- All token type names must be declared;
- The first rule specifies the start symbol;
- %token name \rightarrow #define