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

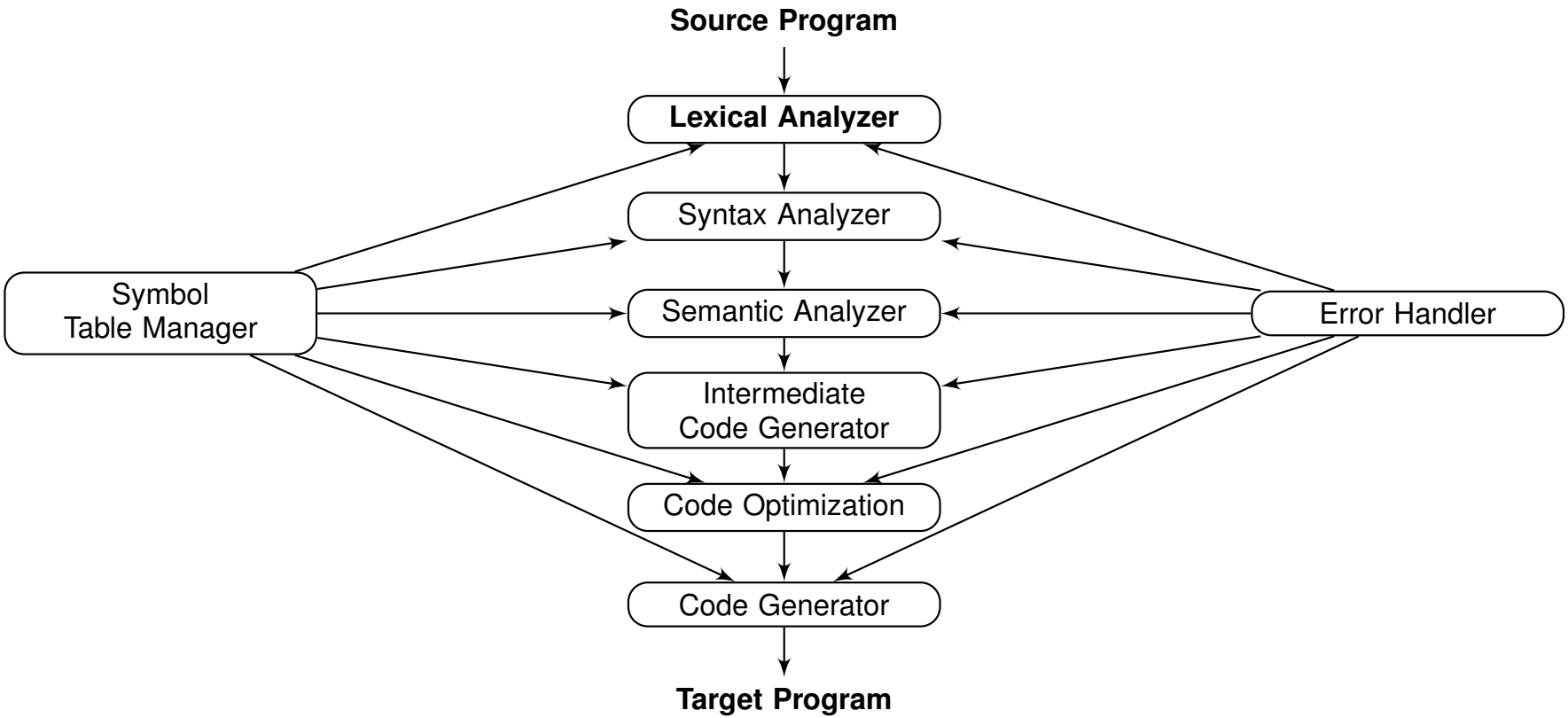
July 16, 2018

Phases Of Compilation

Lexical Analyzer

Assignment

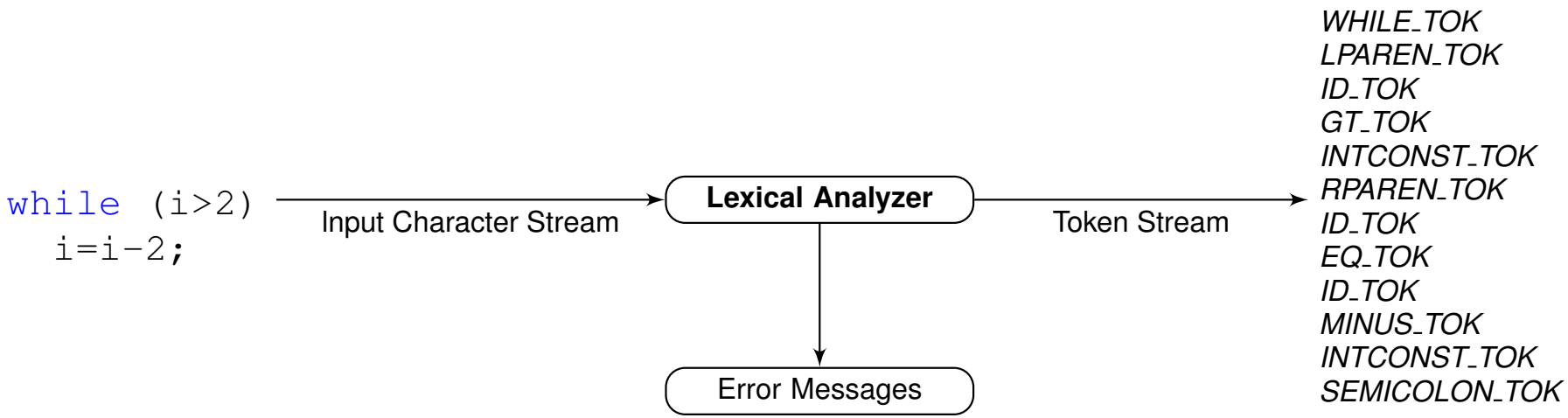
Phases Of Compilation



Lexical Analyzer

- converts the input program into a sequence of Tokens.
- can be implemented with the *Deterministic finite Automata*.

Lexical Analyzer



Programmer's View

```
FILE *yyin;
char *yytext;
main(int argc, char *argv[]){
int token;
if (argc != 2){

}else{
        yyin = fopen(argv[1], "r");
        while(!feof(yyin)){
                token = yylex();
                printf("%d", token);
        }
        fclose(yyin);
    }
}
```

```
int yylex(){
    ...
    ...
}
```

Loop and switch Approach

```
/* Single caharacter lexemes */
#define LPAREN_TOK '('
#define GT_TOK '>'
#define RPAREN_TOK ')'
#define EQ_TOK '='
#define MINUS_TOK '-'
#define SEMICOLON_TOK ';'
/*.....................
.....................*/
/* Reserved words */
#define WHILE_TOK 256
/*.....................
.....................*/
/* Identifier, constants..*/
#define ID_TOK 350
#define INTCONST 351
/*.....................
.....................*/
```

Loop and switch Approach

```
int yylex(){
char ch;
if(yyin == null){
        yyin = stdin;
}
ch = getc(fp);
while (isspace(ch))
        ch = getc(fp);
        // discard any white space
switch(ch){
         case ';': case ',': case '=': //... and other single
             character tokens
         yytext[0] = ch;
         yyleng = 1;
         return ch; // ASCII value is used as token value
         case 'A': case 'B': case 'C': // ... and other upper
             letters .
         case 'a': case 'b': case 'c': // ... and other lower
             letters
     }
}
```

Assignment

Implement a lexical analyzer for the following types of tokens:

- Arithmetic, Relational, Logical, Bitwise and Assignment Operators of C.
- Reserved words: int, float, char, for, while, if and else
- Identifier.
- Integer Constants.
- Parentheses, Curly braces

Follow the ideas of yytext, yyleng, etc. as stated in the study material.