

# Compiler Theory

The 4<sup>th</sup> Program Assignment – LR(1) Parser  
21900628 – Sechang Jang

## Index

Theoretical Foundations .....	2
BNF .....	3

# Theoretical Foundations

An LR parser is a type of bottom-up parser used for analyzing deterministic context-free languages. It's called "LR" because it reads input from Left to right and produces a Rightmost derivation in reverse. There are several types of LR parsers, such as SLR (Simple LR), LALR (Look-Ahead LR), and Canonical LR. Here's a brief overview of the key concepts:

**Bottom-Up Parsing:** LR parsers work by starting from the input symbols and trying to construct the parse tree from the leaves (input symbols) up to the root (start symbol). This contrasts with top-down parsers like LL parsers, which start from the start symbol and try to rewrite it to match the input string.

**Shift-Reduce Parsing:** The main mechanism of LR parsers involves two operations:

**Shift:** Read the next input symbol and push it onto a stack.

**Reduce:** Replace a string of symbols on the stack (which matches the right-hand side of a production rule) with the left-hand side non-terminal of that rule.

**Parse Table:** The parsing decisions are driven by a precomputed table, which guides whether to shift, reduce, or accept the input. This table is constructed based on the grammar of the language being parsed.

**States and Items:** The parser maintains a state, which represents the current position in the parsing process. Each state corresponds to a set of "items," where each item indicates how much of a production rule has been seen. For example, an item might look like this:  $A \rightarrow \alpha \bullet \beta$ , indicating that the parser has seen  $\alpha$  and expects to see  $\beta$  next.

**Lookahead:** LR parsers use lookahead tokens to decide on parsing actions. The lookahead token is the next input symbol that the parser examines to make decisions. Different types of LR parsers use different amounts of lookahead. For example, SLR parsers use a simple one-symbol lookahead, whereas Canonical LR parsers use a more sophisticated method.

## Types of LR Parsers

**SLR (Simple LR):** Simplest form of LR parsing, using basic follow sets for lookahead decisions.

**LALR (Look-Ahead LR):** A more powerful and commonly used form that merges states with identical core items but different lookaheads, reducing the size of the parse table.

**Canonical LR:** The most powerful form, using full lookahead sets for each state, which results in larger parse tables but can handle a broader range of grammars.

## Advantages and Disadvantages

### Advantages:

Can handle a wide range of context-free grammars, including many that are not suitable for LL parsers.  
More powerful than LL parsers, allowing for left-recursive grammars.

### Disadvantages:

More complex to implement and understand compared to LL parsers.  
The parse tables can be large, especially for Canonical LR parsers.

In summary, LR parsers are a robust and powerful tool for syntax analysis in compilers, capable of handling a broad spectrum of grammars with deterministic parsing.

# BNF

PROGRAM -> program IDENTIFIER BLOCK  
IDENTIFIER -> identifier  
BLOCK -> begin STATEMENT end  
STATEMENT -> IF\_BLOCK STMT | ASSIGNMENT STMT | PRINT STMT | DISPLAY STMT | FOR\_STMT STMT | WHILE\_STMT STMT  
IF\_BLOCK -> IF\_STMT | IF\_STMT ELSE\_IF\_STMT | IF\_STMT ELSE\_STMT | IF\_STMT ELSE\_IF\_STMT ELSE\_STMT  
IF\_STMT -> if ( COMPARISON\_STMT ) BLOCK  
ELSE\_IF\_STMT -> else\_if ( COMPARISON\_STMT ) BLOCK NEW\_ELSE\_IF  
NEW\_ELSE\_IF -> ELSE\_IF\_STMT | /eps  
ELSE\_STMT -> else BLOCK  
TYPE -> int | integer | /eps  
COMPARISON\_STMT -> NUMORID COMPARISON\_OP NUMORID  
ASSIGNMENT -> TYPE ASSIGN ;  
ASSIGN -> IDENTIFIER EQUAL ASSIGN'  
ASSIGN' -> , IDENTIFIER EQUAL ASSIGN' | /eps  
EQUAL -> = ASSIGNED ASSIGN'  
ASSIGNED -> NUMERIC ASSIGNED'  
ASSIGNED' -> ARITHOP NUMORID ASSIGNED' | /eps  
NUMORID -> NUMBER | IDENTIFIER  
PRINT -> print\_line ( STRING ) ;  
STRING -> string | identifier  
NUMBER -> number\_literal  
COMPARISON\_OP -> < | <= | != | <= | >=  
ARITHOP -> + | - | / | \*  
INCREMENTALOP -> ++ | /eps  
DISPLAY -> display ( string ) ;  
FOR\_STMT -> for ( ASSIGNMENT COMPARISON\_STMT ; IDENTIFIER INCREMENTALOP ) BLOCK  
WHILE\_STMT -> while ( COMPARISON\_STMT ) BLOCK

First set

First (PROGRAM) = {program}  
First (IDENTIFIER) = {identifier}  
First (BLOCK) = {begin}  
First (TYPE) = {integer, int, /eps}  
First (STATEMENT) = {if, identifier, \$, printline, display, for, while, int, integer}  
First (STMT) = {if, alphabet, \$, printline, display, for, while, int, integer, /eps}  
First (IF\_BLOCK) = {if}  
First (IF\_STMT) = {if}  
First (ELSE\_IF) = {else\_if}  
First (NEW\_ELSE\_IF\_STMT) = {else\_if, /eps}  
First (ELSE\_STMT) = {else}  
First (COMPARISON\_STMT) = {identifier}  
First (ASSIGNMENT) = {alphabet, int, integer}  
First (ASSIGN) = {alphabet, int, integer}  
First (ASSIGN') = {,}  
First (ASSIGNED) = {identifier, number}  
First (ASSIGNED') = {+, -, \*, /, /eps}  
First (PRINT) = {print\_line}  
First (NUMBER) = {number\_literal}  
First (STRING) = {string, identifier}

First (COMPARISONOP)={<, <=, !=, >, >=}  
 First (ARITHOP)={+, -, \*, /}  
 First (INCREMENTALOP)={++, /eps}  
 First (DISPLAY)={display}  
 First (FOR\_STMT)={for}  
 First (WHILE\_STMT)={while}  
 First (Equal)={=}  
 First (NUMORID)={number\_literal | identifier}  
 Follow Set  
 Follow (PROGRAM)={identifier}  
 Follow (DIGIT)={digits, ., \_begin, identifier, ;, , , >, <, >=, <=, ==, !=, +, -, \*, /, ++}  
 Follow (IDENTIFIER)={begin, identifier, ;, , , >, <, >=, <=, ==, !=, +, -, \*, /}  
 Follow (BLOCK)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (TYPE)={identifier}  
 Follow (STATEMENT)={end}  
 Follow (STMT)={end}  
 Follow (IF\_BLOCK)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (IF\_STMT)={identifier, end, display, for, while, break, if, int, integer, print\_line, else\_if, else}  
 Follow (ELSE\_IF)={identifier, end, display, for, while, break, if, int, integer, print\_line, else}  
 Follow (NEW\_ELSE\_IF\_STMT)={identifier, end, display, for, while, break, if, int, integer, print\_line, else}  
 Follow (ELSE\_STMT)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (COMPARISON\_STMT)={}, ;}  
 Follow (ASSIGNMENT)={identifier, end, display, for, while, break, if, int, integer, print\_line, ;}  
 Follow (ASSIGN)={;}  
 Follow (ASSIGN')={alphabet, \$}  
 Follow (ASSIGNED)={;, ,}  
 Follow (ASSIGNED')={;, ,}  
 Follow (PRINT)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (NUMBER)={identifier, >, <, >=, <=, ==, !=, +, -, \*, /, ++, )}  
 Follow (STRING)={})  
 Follow (COMPARISONOP)={digit, identifier}  
 Follow (ARITHOP)={digit, identifier}  
 Follow (INCREMENTALOP)={})  
 Follow (DISPLAY)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (FOR\_STMT)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (WHILE\_STMT)={identifier, end, display, for, while, break, if, int, integer, print\_line}  
 Follow (Equal)={identifier, ;}  
 Follow (NUMORID)={identifier, >, <, >=, <=, ==, !=, +, -, \*, /, ++, )}

# Parsing Table

CSV file attached

# Outputs

```
PS D:\Handong\4-1\compiler Theory\HW4> javac -d bin *.java
PS D:\Handong\4-1\compiler Theory\HW4> java -cp bin LRparser .\testCaseLRparser\test1.txt
[0, program, 2]
[0, program, 2, Test1, 4]
[0, program, 2, IDENTIFIER, 3]
[0, program, 2, IDENTIFIER, 3, begin, 6]
[0, program, 2, IDENTIFIER, 3, begin, 6, int, 21]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, Time.10.24, 43]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, 2206, 83]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMBER, 81]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMORID, 80]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMORID, 80, ASSIGNED', 104]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, Hour.In.Day, 43]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, 0, 83]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMBER, 81]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMORID, 80]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMORID, 80, ASSIGNED', 104]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, $Minute.In.Hour, 43]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, EQUAL, 138]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, EQUAL, 138, ASSIGN', 143]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ASSIGN', 102]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, EQUAL, 138]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ASSIGN', 102]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, EQUAL, 61]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, EQUAL, 61, ASSIGN', 77]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, ASSIGN, 41]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, ASSIGN, 41, ,, 60]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, Time.10.24, 73]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, IDENTIFIER, 71]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, NUMORID, 69]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, NUMORID, 69, <, 93]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, NUMORID, 69, COMPARISON.OP, 91]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, NUMORID, 69, COMPARISON.OP, 91, 10, 119]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, if, 20, (, 48, NUMORID, 69, COMPARISON.OP, 91, NUMBER, 117]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16, (, 44]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16, (, 44, "Good evening.", 64]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16, (, 44, STRING, 63]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16, (, 44, STRING, 63, ), 85]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, print_line, 16, (, 44, STRING, 63, ), 85, ,, 110]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, PRINT, 10]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, PRINT, 10, STMT, 33]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, STATEMENT, 76]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, begin, 59, STATEMENT, 76, end, 100]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, else, 40, BLOCK, 58]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_STMT, 14, ELSE_IF_STMT, 37, ELSE_STMT, 56]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_BLOCK, 25]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, IF_BLOCK, 25, STMT, 49]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, STMT, 32]
[0, program, 2, IDENTIFIER, 3, begin, 6, STATEMENT, 7]
[0, program, 2, IDENTIFIER, 3, begin, 6, STATEMENT, 7, end, 23]
[0, program, 2, IDENTIFIER, 3, BLOCK, 5]
[0, PROGRAM, 1]
Accept
```







```

PS D:\Handong\4-1\Compiler Theory\HW4> java -cp bin LRparser .\testCaseLRparser\test2error.txt
[0, program, 2]
[0, program, 2, Test2, 4]
[0, program, 2, IDENTIFIER, 3]
[0, program, 2, IDENTIFIER, 3, begin, 6]
[0, program, 2, IDENTIFIER, 3, begin, 6, int, 21]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, i.Value.$, 43]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, 0, 83]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMBER, 81]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMORID, 80]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, NUMORID, 80, ASSIGNED', 104]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, $Minute.In.Hour, 43]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, 1, 83]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMBER, 81]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMORID, 80]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, NUMORID, 80, ASSIGNED', 104]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, ASSIGNED, 79]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, =, 62, ASSIGNED, 79, ASSIGN', 102]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, EQUAL, 138]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ,, 103, IDENTIFIER, 125, EQUAL, 138, ASSIGN', 143]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, =, 62, ASSIGNED, 79, ASSIGN', 102]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, EQUAL, 61]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, IDENTIFIER, 42, EQUAL, 61, ASSIGN', 77]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, ASSIGN, 41]
[0, program, 2, IDENTIFIER, 3, begin, 6, TYPE, 15, ASSIGN, 41, ,, 60]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, while, 19]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, while, 19, (, 47]
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, while, 19, (, 47, i.Value.$, 73]
Un definiend
[0, program, 2, IDENTIFIER, 3, begin, 6, ASSIGNMENT, 9, while, 19, (, 47, i.Value.$, 73, null, -1]
Parsing error

```



