<u>HackingOff</u>

- <u>Home</u>
- <u>Blog</u>
- Compiler Construction Toolkit
 - Overview
 - <u>Scanner Generator</u>
 - Regex to NFA & DFA
 - NFA to DFA
 - BNF to First, Follow, & Predict sets
 - 0

0

- <u>Parser Generator Overview</u>
- <u>LL(1) Parser Generator</u>
- o LR(0) Parser Generator
- SLR(1) Parser Generator

LL(1) Parser Generator. First, Follow, & Predict Sets. Table

Overview

Given a grammar in (limited) EBNF, this online tool automatically calculates the first, follow, and predict sets. It also generates LL(1) parser tables from the predict sets, as done by <u>Fischer & LeBlanc</u>.

The sets are shown in two formats: human-friendly tables, and machine-friendly JSON dumps. Use a JSON library to read those tables into your programs to rapidly iterate on your parser's design.

- First Set
- Follow Set
- Predict Set
- LL(1) Table
- Grammar Input

First Set

```
Non-Terminal Symbol
                                                                           First Set
if
                                if
else
                                else
while
                                while
break
                                break
                                }
{\tt continue}
                                continue
return
                                return
&&
                                &&
                                ==
!=
                                !=
                                <
                                <=
                                >=
identifier
                                identifier
                                (
INT-LITERAL
                                INT-LITERAL
BOOL-LITERAL
                                BOOL-LITERAL
```

http://hackingoff.com/compilers/II-1-parser-generator

```
var
                              var
class
                              class
const
                              const
int
                              int
bool
                              hoo1
if-statement
                              if
while-statement
                              while
                              break
break-statement
compound-statement
                              \epsilon, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class
statement-list
continue-statement
                              continue
return-statement
                              return
                              ;, ε, -, !, ++, --
expression-statement
                               ε, -, !, ++, --
expression-list
class-body
variable-declaration-list
                              ε, var
assignment-operator
                              =, *=, /=, +=, -=
condition-or-expression-tail \epsilon, -, !, ++, --
condition-and-expression-tail &&, ε
                              ε, ==, !=
equality-expression-tail
                              \epsilon , <, <=, >, >=
rel-expression-tail
additive-expression-tail
                              ε, +, -
m-d-expression-tail
                              ε. *. /
                              -, !, ++, --
u-expression
post-expression-tail
                              ., ++, --, ε
                              identifier, (, INT-LITERAL, BOOL-LITERAL
primary-expression
para-list
                              (
proper-para-list-tail
                              , ,
arg-list
                              (
proper-arg-list-tail
                              ,, ε
                              identifier
function-declaration
variable-declaration
                              var
class-declaration
                              class
constant-declaration
                              const
init-expression
type-annotation
type
                              int, bool
                              ε, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class
top-level
                              {, while, continue, if, return, break, ;, ε, -, !, ++, --, identifier, var, const, class
statement
m-d-expression
                              -. !. ++. --
post-expression
                              identifier, (, INT-LITERAL, BOOL-LITERAL
para-declaration
                              int, bool
declaration-statement
                              identifier, var, const, class
additive-expression
                              -, !, ++, --
                              int, bool
proper-para-list
                              -, !, ++, --
rel-expression
equality-expression
                              -, !, ++, --
condition-and-expression
                             -, !, ++, --
condition-or-expression
                              -, !, ++, --
assignment-expression
                              -, !, ++, --
expression
arg
                              -, !, ++, --
proper-arg-list
                              -, !, ++, --
```

Follow Set

```
Non-Terminal Symbol
                                                                         Follow Set
statement
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
if-statement
                              $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class,
while-statement
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class,
break-statement
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class,
compound-statement
                             else, $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
{\tt statement-list}
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
continue-statement
return-statement
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
expression-statement
                             $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
expression-list
class-body
```

```
variable-declaration-list
expression
                               ), ;, {, ,
assignment-expression
                              ), ;, {, ,
assignment-operator
condition-or-expression
                              ), ;, {, ,
condition-or-expression-tail ), ;, {, ,
condition-and-expression
                               -, !, ++, --, ), ;, {, ,
condition-and-expression-tail-, !, ++, --, ), ;, {, ,
equality-expression
                              ==, !=, &&, -, !, ++, --, ), ;, {, ,
                               ==, !=, &&, -, !, ++, --, ), ;, {, ,
equality-expression-tail
                              ==, !=, &&, -, !, ++, --, ), ;, {, ,
rel-expression
                               ==, !=, &&, -, !, ++, --, ), ;, {, ,
rel-expression-tail
                               <, <=, >, >=, ==, !=, &&, -, !, ++, --, ), ;, {, ,
additive-expression
                               <, <=, >, >=, ==, !=, &&, -, !, ++, --, ), ;, {, ,
additive-expression-tail
                               +, -, <, <=, >, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
m-d-expression
                               +, -, <, <=, >, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
m-d-expression-tail
                               *, /, +, -, <, <=, >, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
u-expression
                               *, /, +, -, <, <=, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
post-expression
                               *, /, +, -, <, <=, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
post-expression-tail
                               ., ++, --, *, /, +, -, <, <=, >, >=, ==, !=, &&, !, ), ;, {, ,
primary-expression
para-list
proper-para-list
                              )
                              )
proper-para-list-tail
para-declaration
                               ., ++, --, *, /, +, -, <, <=, >, >=, ==, !=, &&, !, ), ;, {, ,
arg-list
proper-arg-list
                               )
proper-arg-list-tail
arg
                               ,,)
declaration-statement
                               $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
                               $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class,
function-declaration
variable-declaration
                               var, $, {, while, continue, if, return, break, ;, -, !, ++, --, identifier, const, class,
                               , {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, } , {, while, continue, if, return, break, ;, -, !, ++, --, identifier, var, const, class, }
class-declaration
constant-declaration
init-expression
type-annotation
                               identifier, ;
type
top-level
```

Predict Set

```
Expression
                                                                                                        Predict
1 statement → compound-statement
2 statement → if-statement
                                                                         if
                                                                         while
3 statement → while-statement
4 statement → break-statement
                                                                         hreak
                                                                         continue
5 statement → continue-statement
6 statement \rightarrow return-statement
                                                                         return
                                                                         ;, -, !, ++, --
7 statement → expression-statement
                                                                         identifier, var, const, class
8 statement → declaration-statement
  if-statement → if expression compound-statement else compound-
  statement
10 while-statement \rightarrow while expression compound-statement
                                                                         while
11 break-statement → break ;
                                                                         break
12 \text{ compound-statement} \rightarrow \{ \text{ statement-list } \}
                                                                         {
13 statement-list \rightarrow \epsilon
                                                                         {, while, continue, if, return, break, ;, -, !, ++, --,
14 statement-list → statement statement-list
                                                                         identifier, var, const, class
15 continue-statement → continue ;
                                                                         continue
16 \, \mathrm{return}\text{-statement} \rightarrow \mathrm{return} \, \mathrm{expression};
                                                                         return
17 return-statement → return ;
                                                                         return
18 expression-statement → expression-list ;
                                                                         -, !, ++, --, ;
19 expression-list → expression
                                                                         -, !, ++, -
20 expression-list → ε
21 class-body → { variable-declaration-list }
22 variable-declaration-list → variable-declaration variable-
                                                                         var
  declaration-list
23 variable-declaration-list → ε
24 expression → assignment-expression
                                                                         -, !, ++, --
25 assignment-expression → condition-or-expression
                                                                         -, !, ++, -
26 assignment-operator → =
```

```
27 assignment-operator → *=
                                                                            /=
28 assignment-operator → /=
29 assignment-operator \rightarrow +=
                                                                            +=
30 assignment-operator \rightarrow -=
                                                                            -=
31 condition—or—expression → condition—and—expression condition—
                                                                            -, !, ++, --
  or-expression-tail
32 condition-or-expression-tail \rightarrow \epsilon
                                                                            ), ;, {, ,
33 condition-or-expression-tail → condition-and-expression
                                                                            -, !, ++, -
  condition-or-expression-tail
_{34} condition-and-expression \rightarrow equality-expression condition-and-expression-tail
                                                                            -, !, ++, --
35 condition—and—expression—tail → && equality—expression
                                                                            &&
  equality-expression-tail
36 condition-and-expression-tail → ε
                                                                            -, !, ++, --, ), ;, {, ,
37 equality-expression \rightarrow rel-expression equality-expression-tail
                                                                            -, !, ++, -
38 equality-expression-tail → ε
                                                                            ==, !=, &&, -, !, ++, --, ), ;, {, ,
_{39} equality-expression-tail \rightarrow == rel-expression equality-
   expression-tail
40 equality-expression-tail \rightarrow != rel-expression equality-expression-tail
                                                                            !=
41 rel-expression → additive-expression rel-expression-tail
                                                                            -, !, ++, --
42 rel-expression-tail → ε
                                                                            ==, !=, &&, -, !, ++, --, ), ;, {, ,
43 rel-expression-tail → < additive-expression rel-expression-tail <
44 rel-expression-tail \rightarrow <= additive-expression rel-expression-tail
45 rel-expression-tail → > additive-expression rel-expression-tail >
46 rel-expression-tail → >= additive-expression rel-expression-
47 additive-expression → m-d-expression additive-expression-tail
48 additive-expression-tail \rightarrow \epsilon
                                                                            <, <=, >, >=, ==, !=, &&, -, !, ++, --, ), ;, {, ,
49 additive-expression-tail → + m-d-expression additive-
  expression-tail
_{50} additive-expression-tail \rightarrow - m-d-expression additive-expression-tail
                                                                            -, !, ++, --
51\,\text{m-d-expression} \rightarrow \text{u-expression} \,\text{m-d-expression-tail}
                                                                               -, <, <=, >, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
52 m-d-expression-tail → ε
53 m-d-expression-tail \rightarrow * u-expression m-d-expression-tail
54\,\mathrm{m-d-expression-tail} \rightarrow / u-expression m-d-expression-tail
55 \text{ u-expression} \rightarrow - \text{ u-expression}
56 \text{ u-expression} \rightarrow ! \text{ u-expression}
57 \text{ u-expression} \rightarrow \text{++ post-expression}
58 u-expression → -- post-expression
59 post-expression → primary-expression
                                                                            identifier, (, INT-LITERAL, BOOL-LITERAL
60 post-expression → primary-expression post-expression-tail
                                                                            identifier, (, INT-LITERAL, BOOL-LITERAL
61 post-expression-tail \rightarrow . identifier post-expression-tail
62 post-expression-tail → ++ post-expression
63 post-expression-tail → -- post-expression
64 post-expression-tail → ε
                                                                            *, /, +, -, <, <=, >, >=, ==, !=, &&, !, ++, --, ), ;, {, ,
65 primary-expression → identifier
                                                                            identifier
                                                                            identifier
^{66} primary-expression \rightarrow identifier arg-list
67 primary-expression \rightarrow ( expression )
68 primary-expression \rightarrow INT-LITERAL
                                                                            INT-LITERAL
69 primary-expression → BOOL-LITERAL
                                                                            BOOL-LITERAL
70 para-list → ()
                                                                            (
71 para-list → ( proper-para-list )
                                                                            (
72 proper-para-list → para-declaration proper-para-list-tail
                                                                            int, bool
_{73} proper-para-list-tail \rightarrow , para-declaration proper-para-list-
  tai1
74 proper-para-list-tail → ε
                                                                            )
75 para-declaration → type identifier
                                                                            int, bool
76 arg-list \rightarrow ()
                                                                            (
77 arg-list → ( proper-arg-list )
78 proper-arg-list → arg proper-arg-list-tail
                                                                               !, ++, --
79 proper-arg-list-tail \rightarrow , arg proper-arg-list-tail
80 proper-arg-list-tail \rightarrow \epsilon
                                                                            -, !, ++, --
81 arg → expression
82 declaration-statement \rightarrow function-declaration
                                                                            identifier
83 declaration-statement \rightarrow constant-declaration
                                                                            const
84 declaration-statement \rightarrow variable-declaration
                                                                            var
85 declaration-statement \rightarrow class-declaration
                                                                            class
86 function-declaration → identifier para-list compound-statement identifier
87 variable-declaration \rightarrow var identifier init-expression;
```

```
88 variable-declaration \rightarrow var identifier type-annotation;
89 class-declaration \rightarrow class identifier init-expression;
                                                                            class
90 class-declaration \rightarrow class identifier type-annotation ;
                                                                            class
91 constant-declaration \rightarrow const identifier init-expression ;
                                                                            const
92 constant-declaration → const identifier type-annotation ;
                                                                            const
93 init-expression \rightarrow = expression
94 type-annotation → : type
95 type → int
                                                                            int
96 type → bool
                                                                            boo1
                                                                             {, while, continue, if, return, break, ;, -, !, ++, --,
97 top-level \rightarrow statement top-level
                                                                            identifier, var, const, class
98 top-level \rightarrow \epsilon
```

LL(1) Parsing Table

On the LL(1) Parsing Table's Meaning and Construction

- The top row corresponds to the columns for all the potential terminal symbols, augmented with \$ to represent the end of the parse.
- The leftmost column and second row are all zero filled, to accommodate the way Fischer and LeBlanc wrote their parser's handling of abs().
- The remaining rows correspond to production rules in the original grammar that you typed in.
- Each entry in that row maps the left-hand-side (LHS) of a production rule onto a line-number. That number is the line in which the LHS had that specific column symbol in its predict set.
- If a terminal is absent from a non-terminal's predict set, an error code is placed in the table. If that terminal is in follow(that non-terminal), the error is a POP error. Else, it's a SCAN error.

```
POP error code = # of predict table productions + 1
SCAN error code = # of predict table productions + 2
```

In practice, you'd want to tear the top, label row off of the table and stick it in a comment, so that you can make sense of your table. The remaining table can be used as is.

LL(1) Parsing Table as JSON (for Easy Import)

```
 \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}  \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}{l}   \begin{tabular}
```

LL(1) Parsing Push-Map (as JSON)

This structure maps each production rule in the expanded grammar (seen as the middle column in the predict table above) to a series of states that the LL parser pushes onto the stack.

```
        \{ "1":[5], "2":[2], "3":[3], "4":[4], "5":[7], "6":[8], "7":[9], "8":[40], "9":[5, -2, 5, 13, -1], "10":[5, 13, -3], "11":[-5, -4], "12": [-7, 6, -6], "14":[6, 1], "15":[-5, -8], "16":[-5, 13, -9], "17":[-5, -9], "18":[-5, 10], "19":[13], "21":[-7, 12, -6], "22":[12, 42], "24":[14], "25": [16], "26":[-10], "27":[-11], "28":[-12], "29":[-13], "30":[-14], "31":[17, 18], "33":[17, 18], "34":[19, 20], "35":[21, 20, -15], "37": [21, 22], "39":[21, 22, -16], "40":[21, 22, -17], "41":[23, 24], "43":[23, 24, -18], "44":[23, 24, -19], "45":[23, 24, -20], "46":[23, 24, -21], "47": [25, 26], "49":[25, 26, -22], "50":[25, 26, -23], "51":[27, 28], "53":[27, 28, -24], "54":[27, 28, -25], "55":[28, -23], "56":[28, -26], "57": [29, -27], "58":[29, -28], "59":[31], "60":[30, 31], "61":[30, -30, -29], "62":[29, -27], "63":[29, -28], "65":[-30], "66":[36, -30], "67": [-32, 13, -31], "68":[-33], "69":[-34], "70":[-32, 31], "71":[-32, 33, -31], "72":[34, 35], "73":[34, 35, -35], "75":[-30, 47], "76":[-32, -31], "77": [-32, 37, -31], "78":[38, 39], "79":[38, 39, -35], "81":[13], "82":[41], "83":[44], "84":[42], "85":[43], "86":[5, 32, -30], "87": [-5, 45, -30, -36], "88":[-5, 46, -30, -36], "89":[-5, 45, -30, -37], "90":[-5, 46, -30, -37], "91":[-5, 45, -30, -38], "92":[-5, 46, -30, -38], "93": [13, -10], "94":[47, -39], "95":[-40], "96":[-41], "97":[48, 1] \}
```

Feed me your delicious grammar, mortal.

```
statement ->
compound-statement
if-statement | while-
statement | break-
statement
            continue
statement
            return-
statement
expression-statement
declaration-
statement
if-statement -> if
expression compound-
statement else
compound-statement
while-statement
while expression
compound-statement
break-statement ->
break
compound-statement ->
Generate LL(1) Parsing Table, First Set, Follow Set, & Predict Set
```

Grammar Specification Requirements

Productions use the following format:

```
Goal -> A
A -> ( A ) | Two
Two -> a
Two -> b
```

- "->" separates the non-terminal on the left-hand-side from what it produces.
- x -> y | z is EBNF short-hand for x -> y | x is EBNF short-hand for x -> y

Be certain to place spaces between things you don't want read as one symbol. (A) eq (A)

About This Tool

Intended Audience

Computer science students & autodidacts studying compilers or parsing.

Purpose

This tool provides rapid feedback loop for learning about grammars. How?

- Rapid visualization of grammars enables convenient tweaking. Botched a production? No problem; tweak it and everything's spit back out.
- Ability to dump LR(0) and SLR(1) tables. Helps with manual parse tracing and hand-writing parsers.
- Assisting with coursework.

Underlying Theory

How to draw NFAs for SLR(0) and LR(1) grammars. Want to learn how it works or how to do it by hand? Read that.

© HackingOff.com 2012