

CSE 3302
Programming Languages



Syntax (cont.)

Chengkai Li, Weimin He
Spring 2008

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

1

What is Parsing?



- Given a grammar and a token string:**
 - determine if the grammar can generate the token string?
 - i.e., is the string a legal program in the language?
- In other words, to construct a parse tree for the token string.**

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

2

What's significant about parse tree?



A parse tree gives a unique syntactic structure

- Leftmost, rightmost derivation
- There is only one leftmost derivation for a parse tree, and symmetrically only one rightmost derivation for a parse tree.

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

3

Example



$\text{expr} \rightarrow \text{expr} + \text{expr} \mid \text{expr} * \text{expr} \mid (\text{expr}) \mid \text{number}$
 $\text{number} \rightarrow \text{number digit} \mid \text{digit}$
 $\text{digit} \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

parse tree <pre> expr +-- expr +-- number +-- digit +-- 3 </pre>	leftmost derivation <pre> expr +-- expr +-- number +-- digit +-- 3 </pre>
---	--

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

4

What's significant about parse tree?



A parse tree has a unique meaning, thus provides basis for semantic analysis.

(Syntax-directed semantics: semantics are attached to syntactic structure.)

expr
 $\text{expr} \rightarrow \text{expr1} + \text{expr2}$
 $\text{expr1.val} = \text{expr1.val} + \text{expr2.val}$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

5

Relationship among language, grammar, parser



- Chomsky Hierarchy
http://en.wikipedia.org/wiki/Chomsky_hierarchy
- A language can be described by multiple grammars, while a grammar defines one language.
- A grammar can be parsed by multiple parsers, while a parser accepts one grammar, thus one language.
- Should design a language that allows simple grammar and efficient parser
- For a language, we should construct a grammar that allows fast parsing
- Given a grammar, we should build an efficient parser

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

6

Ambiguity



- **Ambiguous grammar:** There can be multiple parse trees for the same sentence (program)
 - In other words, multiple leftmost derivations.
- **Why is it bad?**
 - Multiple meanings

Lecture 4 – Syntax (Cont.), Spring 2008

CSE3302 Programming Languages, UT-Arlington

©Chengkai Li, Weimin He, Weimin He 2008

7

Ambiguity



- **Was this ambiguous?**

$$\begin{aligned} \text{number} &\rightarrow \text{number digit} \mid \text{digit} \\ \text{digit} &\rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$

- **How about this?**

$$\text{expr} \rightarrow \text{expr} - \text{expr} \mid \text{number}$$

Lecture 4 – Syntax (Cont.), Spring 2008

CSE3302 Programming Languages, UT-Arlington

©Chengkai Li, Weimin He, Weimin He 2008

8

Deal with Ambiguity



- **Unambiguous Grammar**
 - Rewrite the grammar to avoid ambiguity.

Lecture 4 – Syntax (Cont.), Spring 2008

CSE3302 Programming Languages, UT-Arlington

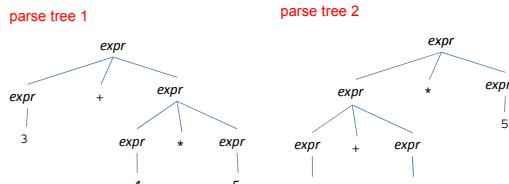
©Chengkai Li, Weimin He, Weimin He 2008

9

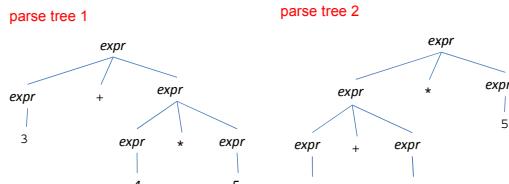
Example of Ambiguity: Precedence

$$\begin{aligned} \text{expr} &\rightarrow \text{expr} + \text{expr} \mid \text{expr} * \text{expr} \mid (\text{expr}) \mid 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$
Two different parse trees for expression $3+4*5$

parse tree 1



parse tree 2

Lecture 4 – Syntax (Cont.), Spring 2008
CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

10

Eliminating Ambiguity for Precedence



- Establish “precedence cascade”: using different structured names for different constructs, adding grammar rules.
 - Higher precedence : lower in cascade

$$\text{expr} \rightarrow \text{expr} + \text{expr} \mid \text{expr} * \text{expr} \mid (\text{expr}) \mid \text{number}$$


$$\text{expr} \rightarrow \text{expr} + \text{expr} \mid \text{term}$$

$$\text{term} \rightarrow \text{term} * \text{term} \mid (\text{expr}) \mid \text{number}$$

Lecture 4 – Syntax (Cont.), Spring 2008

CSE3302 Programming Languages, UT-Arlington

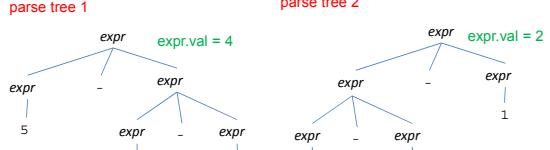
©Chengkai Li, Weimin He, Weimin He 2008

11

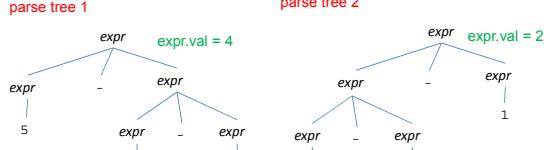
Example of Ambiguity: Associativity

$$\begin{aligned} \text{expr} &\rightarrow \text{expr} - \text{expr} \mid (\text{expr}) \mid 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$
Two different parse trees for expression $5-2-1$

parse tree 1



parse tree 2



- is right-associative, which is against common practice in integer arithmetic

- is left-associative, which is what we usually assume

Lecture 4 – Syntax (Cont.), Spring 2008
CSE3302 Programming Languages, UT-Arlington
©Chengkai Li, Weimin He, Weimin He 2008

12

Associativity



- Left-Associative: + - * /
- Right-Associative: =

What is meant by $a=b=c=1$?

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

13

Eliminating Ambiguity for Associativity



- left-associativity: left-recursion**

$$\begin{aligned} \text{expr} &\rightarrow \text{expr} - \text{expr} \mid (\text{expr}) \mid \text{number} \\ &\quad \xrightarrow{\text{blue arrow}} \\ &\text{expr} \rightarrow \text{expr} - \text{term} \mid \text{term} \\ &\text{term} \rightarrow (\text{expr}) \mid \text{number} \end{aligned}$$

- right-associativity: right-recursion**

$$\begin{aligned} \text{expr} &\rightarrow \text{expr} = \text{expr} \mid a \mid b \mid c \\ &\quad \xrightarrow{\text{blue arrow}} \text{expr} \rightarrow \text{term} = \text{expr} \mid \text{term} \\ &\text{term} \rightarrow a \mid b \mid c \end{aligned}$$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

14

Putting Together



$\text{expr} \rightarrow \text{expr} - \text{expr} \mid \text{expr} / \text{expr} \mid (\text{expr}) \mid \text{number}$
 $\text{number} \rightarrow \text{number digit} \mid \text{digit}$
 $\text{digit} \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

We want to make - left-associative and / has precedence over -

$\text{expr} \rightarrow \text{expr} - \text{term} \mid \text{term}$
 $\text{term} \rightarrow \text{term} / \text{factor} \mid \text{factor}$
 $\text{factor} \rightarrow (\text{expr}) \mid \text{number}$
 $\text{number} \rightarrow \text{number digit} \mid \text{digit}$
 $\text{digit} \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

15

Example of Ambiguity: Dangling-Else



$\text{stmt} \rightarrow \text{if}(\text{expr}) \text{stmt}$
 $\quad \mid \text{if}(\text{expr}) \text{stmt} \text{else} \text{stmt}$
 $\quad \mid \text{other-stmt}$

Two different parse trees for "if(expr) if(expr) stmt else stmt"

parse tree 1

```

    graph TD
      S1[stmt] --> I1[if]
      I1 --> E1(expr)
      I1 --> S1_1[stmt]
      S1_1 --> I1_1[if]
      I1_1 --> E1_1(expr)
      I1_1 --> S1_2[stmt]
      I1_1 --> E1_2[else]
      I1_1 --> S1_3[stmt]
    
```

parse tree 2

```

    graph TD
      S2[stmt] --> I2[if]
      I2 --> E2(expr)
      I2 --> S2_1[stmt]
      S2_1 --> I2_1[if]
      I2_1 --> E2_1(expr)
      I2_1 --> S2_2[stmt]
      I2_1 --> E2_2[else]
      I2_1 --> S2_3[stmt]
    
```

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

16

Eliminating Dangling-Else



$\text{stmt} \rightarrow \text{matched_stmt}$
 $\quad \mid \text{unmatched_stmt}$
 $\text{matched_stmt} \rightarrow \text{if}(\text{expr}) \text{matched_stmt} \text{else} \text{matched_stmt}$
 $\quad \mid \text{other-stmt}$
 $\text{unmatched_stmt} \rightarrow \text{if}(\text{expr}) \text{stmt}$
 $\quad \mid \text{if}(\text{expr}) \text{matched_stmt} \text{else} \text{unmatched_stmt}$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

17

EBNF



- Repetition { }**

$$\begin{aligned} \text{number} &\rightarrow \text{digit} \mid \text{number digit} & \xrightarrow{\text{blue arrow}} \text{number} &\rightarrow \{ \text{digit} \} \\ \text{expr} &\rightarrow \text{expr} - \text{term} \mid \text{term} & \xrightarrow{\text{blue arrow}} \text{expr} &\rightarrow \text{term} \{ - \text{term} \} \end{aligned}$$

- Option []**

$$\begin{aligned} \text{signed-number} &\rightarrow \text{sign} \text{ number} \\ &\quad \mid \text{number} & \xrightarrow{\text{blue arrow}} \text{signed-number} &\rightarrow [\text{sign}] \text{ number} \end{aligned}$$

$\text{if-stmt} \rightarrow \text{if}(\text{expr}) \text{stmt}$ $\xrightarrow{\text{blue arrow}}$ $\text{if-stmt} \rightarrow \text{if}(\text{expr}) \text{stmt} [\text{else} \text{stmt}]$
 $\quad \mid \text{if}(\text{expr}) \text{stmt} \text{else} \text{stmt}$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He, Weimin He 2008

18

Syntax Diagrams



- Written from EBNF, not BNF
- If-statement
(more examples on page 101)

```

if-statement → if "(" expression ")"
if-statement → if "(" statement "else" statement ")"
  
```

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 19

Parsing Techniques



- Intuitive analysis and conclusion
- No formal theorems and rigorous proofs
- More details: compilers, automata theory

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 20

Parsing



- Parsing:**
 - Determine if a grammar can generate a given token string.
 - That is, to construct a parse tree for the token string.
- Two ways of constructing the parse tree**
 - Top-down (from root towards leaves)
 - Can be constructed more easily by hand
 - Bottom-up (from leaves towards root)
 - Can handle a larger class of grammars
 - Parser generators tend to use bottom-up methods

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 21

Top-Down Parser



- Recursive-descent parser:
 - A special kind of top-down parser: single left-to-right scan, with one lookahead symbol.
 - Backtracking (trial-and-error) may happen
- Predictive parser:
 - The lookahead symbol determines which production to apply, without backtracking

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 22

Recursive-Descent Parser



- Types in Pascal

$$\begin{aligned} \text{type} &\rightarrow \text{simple} \mid \text{array} [\text{simple}] \text{ of type} \\ \text{simple} &\rightarrow \text{char} \mid \text{integer} \end{aligned}$$

Input: array [integer] of char

Parse tree:

```

graph TD
    type --- simple1
    type --- array1
    type --- brackets1
    type --- of1
    type --- char1
    array1 --- brackets2
    brackets2 --- simple2
    brackets2 --- integer1
    brackets1 --- brackets2
    brackets2 --- brackets3
    brackets3 --- simple3
    brackets3 --- char1
    brackets1 --- of2
    of2 --- type2
    type2 --- simple4
    simple4 --- minus1
    simple4 --- term1
    minus1 --- expr1
    term1 --- minus2
    term1 --- term2
    minus2 --- expr2
    expr2 --- minus3
    expr2 --- term3
    minus3 --- expr3
    expr3 --- minus4
    expr3 --- term4
    minus4 --- expr4
    expr4 --- minus5
    expr4 --- term5
    minus5 --- expr5
    expr5 --- minus6
    expr5 --- term6
    minus6 --- expr6
    expr6 --- minus7
    expr6 --- term7
    minus7 --- expr7
    expr7 --- minus8
    expr7 --- term8
    minus8 --- expr8
    expr8 --- minus9
    expr8 --- term9
    minus9 --- expr9
    expr9 --- minus10
    expr9 --- term10
    minus10 --- expr10
    expr10 --- minus11
    expr10 --- term11
    minus11 --- expr11
    expr11 --- minus12
    expr11 --- term12
    minus12 --- expr12
    expr12 --- minus13
    expr12 --- term13
    minus13 --- expr13
    expr13 --- minus14
    expr13 --- term14
    minus14 --- expr14
    expr14 --- minus15
    expr14 --- term15
    minus15 --- expr15
    expr15 --- minus16
    expr15 --- term16
    minus16 --- expr16
    expr16 --- minus17
    expr16 --- term17
    minus17 --- expr17
    expr17 --- minus18
    expr17 --- term18
    minus18 --- expr18
    expr18 --- minus19
    expr18 --- term19
    minus19 --- expr19
    expr19 --- minus20
    expr19 --- term20
    minus20 --- expr20
    expr20 --- minus21
    expr20 --- term21
    minus21 --- expr21
    expr21 --- minus22
    expr21 --- term22
    minus22 --- expr23
    expr23 --- minus24
    expr23 --- term25
    minus24 --- expr24
    expr24 --- minus25
    expr24 --- term26
    minus25 --- expr25
    expr25 --- minus26
    expr25 --- term27
    minus26 --- expr26
    expr26 --- minus27
    expr26 --- term28
    minus27 --- expr27
    expr27 --- minus28
    expr27 --- term29
    minus28 --- expr28
    expr28 --- minus29
    expr28 --- term30
    minus29 --- expr29
    expr29 --- minus30
    expr29 --- term31
    minus30 --- expr31
    expr31 --- minus32
    expr31 --- term33
    minus32 --- expr32
    expr32 --- minus33
    expr32 --- term34
    minus33 --- expr34
    expr34 --- minus35
    expr34 --- term36
    minus35 --- expr36
    expr36 --- minus37
    expr36 --- term38
    minus37 --- expr38
    expr38 --- minus39
    expr38 --- term39
    minus39 --- expr40
    expr40 --- minus41
    expr40 --- term42
    minus41 --- expr42
    expr42 --- minus43
    expr42 --- term44
    minus43 --- expr44
    expr44 --- minus45
    expr44 --- term46
    minus45 --- expr46
    expr46 --- minus47
    expr46 --- term48
    minus47 --- expr48
    expr48 --- minus49
    expr48 --- term50
    minus49 --- expr50
    expr50 --- minus51
    expr50 --- term52
    minus51 --- expr52
    expr52 --- minus53
    expr52 --- term54
    minus53 --- expr54
    expr54 --- minus55
    expr54 --- term56
    minus55 --- expr56
    expr56 --- minus57
    expr56 --- term58
    minus57 --- expr58
    expr58 --- minus59
    expr58 --- term60
    minus59 --- expr60
    expr60 --- minus61
    expr60 --- term62
    minus61 --- expr62
    expr62 --- minus63
    expr62 --- term64
    minus63 --- expr64
    expr64 --- minus65
    expr64 --- term66
    minus65 --- expr66
    expr66 --- minus67
    expr66 --- term68
    minus67 --- expr68
    expr68 --- minus69
    expr68 --- term70
    minus69 --- expr70
    expr70 --- minus71
    expr70 --- term72
    minus71 --- expr72
    expr72 --- minus73
    expr72 --- term74
    minus73 --- expr74
    expr74 --- minus75
    expr74 --- term76
    minus75 --- expr76
    expr76 --- minus77
    expr76 --- term78
    minus77 --- expr78
    expr78 --- minus79
    expr78 --- term80
    minus79 --- expr80
    expr80 --- minus81
    expr80 --- term82
    minus81 --- expr82
    expr82 --- minus83
    expr82 --- term84
    minus83 --- expr84
    expr84 --- minus85
    expr84 --- term86
    minus85 --- expr86
    expr86 --- minus87
    expr86 --- term88
    minus87 --- expr88
    expr88 --- minus89
    expr88 --- term90
    minus89 --- expr90
    expr90 --- minus91
    expr90 --- term92
    minus91 --- expr92
    expr92 --- minus93
    expr92 --- term94
    minus93 --- expr94
    expr94 --- minus95
    expr94 --- term96
    minus95 --- expr96
    expr96 --- minus97
    expr96 --- term98
    minus97 --- expr98
    expr98 --- minus99
    expr98 --- term100
    minus99 --- expr100
    expr100 --- minus101
    expr100 --- term102
    minus101 --- expr102
    expr102 --- minus103
    expr102 --- term104
    minus103 --- expr104
    expr104 --- minus105
    expr104 --- term106
    minus105 --- expr106
    expr106 --- minus107
    expr106 --- term108
    minus107 --- expr108
    expr108 --- minus109
    expr108 --- term110
    minus109 --- expr110
    expr110 --- minus111
    expr110 --- term112
    minus111 --- expr112
    expr112 --- minus113
    expr112 --- term114
    minus113 --- expr114
    expr114 --- minus115
    expr114 --- term116
    minus115 --- expr116
    expr116 --- minus117
    expr116 --- term118
    minus117 --- expr118
    expr118 --- minus119
    expr118 --- term120
    minus119 --- expr120
    expr120 --- minus121
    expr120 --- term122
    minus121 --- expr122
    expr122 --- minus123
    expr122 --- term124
    minus123 --- expr124
    expr124 --- minus125
    expr124 --- term126
    minus125 --- expr126
    expr126 --- minus127
    expr126 --- term128
    minus127 --- expr128
    expr128 --- minus129
    expr128 --- term130
    minus129 --- expr130
    expr130 --- minus131
    expr130 --- term132
    minus131 --- expr132
    expr132 --- minus133
    expr132 --- term134
    minus133 --- expr134
    expr134 --- minus135
    expr134 --- term136
    minus135 --- expr136
    expr136 --- minus137
    expr136 --- term138
    minus137 --- expr138
    expr138 --- minus139
    expr138 --- term140
    minus139 --- expr140
    expr140 --- minus141
    expr140 --- term142
    minus141 --- expr142
    expr142 --- minus143
    expr142 --- term144
    minus143 --- expr144
    expr144 --- minus145
    expr144 --- term146
    minus145 --- expr146
    expr146 --- minus147
    expr146 --- term148
    minus147 --- expr148
    expr148 --- minus149
    expr148 --- term150
    minus149 --- expr150
    expr150 --- minus151
    expr150 --- term152
    minus151 --- expr152
    expr152 --- minus153
    expr152 --- term154
    minus153 --- expr154
    expr154 --- minus155
    expr154 --- term156
    minus155 --- expr156
    expr156 --- minus157
    expr156 --- term158
    minus157 --- expr158
    expr158 --- minus159
    expr158 --- term160
    minus159 --- expr160
    expr160 --- minus161
    expr160 --- term162
    minus161 --- expr162
    expr162 --- minus163
    expr162 --- term164
    minus163 --- expr164
    expr164 --- minus165
    expr164 --- term166
    minus165 --- expr166
    expr166 --- minus167
    expr166 --- term168
    minus167 --- expr168
    expr168 --- minus169
    expr168 --- term170
    minus169 --- expr170
    expr170 --- minus171
    expr170 --- term172
    minus171 --- expr172
    expr172 --- minus173
    expr172 --- term174
    minus173 --- expr174
    expr174 --- minus175
    expr174 --- term176
    minus175 --- expr176
    expr176 --- minus177
    expr176 --- term178
    minus177 --- expr178
    expr178 --- minus179
    expr178 --- term180
    minus179 --- expr180
    expr180 --- minus181
    expr180 --- term182
    minus181 --- expr182
    expr182 --- minus183
    expr182 --- term184
    minus183 --- expr184
    expr184 --- minus185
    expr184 --- term186
    minus185 --- expr186
    expr186 --- minus187
    expr186 --- term188
    minus187 --- expr188
    expr188 --- minus189
    expr188 --- term190
    minus189 --- expr190
    expr190 --- minus191
    expr190 --- term192
    minus191 --- expr192
    expr192 --- minus193
    expr192 --- term194
    minus193 --- expr194
    expr194 --- minus195
    expr194 --- term196
    minus195 --- expr196
    expr196 --- minus197
    expr196 --- term198
    minus197 --- expr198
    expr198 --- minus199
    expr198 --- term200
    minus199 --- expr200
    expr200 --- minus201
    expr200 --- term202
    minus201 --- expr202
    expr202 --- minus203
    expr202 --- term204
    minus203 --- expr204
    expr204 --- minus205
    expr204 --- term206
    minus205 --- expr206
    expr206 --- minus207
    expr206 --- term208
    minus207 --- expr208
    expr208 --- minus209
    expr208 --- term210
    minus209 --- expr210
    expr210 --- minus211
    expr210 --- term212
    minus211 --- expr212
    expr212 --- minus213
    expr212 --- term214
    minus213 --- expr214
    expr214 --- minus215
    expr214 --- term216
    minus215 --- expr216
    expr216 --- minus217
    expr216 --- term218
    minus217 --- expr218
    expr218 --- minus219
    expr218 --- term220
    minus219 --- expr220
    expr220 --- minus221
    expr220 --- term222
    minus221 --- expr222
    expr222 --- minus223
    expr222 --- term224
    minus223 --- expr224
    expr224 --- minus225
    expr224 --- term226
    minus225 --- expr226
    expr226 --- minus227
    expr226 --- term228
    minus227 --- expr228
    expr228 --- minus229
    expr228 --- term230
    minus229 --- expr230
    expr230 --- minus231
    expr230 --- term232
    minus231 --- expr232
    expr232 --- minus233
    expr232 --- term234
    minus233 --- expr234
    expr234 --- minus235
    expr234 --- term236
    minus235 --- expr236
    expr236 --- minus237
    expr236 --- term238
    minus237 --- expr238
    expr238 --- minus239
    expr238 --- term240
    minus239 --- expr240
    expr240 --- minus241
    expr240 --- term242
    minus241 --- expr242
    expr242 --- minus243
    expr242 --- term244
    minus243 --- expr244
    expr244 --- minus245
    expr244 --- term246
    minus245 --- expr246
    expr246 --- minus247
    expr246 --- term248
    minus247 --- expr248
    expr248 --- minus249
    expr248 --- term250
    minus249 --- expr250
    expr250 --- minus251
    expr250 --- term252
    minus251 --- expr252
    expr252 --- minus253
    expr252 --- term254
    minus253 --- expr254
    expr254 --- minus255
    expr254 --- term256
    minus255 --- expr256
    expr256 --- minus257
    expr256 --- term258
    minus257 --- expr258
    expr258 --- minus259
    expr258 --- term260
    minus259 --- expr260
    expr260 --- minus261
    expr260 --- term262
    minus261 --- expr262
    expr262 --- minus263
    expr262 --- term264
    minus263 --- expr264
    expr264 --- minus265
    expr264 --- term266
    minus265 --- expr266
    expr266 --- minus267
    expr266 --- term268
    minus267 --- expr268
    expr268 --- minus269
    expr268 --- term270
    minus269 --- expr270
    expr270 --- minus271
    expr270 --- term272
    minus271 --- expr272
    expr272 --- minus273
    expr272 --- term274
    minus273 --- expr274
    expr274 --- minus275
    expr274 --- term276
    minus275 --- expr276
    expr276 --- minus277
    expr276 --- term278
    minus277 --- expr278
    expr278 --- minus279
    expr278 --- term280
    minus279 --- expr280
    expr280 --- minus281
    expr280 --- term282
    minus281 --- expr282
    expr282 --- minus283
    expr282 --- term284
    minus283 --- expr284
    expr284 --- minus285
    expr284 --- term286
    minus285 --- expr286
    expr286 --- minus287
    expr286 --- term288
    minus287 --- expr288
    expr288 --- minus289
    expr288 --- term290
    minus289 --- expr290
    expr290 --- minus291
    expr290 --- term292
    minus291 --- expr292
    expr292 --- minus293
    expr292 --- term294
    minus293 --- expr294
    expr294 --- minus295
    expr294 --- term296
    minus295 --- expr296
    expr296 --- minus297
    expr296 --- term298
    minus297 --- expr298
    expr298 --- minus299
    expr298 --- term300
    minus299 --- expr300
    expr300 --- minus301
    expr300 --- term302
    minus301 --- expr302
    expr302 --- minus303
    expr302 --- term304
    minus303 --- expr304
    expr304 --- minus305
    expr304 --- term306
    minus305 --- expr306
    expr306 --- minus307
    expr306 --- term308
    minus307 --- expr308
    expr308 --- minus309
    expr308 --- term310
    minus309 --- expr310
    expr310 --- minus311
    expr310 --- term312
    minus311 --- expr312
    expr312 --- minus313
    expr312 --- term314
    minus313 --- expr314
    expr314 --- minus315
    expr314 --- term316
    minus315 --- expr316
    expr316 --- minus317
    expr316 --- term318
    minus317 --- expr318
    expr318 --- minus319
    expr318 --- term320
    minus319 --- expr320
    expr320 --- minus321
    expr320 --- term322
    minus321 --- expr322
    expr322 --- minus323
    expr322 --- term324
    minus323 --- expr324
    expr324 --- minus325
    expr324 --- term326
    minus325 --- expr326
    expr326 --- minus327
    expr326 --- term328
    minus327 --- expr328
    expr328 --- minus329
    expr328 --- term330
    minus329 --- expr330
    expr330 --- minus331
    expr330 --- term332
    minus331 --- expr332
    expr332 --- minus333
    expr332 --- term334
    minus333 --- expr334
    expr334 --- minus335
    expr334 --- term336
    minus335 --- expr336
    expr336 --- minus337
    expr336 --- term338
    minus337 --- expr338
    expr338 --- minus339
    expr338 --- term340
    minus339 --- expr340
    expr340 --- minus341
    expr340 --- term342
    minus341 --- expr342
    expr342 --- minus343
    expr342 --- term344
    minus343 --- expr344
    expr344 --- minus345
    expr344 --- term346
    minus345 --- expr346
    expr346 --- minus347
    expr346 --- term348
    minus347 --- expr348
    expr348 --- minus349
    expr348 --- term350
    minus349 --- expr350
    expr350 --- minus351
    expr350 --- term352
    minus351 --- expr352
    expr352 --- minus353
    expr352 --- term354
    minus353 --- expr354
    expr354 --- minus355
    expr354 --- term356
    minus355 --- expr356
    expr356 --- minus357
    expr356 --- term358
    minus357 --- expr358
    expr358 --- minus359
    expr358 --- term360
    minus359 --- expr360
    expr360 --- minus361
    expr360 --- term362
    minus361 --- expr362
    expr362 --- minus363
    expr362 --- term364
    minus363 --- expr364
    expr364 --- minus365
    expr364 --- term366
    minus365 --- expr366
    expr366 --- minus367
    expr366 --- term368
    minus367 --- expr368
    expr368 --- minus369
    expr368 --- term370
    minus369 --- expr370
    expr370 --- minus371
    expr370 --- term372
    minus371 --- expr372
    expr372 --- minus373
    expr372 --- term374
    minus373 --- expr374
    expr374 --- minus375
    expr374 --- term376
    minus375 --- expr376
    expr376 --- minus377
    expr376 --- term378
    minus377 --- expr378
    expr378 --- minus379
    expr378 --- term380
    minus379 --- expr380
    expr380 --- minus381
    expr380 --- term382
    minus381 --- expr382
    expr382 --- minus383
    expr382 --- term384
    minus383 --- expr384
    expr384 --- minus385
    expr384 --- term386
    minus385 --- expr386
    expr386 --- minus387
    expr386 --- term388
    minus387 --- expr388
    expr388 --- minus389
    expr388 --- term390
    minus389 --- expr390
    expr390 --- minus391
    expr390 --- term392
    minus391 --- expr392
    expr392 --- minus393
    expr392 --- term394
    minus393 --- expr394
    expr394 --- minus395
    expr394 --- term396
    minus395 --- expr396
    expr396 --- minus397
    expr396 --- term398
    minus397 --- expr398
    expr398 --- minus399
    expr398 --- term400
    minus399 --- expr400
    expr400 --- minus401
    expr400 --- term402
    minus401 --- expr402
    expr402 --- minus403
    expr402 --- term404
    minus403 --- expr404
    expr404 --- minus405
    expr404 --- term406
    minus405 --- expr406
    expr406 --- minus407
    expr406 --- term408
    minus407 --- expr408
    expr408 --- minus409
    expr408 --- term410
    minus409 --- expr410
    expr410 --- minus411
    expr410 --- term412
    minus411 --- expr412
    expr412 --- minus413
    expr412 --- term414
    minus413 --- expr414
    expr414 --- minus415
    expr414 --- term416
    minus415 --- expr416
    expr416 --- minus417
    expr416 --- term418
    minus417 --- expr418
    expr418 --- minus419
    expr418 --- term420
    minus419 --- expr420
    expr420 --- minus421
    expr420 --- term422
    minus421 --- expr422
    expr422 --- minus423
    expr422 --- term424
    minus423 --- expr424
    expr424 --- minus425
    expr424 --- term426
    minus425 --- expr426
    expr426 --- minus427
    expr426 --- term428
    minus427 --- expr428
    expr428 --- minus429
    expr428 --- term430
    minus429 --- expr430
    expr430 --- minus431
    expr430 --- term432
    minus431 --- expr432
    expr432 --- minus433
    expr432 --- term434
    minus433 --- expr434
    expr434 --- minus435
    expr434 --- term436
    minus435 --- expr436
    expr436 --- minus437
    expr436 --- term438
    minus437 --- expr438
    expr438 --- minus439
    expr438 --- term440
    minus439 --- expr440
    expr440 --- minus441
    expr440 --- term442
    minus441 --- expr442
    expr442 --- minus443
    expr442 --- term444
    minus443 --- expr444
    expr444 --- minus445
    expr444 --- term446
    minus445 --- expr446
    expr446 --- minus447
    expr446 --- term448
    minus447 --- expr448
    expr448 --- minus449
    expr448 --- term450
    minus449 --- expr450
    expr450 --- minus451
    expr450 --- term452
    minus451 --- expr452
    expr452 --- minus453
    expr452 --- term454
    minus453 --- expr454
    expr454 --- minus455
    expr454 --- term456
    minus455 --- expr456
    expr456 --- minus457
    expr456 --- term458
    minus457 --- expr458
    expr458 --- minus459
    expr458 --- term460
    minus459 --- expr460
    expr460 --- minus461
    expr460 --- term462
    minus461 --- expr462
    expr462 --- minus463
    expr462 --- term464
    minus463 --- expr464
    expr464 --- minus465
    expr464 --- term466
    minus465 --- expr466
    expr466 --- minus467
    expr466 --- term468
    minus467 --- expr468
    expr468 --- minus469
    expr468 --- term470
    minus469 --- expr470
    expr470 --- minus471
    expr470 --- term472
    minus471 --- expr472
    expr472 --- minus473
    expr472 --- term474
    minus473 --- expr474
    expr474 --- minus475
    expr474 --- term476
    minus475 --- expr476
    expr476 --- minus477
    expr476 --- term478
    minus477 --- expr478
    expr478 --- minus479
    expr478 --- term480
    minus479 --- expr480
    expr480 --- minus481
    expr480 --- term482
    minus481 --- expr482
    expr482 --- minus483
    expr482 --- term484
    minus483 --- expr484
    expr484 --- minus485
    expr484 --- term486
    minus485 --- expr486
    expr486 --- minus487
    expr486 --- term488
    minus487 --- expr488
    expr488 --- minus489
    expr488 --- term490
    minus489 --- expr490
    expr490 --- minus491
    expr490 --- term492
    minus491 --- expr492
    expr492 --- minus493
    expr492 --- term494
    minus493 --- expr494
    expr494 --- minus495
    expr494 --- term496
    minus495 --- expr496
    expr496 --- minus497
    expr496 --- term498
    minus497 --- expr498
    expr498 --- minus499
    expr498 --- term500
    minus499 --- expr500
    expr500 --- minus501
    expr500 --- term502
    minus501 --- expr502
    expr502 --- minus503
    expr502 --- term504
    minus503 --- expr504
    expr504 --- minus505
    expr504 --- term506
    minus505 --- expr506
    expr506 --- minus507
    expr506 --- term508
    minus507 --- expr508
    expr508 --- minus509
    expr508 --- term510
    minus509 --- expr510
    expr510 --- minus511
    expr510 --- term512
    minus511 --- expr512
    expr512 --- minus513
    expr512 --- term514
    minus513 --- expr514
    expr514 --- minus515
    expr514 --- term516
    minus515 --- expr516
    expr516 --- minus517
    expr516 --- term518
    minus517 --- expr518
    expr518 --- minus519
    expr518 --- term520
    minus519 --- expr520
    expr520 --- minus521
    expr520 --- term522
    minus521 --- expr522
    expr522 --- minus523
    expr522 --- term524
    minus523 --- expr524
    expr524 --- minus525
    expr524 --- term526
    minus525 --- expr526
    expr526 --- minus527
    expr526 --- term528
    minus527 --- expr528
    expr528 --- minus529
    expr528 --- term530
    minus529 --- expr530
    expr530 --- minus531
    expr530 --- term532
    minus531 --- expr532
    expr532 --- minus533
    expr532 --- term534
    minus533 --- expr534
    expr534 --- minus535
    expr534 --- term536
    minus535 --- expr536
    expr536 --- minus537
    expr536 --- term538
    minus537 --- expr538
    expr538 --- minus539
    expr538 --- term540
    minus539 --- expr540
    expr540 --- minus541
    expr540 --- term542
    minus541 --- expr542
    expr542 --- minus543
    expr542 --- term544
    minus543 --- expr544
    expr544 --- minus545
    expr544 --- term546
    minus545 --- expr546
    expr546 --- minus547
    expr546 --- term548
    minus547 --- expr548
    expr548 --- minus549
    expr548 --- term550
    minus549 --- expr550
    expr550 --- minus551
    expr550 --- term552
    minus551 --- expr552
    expr552 --- minus553
    expr552 --- term554
    minus553 --- expr554
    expr554 --- minus555
    expr554 --- term556
    minus555 --- expr556
    expr556 --- minus557
    expr556 --- term558
    minus557 --- expr558
    expr558 --- minus559
    expr558 --- term560
    minus559 --- expr560
    expr560 --- minus561
    expr560 --- term562
    minus561 --- expr562
    expr562 --- minus563
    expr562 --- term564
    minus563 --- expr564
    expr564 --- minus565
    expr564 --- term566
    minus565 --- expr566
    expr566 --- minus567
    expr566 --- term568
    minus567 --- expr568
    expr568 --- minus569
    expr568 --- term570
    minus569 --- expr570
    expr570 --- minus571
    expr570 --- term572
    minus571 --- expr572
    expr572 --- minus573
    expr572 --- term574
    minus573 --- expr574
    expr574 --- minus575
    expr574 --- term576
    minus575 --- expr576
    expr576 --- minus577
    expr576 --- term578
    minus577 --- expr578
    expr578 --- minus579
    expr578 --- term580
    minus579 --- expr580
    expr580 --- minus581
    expr580 --- term582
    minus581 --- expr582
    expr582 --- minus583
    expr582 --- term584
    minus583 --- expr584
    expr584 --- minus585
    expr584 --- term586
    minus585 --- expr586
    expr586 --- minus587
    expr586 --- term588
    minus587 --- expr588
    expr588 --- minus589
    expr588 --- term590
    minus589 --- expr590
    expr590 --- minus591
    expr590 --- term592
    minus591 --- expr592
    expr592 --- minus593
    expr592 --- term594
    minus593 --- expr594
    expr594 --- minus595
    expr594 --- term596
    minus595 --- expr596
    expr596 --- minus597
    expr596 --- term598
    minus597 --- expr598
    expr598 --- minus599
    expr598 --- term600
    minus599 --- expr600
    expr600 --- minus601
    expr600 --- term602
    minus601 --- expr602
    expr602 --- minus603
    expr602 --- term604
    minus603 --- expr604
    expr604 --- minus605
    expr604 --- term606
    minus605 --- expr606
    expr606 --- minus607
    expr606 --- term608
    minus607 --- expr608
    expr608 --- minus609
    expr608 --- term610
    minus609 --- expr610
    expr610 --- minus611
    expr610 --- term612
    minus611 --- expr612
    expr612 --- minus613
    expr612 --- term614
    minus613 --- expr614
    expr614 --- minus615
    expr614 --- term616
    minus615 --- expr616
    expr616 --- minus617
    expr616 --- term6
```

Eliminating Left-Recursion

A

$expr \rightarrow expr - term \mid term$
 $term \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

→

$expr \rightarrow term \ expr' \quad expr' \rightarrow - \ term \ expr' \mid \epsilon$
 $term \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

→ (EBNF)

$expr \rightarrow term \{ - term \} \quad \rightarrow \begin{aligned} void \ expr(void) \\ \{ \ term(); \\ \quad \text{while } (\text{token} == '-') \\ \quad \{ \ \text{match}('>'); \\ \quad \quad \text{term}(); \\ \} \end{aligned}$
 $term \rightarrow 0 \mid 1 \mid \dots \mid 9$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 25

Challenge 2: Backtracking is Inefficient

A

- **Backtracking: trial-and-error**

$type \rightarrow simple \mid array [simple] \ of \ type$ (Types in Pascal)
 $simple \rightarrow \text{char} \mid \text{integer}$

Input: $\text{array} \ [\ \text{integer} \] \ \text{of} \ \text{char}$

Parse tree

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 26

Challenge 2: Backtracking is Inefficient

A

$subscription \rightarrow term \mid term .. term$
 $term \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

- We cannot avoid backtracking if the grammar has multiple productions to apply, given a lookahead symbol.
- Solution:
 - Change the grammar so that there is only one applicable production that is unambiguously determined by lookahead symbol.

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 27

Avoiding Backtracking by Left Factoring

A

$A \rightarrow \alpha \beta_1 \mid \alpha \beta_2$
→
 $A \rightarrow \alpha A'$
 $A' \rightarrow \beta_1 \mid \beta_2$

$expr \rightarrow term \mid term .. term$
 $term \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

→

$expr \rightarrow term \ rest$
 $rest \rightarrow .. \ term \mid \epsilon$
 $term \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 28

Left Factoring Using EBNF

A

$expr \rightarrow term @ expr \mid term$
→
 $expr \rightarrow term [@ expr]$

$if\text{-statement} \rightarrow \text{if} (\ expr) \ statement$
 $\mid \text{if} (\ expr) \ statement \ \text{else} \ statement$

→
 $if\text{-statement} \rightarrow \text{if} (\ expr) \ statement \ [\ \text{else} \ statement]$

Lecture 4 – Syntax (Cont.), Spring 2008 CSE3302 Programming Languages, UT-Arlington ©Chengkai Li, Weimin He 2008 29