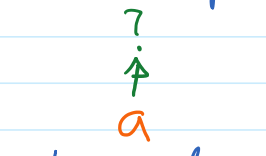
11 Grammar Issues For Top-Down Parsers

Monday, October 9, 2023 11:00 AM

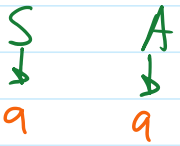
- AMBIGUITY

Not an issue for top-down parsers:

$S \to a \mid A$
$A \to a$



reduce-reduce conflict.



- LACK OF PAIRWISSE DISJOINTNESS

for a non-terminal A
more than one body of a rule with A in the head
begins with the same prefix.

E.G

$A \to aB \mid aC \mid aaD \mid b$

```
FUNCTION parse_A()
  IF token = 'a' THEN
    getToken()
    ???
```

the token does not gives you enough information on which rule body to apply.

Solution 1: use the next token.
use more look-ahead symbols.

Solution 2: Change Grammar.

REMEDY: LEFT FACTORING

Replace: $A \to \alpha\beta_1 \mid \alpha\beta_2 \mid \alpha\beta_3 \mid \ldots \mid \Gamma_1 \mid \Gamma_2 \mid \ldots$

α is a common prefix

with: $A \to \alpha A' \mid \Gamma_1 \mid \Gamma_2 \mid \ldots$

$A' \to \beta_1 \mid \beta_2 \mid \beta_3 \ldots$

e.g. $A \to \underbrace{aB}_{\beta} \mid \underbrace{aC}_{\beta} \mid \underbrace{aaD}_{\beta} \mid \underbrace{b}_{\Gamma}$

$\alpha = a$

Left factor

$A \to aA' \mid b$

$A' \to B \mid C \mid aD$

```
FUNCTION parse_A()
  IF token = 'a' THEN
    getToken()
    parse_Aprime()
  ELSIF token = 'b' THEN
    getToken()
```

Excersize:

⊗ $S \to zzxAx \mid zzyz \mid zzy \mid zzx \mid yyA \mid yyy$

⊗ $A \to xxAy \mid xzyA \mid yzx$

⊗ $S \to zzyS' \mid zzx \mid yyA \mid yyy$

$S' \to Ay \mid z \mid \lambda$

$A \to xA' \mid yzx$

$A' \to xAy \mid zyA$

⊗ $S \to zzS'' \mid yyA \mid yyy$

$S'' \to yS' \mid x$

$S \to zzS'' \mid yyS'''$

$S''' \to A \mid y$

- LEFT RECURSION

  - Direct
  - Indirect.

$A \to Aa \mid Ab \mid c$

$\{c, ca, caa, caba, \ldots\}$

```
FUNCTION parse_A()
  parse_A()
  ...
  ...
```

$A \to Bx$
$B \to Cy$
$C \to Az \mid w$

parse_A()
↳ parse_B()
↳ parse_C()
↳ parse_A()

REMEDY: DIRECT LEFT RECURSION ELIMINATION

Replace $A \to A\alpha_1 \mid A\alpha_2 \mid A\alpha_3 \mid \ldots \mid \beta_1 \mid \beta_2 \mid \beta_3 \mid \ldots$

(none begin with A)

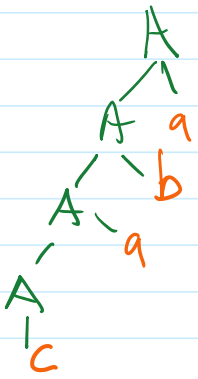
by:

$A \to \beta_1 A' \mid \beta_2 A' \mid \beta_3 A' \mid \ldots$
$A' \to \alpha_1 A' \mid \alpha_2 A' \mid \alpha_3 A' \mid \ldots \mid \lambda$

e.g.1

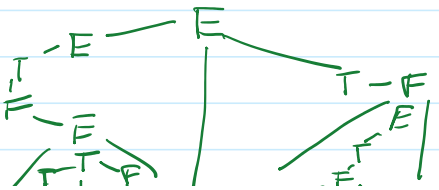
$A \to \underbrace{Aa}_{\alpha_1} \mid \underbrace{Ab}_{\alpha_2} \mid \underbrace{c}_{\beta_1}$ $\Rightarrow$

$A \to cA'$
$A' \to aA' \mid bA' \mid \lambda$



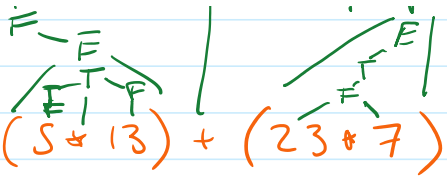
E.6.

$E \to \underbrace{E + T}_{\alpha} \mid \underbrace{T}_{\beta}$
$T \to T * F \mid F$
$F \to (E) \mid \text{int}$



T → T * F | F

F → (E) | int



E → TE'

- E' → + TE' | λ

T → FT'

- T' → * FT' | λ

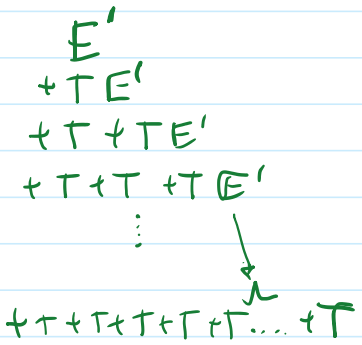
F → (E) | int

Extended BNF

E → T { + T }

T → F { * F }

F → (E) | int



—o— EOF