Chp. 5. Parsers

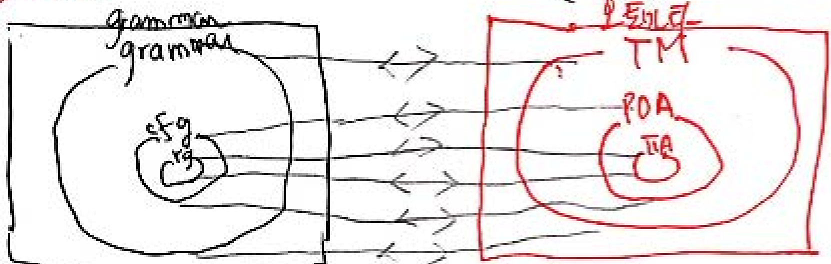
regular grammar (type 3) — 정규문법 $A \to x$ or $A \to xB \in P$
unique (rightmost) Nonterminal, if any! regular
$A, B \in N$, $x \in \Sigma^*$ FA (NFA, ε-NFA, DFA)

context-free grammar (type 2) — 문맥자유문법 $\forall A \to \alpha \in P$ FA + stack = pushdown Automata
many nonterminals grammar. context-free $A \in N$, $\alpha \in (N \cup \Sigma)^*$

~~context-sensitive grammar~~ (type 1) — 문맥의존문법 ~~오토마타~~ ~~FA + (bounded) tape~~
$\alpha A \beta \to \alpha \gamma \beta \in P$ $A \to \gamma$ context_sensitive

≡ (type 0) grammar $\gamma \to \beta \in P$ FA + tape(memory) = Turing machine
$\alpha \in (N \cup \Sigma)^+$, $\beta \in (N \cup \Sigma)^*$ 오토마타

문법
(unrestricted) grammar.



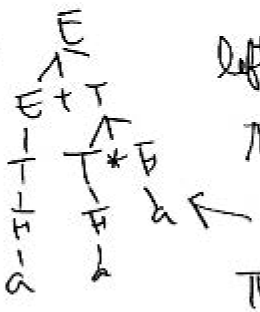
Chomsky's Language Hierarchy

parse tree $\xleftrightarrow{1:1}$ left parse -- a seq. of rules used in the ~~left~~most derivation
right parse 좌우선유도(lm)에서 사용된 rule의 string

$$S \overset{\pi_L}{\underset{lm}{\Rightarrow}} x \in \Sigma^*, \ \pi_L \in P^*$$

넌터미널중 우선으로 바꿀 놈의 정해지면 그 규칙(rule)의 parse tree는
(= parsed tree) 순서가 (leftmost or rightmost) unique하게 결정된다

Ex



$E \to E+T$ ①, $E \to T$ ②
$T \to T*F$ ③, $T \to F$ ④
$F \to a$ ⑤, $F \to (E)$

left parse for $a+a*a$

$$\pi_L = (E \to E+T) \cdot (E \to T) \cdot (T \to F)(F \to a)(T \to T*F) \cdot (T \to F)(F \to a)(F \to a)$$

$$= 1 \cdot 2 \cdot 4 \cdot 5 \cdot 3 \cdot 4 \cdot 5 \cdot 5$$

$$\pi_R = 1 \cdot 3 \cdot 5 \cdot 4 \cdot 5 \cdot 2 \cdot 4 \cdot 5$$

3/ ~~Prop. 5.5~~ guess & verify / produce & shift → left parse … Top-down Parsing

Left (Predictive) Parser for a cfg $G=(N,\Sigma,P,S)$.

$M_P^L = (N\cup\Sigma, \Sigma, \Gamma_P^L, S, \{\varepsilon\}, \$, 1)$ ⇒ see page 18

stack alph. (Voc) | input voc. | initial stack content | set of final stack contents | end marker delimeters.

Parse tree ↓ det. SLL(k), LL(k)

Γ_P^L (pa) $A| \to \omega^R|$ $A\to\omega\in P$. guess A as ω $(A\to\omega\in P)$ $(A\in N, \omega\in(N\cup\Sigma)^*)$

ω^R → nondet.

(sa) $a|a \to |$ $a\in\Sigma$, verify $a\in\Sigma$

→ det.

$G_{Exp}^{X unit}$

$$E\to E+T \mid T*F \mid a \mid (E)$$
$$T\to T*F \mid a \mid (E)$$
$$F\to a \mid (E)$$

unit production-free

vs G_{Exp}^{unit}

$$E\to E+T \mid T$$
$$T\to T*F \mid F$$
$$F\to a \mid (E)$$

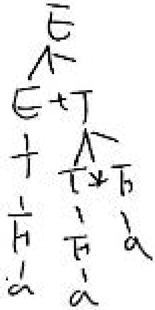
unit-production rule $(E\to T, T\to F)$

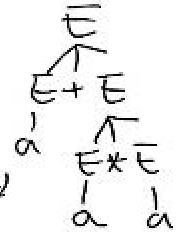
G_{UExp}

$$E\to E+E \mid E*E \mid a \mid (E)$$

precedence $+ < *$

associativity left $+, *$





SLR(k) LALR(k) LR(k)

let ⇑ yacc!

Right (shift-reduce) Parser

$M_P^R = (N\cup\Sigma, \Sigma, \Gamma_P^R, \varepsilon, \{S\}, \$, 1)$ $G_{Exp}^{X unit}$ ⇒ See Parse tree in Page 19

Γ_P^R (sa) $|a \to a|$ $a\in\Sigma$

(ra) $\omega| \to A|$ $A\to\omega\in P$.

right parse ≡ a seq. of rule used in rmd in reversed order.

$\$a+a*a|\$ \overset{a}{\underset{R}{\Rightarrow}} \$+a*a|a\$ \overset{E\to a}{\underset{R}{\Rightarrow}} \$+a*a|E\$ \overset{+}{\Rightarrow} \$a*a|+E\$ \overset{a}{\Rightarrow} \$*a|a+E\$$

$\overset{T\to a}{\Rightarrow} \$*a|T+E\$ \overset{*}{\Rightarrow} \$a|*T+E\$ \overset{a}{\Rightarrow} \$|a*T+E\$ \overset{F\to a}{\Rightarrow} \$|F*T+E\$$

$\overset{T\to T*F}{\Rightarrow} \$|T+E\$ \overset{E\to E+T}{\Rightarrow} \$|E\$$

final configuration

$\pi_R = (E\to a)\cdot(T\to a)\cdot(F\to a)\cdot(T\to T*F)\cdot(E\to E+T)$ 의 역순

※ right most derivation 순서 $(E\to E+T)\cdot(T\to T*F)\cdot(F\to a)\cdot(T\to a)\cdot(E\to a)$ (in reversed order) Bottom-Up Parsing