Lecture 23

Friday, March 15, 2024 11:28 AM

Converting a CFG into CNF

$S \to aSb \mid \varepsilon$ (getting rid of ε-prod?)

$S \to aSb \mid ab$

$S \to aAb \mid A$

$A \to S \mid ab$

Give CFG G:

1. Add "some productions" $\to \hat{G}$
2. Remove unit and ε-productions $\to \hat{\hat{G}}$
3. "Simplify" to CNF $\to G'$

---

step 1: $\hat{P}$ is initialized with P (productions), $\gamma \in (N \cup \Sigma)^*$, $A, B \in N$

Repeat the following two steps until no further updates are made:

* if $A \to B$ and $B \to \gamma$ are in $\hat{P}$, then add $A \to \gamma$ to $\hat{P}$.
* if $A \to \alpha B \gamma$ and $B \to \varepsilon$ are in $\hat{P}$, then add $A \to \alpha\gamma$ to $\hat{P}$.

Step 2: Remove all unit and ε-productions

Also, remove unreachable non-terminals.

$S \to aTb \mid \varepsilon \mid ab \mid$

$T \to S \mid ab \mid \varepsilon \mid aTb$

step 1

$\Downarrow$

step 2:

$S \to aTb \mid ab$

$T \to ab \mid aTb$

---

step 3: For each terminal a, create a new non terminal (say A) and update the rules.

step 3.1:

and update ...

step 3.1:

$S \to ATB \mid AB$

$T \to AB \mid ATB$

$A \to a$

$B \to b$

Then break the long sequences by introducing new non-terminals.

$S \to LB \mid AB$

$T \to AB \mid LB$

$L \to AT$

$A \to a$

$B \to b$

note: we could have merged S and T but not necessary

$S \to SS \mid [S] \mid \varepsilon \mid \underbrace{[\,] \mid S}_{\text{step 1}}$

step 2: $S \to SS \mid [S] \mid [\,]$

step3: $S \to SS \mid AR \mid LR$

$A \to LS$

$L \to [$

$R \to J$

Closure properties of CFLs

* Union: if $A = L(G_1)$ and $B = L(G_2)$ for CFGs G_1 and G_2, then $A \cup B$ is CFG? yes

$S \to S_1 \mid S_2 \to$ start nonterminal of G_2.

S_1: start symbol of G_1

And then $S_1 \to \cdots$
$S_2 \to \cdots$

* Concatanation:

$S \to S_1 S_2$ yes!

* Intersection:

the intersection of two CFLs is not necessarily CF.

$C = \{a^n b^n c^n : n \geq 0\} \to$ not CFG

$A = \{a^n b^n c^k, n \geq 0, k \geq 0\}$

$B = \{a^k b^n c^n : n \geq 0, k \geq 0\}$

$A \cap B = C$

CFL CFL not CFL