27/03/13

Eliminating ϵ-production

$L(G) - \epsilon$ as ϵ is not accepted by any normal form.

Any production leading to ϵ is called nullable.

Eg:

$S \to A$

$A \to aA \mid \epsilon$

$B \to bA$ here, A is nullable variable and as S derives A, ∴ S is also nullable.

Algorithm

Step 1: Initialize P_1 to P.

Step 2: Find all nullable variables in P.

Step 3: For every production $A \to \alpha$ in P, add to P_1 every production that can be obtained from it by deleting from α one or more occurrences of nullable variable.

Step 4: Delete all ϵ-productions from P_1 and delete all duplicate productions and productions of the form $A \to A$

Algo. to find nullable variables :

$N_0 = \{A \in V \mid P \text{ productions } A \to \epsilon\}$

$i = 0;$

do

$i = i + 1;$

$N_i = N_{i-1} \cup \{A \mid A \to \alpha, \alpha \in N_{i-1}^*\}$

while $N_i \neq N_{i-1}$

N_i is set of nullable variables.

Q. (1) $S \to AB$

$A \to aAA \mid \epsilon$

$B \to bBB \mid \epsilon$

here, $N_0 = \{A, B\}$

$N_1 = N_0 \cup S = \{A, B, S\}$

Step 1 ~

$$P_1 = \begin{cases} S \rightarrow \\ A \rightarrow \\ B \rightarrow \end{cases}$$

Step 2 ~ Nullable : A, B, S

Step 3 ~

$$P_1 = \begin{cases} S \rightarrow AB \mid B \mid A \mid \cancel{\epsilon} \\ A \rightarrow aAA \mid \cancel{\epsilon} \mid aA \mid \cancel{aA} \mid a \\ B \rightarrow bBB \mid \cancel{\epsilon} \mid bB \mid \cancel{bB} \mid b \end{cases}$$

Step 4 ~

$$P_1 = \begin{cases} S \rightarrow AB \mid B \mid A \\ A \rightarrow aAA \mid aA \mid a \\ B \rightarrow bBB \mid bB \mid b \end{cases}$$

(2) $S \rightarrow ABBAD$
$A \rightarrow aB \mid \epsilon$
$B \rightarrow bA \mid \epsilon$
$D \rightarrow a$

$N = \{A, B\}$

$S \rightarrow ABBAD \mid BBAD \mid ABBD \mid ABAD \mid \cancel{ABAD} \mid BAD \mid AAD \mid ABD \mid BBD \mid$
$AD \mid BD \mid D$

$A \rightarrow aB \mid \cancel{\epsilon} \mid a$

$B \rightarrow bA \mid \cancel{\epsilon} \mid b$

$D \rightarrow a$

(3) $S \rightarrow aAa \mid AB$
$A \rightarrow BS \mid aBa \mid \epsilon$
$B \rightarrow \epsilon \mid ab$

$N = \{A, B, S\}$

$S \rightarrow aAa \mid AB \mid aa \mid B \mid A \mid \cancel{\epsilon}$
$A \rightarrow BS \mid aBa \mid \cancel{\epsilon} \mid S \mid aa \mid B$
$B \rightarrow \cancel{\epsilon} \mid ab$

Chomsky normal form (CNF) –

A production is said to be in CNF if it is of the form :

$A \to BC$

$A \to a$

where, $A, B, C \in V$ and $a \in T$

V : variable
T : terminal

Why to convert in NF?

Easy to implement.

Q. Convert the following grammar in CNF :

(1)
$S \to AACD$
$A \to aAb \mid \epsilon \mid ab$
$C \to aC \mid a$
$D \to aDa \mid bDb \mid \epsilon$

$S \to S+T$

$S \to SS_1$
$S_1 \to t_1 S_1$
$t_1 \to t$
} CNF

(1) No useless productions.

(2) A and D are nullable.

$S \to AACD \mid ACD \mid AAC \mid CD \mid AC \mid C$ ↳ unit prod^n ∴ write aC|a instead.
$A \to aAb \mid \cancel{\epsilon} \mid ab$
$C \to aC \mid a$
$D \to aDa \mid aa \mid bDb \mid bb$

(3)
$S \to A_1A_2 \mid AA_2 \mid A_1C \mid CD \mid AC \mid A_3C \mid a$
$A_1 \to AA$
$A_2 \to CD$
$A_3 \to a$
$A_4 \to A_3A$
$A_5 \to b$

$A \to A_4A_5 \mid A_3A_5$
$C \to A_3C \mid a$
$D \to A_6A_3 \mid A_3A_3 \mid A_7A_5 \mid A_5A_5$

$A_6 \to A_3D$
$A_7 \to A_5D$

So there are 4 steps involved :

(1) Remove useless productions/symbols
(2) Eliminate ϵ productions
(3) Eliminate unit productions
(4) Convert to CNF

(2) $S \to aSb \mid ab$

(3) $S \to ab \mid aS \mid aaS$

(4) $S \to ABb \mid a$

$A \to aaA \mid B$

$B \to bAb$

(2) $S \to S_1S_2 \mid S_3S_4$

$S_1 \to S_3S$

$S_2 \to b$

$S_3 \to a$

(3) ~~$S \to ab \mid a$~~

$S \to S_1S_2 \mid S_3 \mid S_1S_3$

$S_1 \to a$

$S_2 \to b$

$S_3 \to aS$

(4) A and B are useless productions (do not reach a terminal)

$\therefore S \to a$

05/04/13

(1) $S \to ABaC$

$A \to BC \mid \epsilon$

$B \to b \mid \epsilon$

$C \to D \mid \epsilon$

$D \to d$

(2) $S \to a \mid aA \mid B \mid C$

$A \to aB \mid \epsilon$

$B \to Aa$

$C \to cCD$

$D \to ddd$

(3) $S \to AaB \mid aaB$

$A \to \epsilon$

$B \to bbA \mid \epsilon$

(4) $S \to aA \mid aBB$

$A \to aaA \mid \epsilon$

$B \to bB \mid bbC$

$C \to B$

(5) $S \to abAB$

$A \to bAB \mid \epsilon$

$B \to BAa \mid A \mid \epsilon$

(6) $S \to AaA \mid CA \mid BaB$

$A \to aaBa \mid CDA \mid aa \mid DC$

$B \to bB \mid bAB \mid bb \mid aS$

$C \to Ca \mid bC \mid D$

$D \to bD \mid \epsilon$

1) After removing useless productions,

$S \to ABaC$
$A \to BC$
$B \to b \mid \epsilon$
$C \to D \mid \epsilon$
$D \to d$

} no useless pd^n(s)

A, B, C are nullable variables

$S \to ABaC \mid AaC \mid ABa \mid Aa \mid a \mid BaC \mid \underline{AC} \mid Ba$
$A \to BC \mid B \mid C \mid \cancel{\epsilon}$
$B \to b \mid \cancel{\epsilon}$
$C \to D \mid \cancel{\epsilon}$
$D \to d$

After eliminating unit productions :

$S \to ABaC \mid AaC \mid BaC \mid ABa \mid \cancel{BaC} \mid aC \mid Ba \mid Aa \mid a$
$A \to BC \mid b \mid d$
$B \to b$
$C \to d$
$D \to d$

CNF :

$S \to A_1A_2 \mid AA_2 \mid BA_2 \mid A_1A_3 \mid A_3C \mid BA_3 \mid AA_3 \mid a$
$A_1 \to AB$
$A_2 \to A_3C$
$A_3 \to a$
$A \to BC \mid b \mid d$
$B \to b$

$C \to d$
$D \to d$

2) After removing useless productions,

$S \to a \mid aA \mid B$
$A \to aB \mid \epsilon$
$B \to Aa$

A is nullable.

$S \to a \mid aA \mid B \mid a$
$A \to aB \mid \cancel{\epsilon}$

$B \to Aa \mid a$

After eliminating unit pdⁿ :

$S \to a \mid aA \mid Aa$

$A \to aB$

$B \to Aa \mid a$

CNF :

$S \to a \mid A_1A \mid AA_1$

$A \to A_1B$

$B \to AA_1 \mid a$

$A_1 \to a$

(3)

No useless productions

A and B are nullable variables

$S \to AaB \mid aaB \mid aB \mid Aa \mid a \mid aa$

$A \to \cancel{\epsilon}$

$B \to bbA \mid \cancel{\epsilon} \mid bb$

Now,

$S \to aaB \mid aB \mid a \mid aa$

$B \to bb$

No unit productions.

CNF :

$S \to A_1A_2 \mid A_1B \mid a \mid A_1A_1$

$B \to B_1B_1$

$A_1 \to a$

$A_2 \to A_1B$

$B_1 \to b$

(4)

B and C are useless

$S \to aA$

$A \to aaA \mid \epsilon$

A is nullable

$S \to aA \mid a$

$A \to aaA \mid aa$ No unit productions.

$S \to A_1 A \mid a$

$A_1 \to a$

$A \to A_1 A_2 \mid A_1 A_1$

$A_2 \to A_1 A$

6~ No useless production

$D, A, C, S \to$ nullable variables

$S \to AaA \mid Aa \mid aA \mid a \mid CA \mid C \mid A \mid BaB \mid \cancel{\epsilon}$

$A \to aaBa \mid CDA \mid CD \mid DA \mid CA \mid C \mid D \mid A \mid aa \mid DC \mid \cancel{D} \mid \cancel{\epsilon}$

$B \to bB \mid bAB \mid \cancel{bB} \mid bb \mid aS \mid a$

$C \to Ca \mid a \mid bC \mid b \mid D$

$D \to bD \mid b$

Eliminating unit productions :

$S \to AaA \mid Aa \mid aA \mid a \mid CA \mid \underbrace{Ca \mid a \mid bC \mid b}_{C} \mid \underbrace{bD \mid b}_{D} \mid \underbrace{aaBa \mid CDA \mid CD \mid DA \mid CA \mid aa \mid DC}_{A} \mid BaB$

$A \to aaBa \mid CDA \mid CD \mid DA \mid CA \mid \underbrace{Ca \mid a \mid bC \mid b \mid bD}_{C} \mid aa \mid DC$

$B \to bB \mid bAB \mid bb \mid aS \mid a$

$C \to Ca \mid a \mid bC \mid b \mid bD$

$D \to bD \mid b$

Converting to CNF :

$A_1 \to AA_2$	$A_5 \to A_2A_2$
$A_2 \to a$	$A_6 \to BA_2$
$A_3 \to CA$	$A_7 \to CD$
$A_4 \to b$	$A_8 \to AB$

$S \to A_1A \mid AA_2 \mid A_2A \mid a \mid CA \mid CA_2 \mid A_4C \mid b \mid A_4D \mid A_3A_6 \mid A_7A \mid CD \mid DA \mid CA \mid A_2A_2 \mid DC \mid A_6B \mid BA_2 \mid A_2B$

$B \to A_4B \mid A_4A_8 \mid A_4A_4 \mid A_2S \mid \cancel{A}\ a$

$C \to CA_2 \mid a \mid A_4C \mid b \mid A_4D$

$D \to A_4D \mid b$

$A \to A_5A_6 \mid A_7A \mid CD \mid DA \mid CA \mid CA_2 \mid a \mid A_4C \mid b \mid A_4D \mid A_2A_2 \mid DC$