Нормална форма на Чомски

Деф За всяка граматика $G=(N,\Sigma,S,R)$ ще казваме, че е в НФЧ, ако $(\forall A \to_G \alpha \in R)[|\alpha|=2]$

Теорема За вс. КСГ G, съществува КСГ G' в НФЧ, такава че за вс. w с $|w|\geq 2$ е в сила, че $w\in L(G) \Longleftrightarrow w\in L(G')$

Алгоритъм: Нека $G=(N,\Sigma,S,R)$ е КСГ.

I. Заменяме правила от вида $A\to X_1\dots X_n$ с $A\to X_1A_1$, $A_1\to X_2A_2$, ..., $A_{n-2}\to X_{n-1}X_n$ за $A_1\dots A_{n-2}\notin N$.

II. Махане ε-правилата.

1. Образуване $\mathcal{E}=\{B \mid B\vdash^{*}\varepsilon\}$
2. Премахване вс. правила от вида $A\to\varepsilon$
3. За вс. $B\in\mathcal{E}$ и вс. $A\to BC$ или $A\to CB$ добавяме $A\to C$.

III. Премахване на правила с дясна страна с дължина 1.

1. За $\forall X \in N \cup \Sigma$ образуваме

$$D(X) = \{Y \in N \cup \Sigma \mid X \vdash^* Y\}$$

2. Премахване вс. преименуващи правила

3. Заместване вс. правило $A \to XY$ с $A \to X'Y'$ за $X' \in D(X)$ и $Y' \in D(Y)$

4. За начален символ S добавяме $S \to X'Y'$ за вс. $A \to X'Y'$, когато $A \in D(S)$

<u>Пример:</u> Имаме $\Gamma = (\{S, A, B\}, \{a, b\}, S, \{S \to A|B,$
$A \to \varepsilon \mid Aab,\ B \to S \mid aBb\})$

I. Премахване правилата, които са с дясна част > 2:

$S \to A|B$

$A \to \varepsilon \mid AA_1$

$A_1 \to ab$

$B \to S \mid aB_1$

$B_1 \to Bb$

II. Махане ε-правилата.

1. $E = \{A, S, B\}$

2. Премахване ε-правилата

$S \to A \mid B$

$A \to AA_1$

$A_1 \to ab$

$B \to S \mid aB_1$

$B_1 \to Bb$

3. За вс. $B \in E$ и вс. $A \to BC$ или $A \to CB$ добавяме $A \to C$

$S \to A \mid B$

$A \to AA_1 \mid A_1$

$A_1 \to ab$

$B \to S \mid aB_1$

$B_1 \to Bb \mid b$

III. Премахване правилата с дясна страна $= 1$

1. $D(a) = \{a\}$

$D(b) = \{b\}$

$D(S) = \{A, B, S, A_1\}$

$D(A) = \{A, A_1\}$

$D(A_1) = \{A_1\}$

$D(B) = \{S, A, B, A_1\}$

$D(B_1) = \{B_1, b\}$

2. Премахване "преименуващите" правила

$A \to AA_1$

$A_1 \to ab$

$B \to aB_1$

$B_1 \to Bb$

3. Заместване вс. $A \to XY$ с $A \to X'Y'$ за $X' \in D(X)$ и $Y' \in D(Y)$

$A \to AA_1 \mid A_1A_1$

$A_1 \to ab$

$B \to aB_1 \mid ab$

$B_1 \to Bb \mid Sb \mid Ab \mid A_1b$

4. Добавяне $S \to X'Y'$ за вс. $A \to X'Y'$ за $A \in D(S)$

$S \to AA_1 \mid A_1A_1 \mid aB_1 \mid ab$

$A \to AA_1 \mid A_1A_1$

$A_1 \to ab$

$B \to aB_1 \mid ab$

$B_1 \to Bb \mid Sb \mid Ab \mid A_1b$

CYK Алгоритъм

Нека $G = (N, \Sigma, S, R)$ е КСГ в НФЧ.

Нека $w \in \Sigma^*$ е дума с $|w| = n$, където $n \geq 2$ и нека $w = a_1 \ldots a_n$.

1. За вс. $i, j \in \{1, \ldots, n\}$, $i \leq j$ образуваме

$$N[i][j] = \{B \in N \cup \Sigma \mid B \vdash_{\Gamma}^{*} a_i \ldots a_j\},$$

където $N[i][i] = \{a_i\}$

2. От това, че G е в НФЧ имаме, че за вс. $i < j$

$$A \vdash_G^* a_i \ldots a_j \text{ т.с.т.к. } (\exists A \to_G XY)[X \vdash_G^* a_i \ldots a_k \;\&\; Y \vdash_G^* a_{k+1} \ldots a_j]$$

Следователно можем да изведем връзката

$$N[i][j] = \bigcup_{k=i}^{j} \{B \mid (\exists X \in N[i][k])(\exists Y \in N[k+1][j])[B \to_G XY]\}$$

3. Тогава имаме, че $w \in L(G)$ т.с.т.к. $S \in N[1][|w|]$

Пример:

$S \to AA_1 \mid A_1A_1 \mid aB_1 \mid ab$

$A \to AA_1 \mid A_1A_1$

$A_1 \to ab$

$B \to aB_1 \mid ab$

$B_1 \to Bb \mid Sb \mid Ab \mid A_1b$

Искаме да проверим дали $w = aabb \in L(G)$

$i \backslash j$	1	2	3	4
1	a	$\emptyset$	$\emptyset$	S,....
2		a	S, A_1, B	B_1
3			b	$\emptyset$
4				b

$S \in N[1][4]$
$\Rightarrow w \in L(G)$

$$N[1][3] = \{B \mid B \to XY,\ X \in N[1][1],\ Y \in N[2][3]\}$$
$$\cup \{B \mid B \to XY,\ X \in N[1][2],\ Y \in N[3][3]\}$$

$$N[2][4] = \{B \mid B \to XY,\ X \in N[2][2],\ Y \in N[3][4]\}$$
$$\cup \{B \mid B \to XY,\ X \in N[2][3],\ Y \in N[4][4]\}$$

$$N[1][4] = \{B \mid B \to XY,\ X \in N[1][1],\ Y \in N[2][4]\}$$
$$\cup \{B \mid B \to XY,\ X \in N[1][2],\ Y \in N[3][4]\}$$
$$\cup \{B \mid B \to XY,\ X \in N[1][3],\ Y \in N[4][4]\}$$

Забележка: На много места, за да се прилага CYK е възможна следната дефиниция на НФЧ.

def// Нека $\Gamma = (N, T, S, P)$ е КСГ. Казваме, че Γ е в НФЧ, ако вс. правило е от вида 1) $X \to a$
2) $X \to XY$

$\boxed{\text{CYK в този случай}}$

1. $N[i][i] = \{A \in N \mid X \to w_i \in P\}$ за $i = \overline{1, |w|}$

2. $N[i][j] = \{A \in N \mid (\exists k \in \{i, \dots, j-1\}) (\exists A \to XY \in P) [X \in N[i][k] \,\&\, Y \in N[k+1][j]]\}$

Идея: Когато пресмятаме $N[i][j]$

$\alpha_i \dots \alpha_k | \alpha_{k+1} \dots \alpha_j$, трябва да има

$X \models \alpha_i \dots \alpha_k \,\&\, Y \models \alpha_{k+1} \dots \alpha_j$, т.е.

$X \in N[i][k] \,\&\, Y \in N[k+1][j]$

Задача / Пример

Като използвате CYK, проверете дали baaba $\in L(\Gamma)$, когато

$\Gamma = (\{S, A, B, C\}, \{a, b\}, S, \{$

$S \to AB \mid BC,$

$A \to BA \mid a,$

$B \to CC \mid b,$

$C \to AB \mid a\})$

	b	a	a	b	a
b	$\{B\}$	$\{A, S\}$	$\emptyset$	$\emptyset$	$\{S, A, C\}$
a		$\{A, C\}$	$\{B\}$	$\{B\}$	$\{S, A, C\}$
a			$\{A, C\}$	$\{S, C\}$	$\{B\}$
b				$\{B\}$	$\{A, S\}$
a					$\{A, C\}$

$\Rightarrow$ baaba $\in L(G)$

Сложност в най-лошия случай