НОРМАЛЬНА ФОРМА ХОМСЬКОГО ДЛЯ КВ-ГРАМАТИК.

Граматика в нормальній формі Хомського (граматика в бінарній нормальній формі, квадратична граматика, grammar in Chomsky normal form) - контекстно-вільна граматика $\langle N, \Sigma, S, P\rangle$, в якій кожне правило є одним з таких трьох видів:

$$S \to \varepsilon, \quad S - \text{аксіома}$$
$$A \to a, \quad A \in N, a \in \Sigma$$
$$A \to BC, \quad A \in N, \quad B \in N - \{S\}, C \in N - \{S\}$$

Приклад

Граматика

$$S \to RR; \quad S \to AB; \quad R \to RR; \quad R \to AB;$$
$$A \to a; \quad B \to RB; \quad B \to b$$

є граматикою в нормальній формі Хомського.

Кожна контекстно-вільна граматика еквівалентна деякій граматиці в нормальній формі Хомського.

Перетворення КВ-граматики до нормальної форми Хомського.

Нехай дано контекстно-вільну граматику $G = \langle N, \Sigma, S, P \rangle$. Проведемо ряд перетворень цієї граматики так, щоб породжувана нею мова залишалася незмінною.

Перетворення КВ-граматики до нормальної форми Хомського.

(1) Якщо права частина деякого правила містить символ S, то замінимо граматику N, Σ, S, P на граматику

$$\langle N \cup \{S_0\}, \Sigma, S_0, P \cup \{S_0 \to S\}\rangle,$$

де S_0 – нова аксіома граматики і відповідно новий нетермінал, що не належить множині $N \cup \Sigma$. Так ми позбуваємось випадку, коли аксіома граматики зустрічається в правих частинах правил.

(2) Замінимо у всіх правилах кожен термінальний символ a на новий нетермінальний символ T_a і додамо до множини P правила $T_a \to a$ для всіх $a \in \Sigma$.

Перетворення КВ-граматики до нормальної форми Хомського.

(3) Видалимо правила вигляду $A \to \alpha$, де $|\alpha| > 2$, замінивши кожне з них на ряд коротких правил по два нетермінали кожному (при цьому додаються нові нетермінальні символи).

(4) Тепер видалимо всі правила вигляду $A \to \varepsilon$, де A не є початковим символом. Це можна зробити так.

(а) Якщо для якихось $A \in N$, $B \in N$, $\alpha, \beta \in (N \cup T)^*$, множина P містить правила $B \to \alpha A \beta$ і $A \to \varepsilon$, але не містить правила $B \to \alpha\beta$, то додамо це правило в P. Повторюємо цю процедуру для всіх ε-породжуючих нетерміналів.

(б) Тепер виключимо з множини P всі правила з ε-породжуючими нетерміналами в лівій частині правил граматики. Отримана граматика породжує мову $L - \{\varepsilon\}$.

(5) Якщо для будь-яких $A, B \in N$ і $\alpha \in (N \cup \Sigma)^*$ множина P містить правила $A \to B$ і $B \to \alpha$, але не містить правила $A \to \alpha$, то додаємо це правило в P. Правила $A \to B$ називаються **ланцюговими**. Повторюємо цю процедуру, доки можливо. Після цього виключимо з множини P всі правила вигляду $A \to B$.

Перетворення КВ-граматики до нормальної форми Хомського.

Зауваження. Перетворення КВ-граматик необхідно здійснювати САМЕ В ТАКІЙ ПОСЛІДОВНОСТІ:

а) видалити ε-продукції;

б) видалити ланцюгові продукції;

в) видалити некорисні символи;

г) ввести нову аксіому, якщо аксіома попередньої граматики вживалась в правих частинах правил;

д) замінити правила вигляду $A \rightarrow \alpha$, де $|\alpha| > 2$.