

Классификация грамматик по Хомскому.

### Тип 0.

Любая порождающая грамматика является грамматикой типа 0. На вид правил грамматик этого типа не накладывается никаких дополнительных ограничений.

### Тип 1.

Грамматикой типа 1 будем называть неукорачивающую грамматику.

Грамматика  $G = \langle T, N, P, S \rangle$  называется **неукорачивающей**, если правая часть каждого правила из  $P$  не короче левой части (т. е. для любого правила  $\alpha \rightarrow \beta \in P$  выполняется неравенство  $|\alpha| \leq |\beta|$ ). В виде исключения в неукорачивающей грамматике допускается наличие правила  $S \rightarrow \epsilon$ , при условии, что  $S$  (начальный символ) не встречается в правых частях правил.

#### Замечание:

Также грамматику типа 1 можно определить как контекстно-зависимую грамматику. (*Из определений видно, что контекстно-зависимая грамматика является неукорачивающей*).

Грамматика  $G = \langle T, N, P, S \rangle$  называется **контекстно-зависимой (КЗ)**, если каждое правило из  $P$  имеет вид  $\alpha \rightarrow \beta$ , где  $\alpha = \xi_1 A \xi_2$ ,  $\beta = \xi_1 \gamma \xi_2$ ,  $A \in N$ ,  $\gamma \in (T \cup N)^+$ ,  $\xi_1, \xi_2 \in (T \cup N)^*$ . В виде исключения в КЗ-грамматике допускается наличие правила с пустой правой частью  $S \rightarrow \epsilon$ , при условии, что  $S$  (начальный символ) не встречается в правых частях правил.

### Тип 2.

Грамматикой *типа 2* будем называть контекстно-свободную грамматику.

Грамматика  $G = \langle T, N, P, S \rangle$  называется **контекстно-свободной (КС)**, если каждое правило из  $P$  имеет вид  $A \rightarrow \beta$ , где  $A \in N$ ,  $\beta \in (T \cup N)^*$ .

#### Замечание:

Для любой КС-грамматики  $G$  существует эквивалентная ей неукорачивающая КС-грамматика, то есть такая КС-грамматика, в которой не более одного правила с пустой правой частью. Правило с пустой правой частью (если оно присутствует) имеет вид  $S \rightarrow \epsilon$ , где  $S$  - начальный символ, и при этом никакое правило из  $P$  не содержит  $S$  в своей правой части..

### Тип 3.

Тип 3 составляют регулярные грамматики. Регулярную грамматику можно определить либо как праволинейную, либо как леволинейную.

Грамматика  $G = \langle T, N, P, S \rangle$  называется **праволинейной**, если каждое правило из  $P$  имеет вид  $A \rightarrow wB$  либо  $A \rightarrow w$ , где  $A, B \in N$ ,  $w \in T^*$ .

Грамматика  $G = \langle T, N, P, S \rangle$  называется **леволинейной**, если каждое правило из  $P$  имеет вид  $A \rightarrow Bw$  либо  $A \rightarrow w$ , где  $A, B \in N$ ,  $w \in T^*$ .

#### Замечание:

**Автоматной** грамматикой называется праволинейная (леволинейная) грамматика для правил которой введено дополнительное ограничение: если правая часть правила непустая, то она имеет вид:  $A \rightarrow a$  либо  $A \rightarrow aB$  (для леволинейной, соответственно,  $A \rightarrow a$  либо  $A \rightarrow Ba$ ), где  $A, B \in N$ ,  $a \in T$ .

Автоматная грамматика является более простой формой регулярной грамматики. Существует алгоритм, позволяющий по регулярной (право- или леволинейной) грамматике построить соответствующую автоматную грамматику.

Для лексического анализа программ используется подмножество регулярных грамматик-автоматные грамматики без пустых правых частей.