

Грамматики

Цепочка

Цепочка над алфавитом - произвольная конечная последовательность символов из алфавита

Язык

Язык над алфавитом - произвольное множество цепочек, составленных из символов.

Язык Дика

К языку Дика принадлежат цепочки содержащие правильные скобочные последовательности.

Правила языка Дика: $S \rightarrow (S)$, $S \rightarrow SS$, $S \rightarrow e$ (e – пустая цепочка)

Порождающая грамматика

Порождающей грамматикой называется четверка: $G = (T, N, P, S)$, где:

- T - конечное множество *терминальных* символов - основной алфавит.
- N - конечное множество *нетерминальных* символов - вспомогательный алфавит.
Нетерминалы, по другому, - понятия грамматики языка, которые используются при его описании.
- P – конечное множество правил вывода, называемых также *продукциями*. Каждое правило вывода имеет вид $a \rightarrow b$, где a и b – цепочки терминальных и нетерминальных символов.
Правило $a \rightarrow b$ определяет возможность подстановки b вместо a в процессе порождения цепочек языка.
- S – начальный символ грамматики – один из нетерминальных символов, начальный(стартовый) нетерминал.

Сентенциальная форма

Сентенциальная форма грамматики называется цепочка, выводимая из начального нетерминала этой грамматики.

Сентенция – сентенциальная форма, состоящая только из терминальных символов

Разновидности алгоритмов разбора

- Нисходящий анализ.
Синтаксическое дерево строится от начального нетерминала в сторону цепочки.
- Восходящий анализ.
Дерево строится от цепочки в сторону корня дерева.

Можно в начале подбирать домино для левых символов терминальной цепочки, а можно в начале для правых – соответственно говорят о левосторонних и правосторонних алгоритмах разбора.

Иерархия грамматик Хомского

0. Произвольные грамматики. На вид их правил не накладывается каких-либо ограничений.
1. Контекстно-зависимые грамматики. Правила таких грамматик имеют вид $aAb \rightarrow ayb$, где a, b, y – цепочки терминалов и нетерминалов, A – нетерминальный символ. Такой вид означает, что нетерминал A может быть заменен цепочкой y только в контексте, образуемом цепочками a и b .
2. Контекстно-свободные грамматики. Их правила имеют вид $A \rightarrow y$, где A – нетерминал, y – цепочка терминалов и нетерминалов. В левой части правил всегда один нетерминальный символ.
3. Автоматные грамматики. Все правила имеют одну из трех форм: $A \rightarrow aB$, $A \rightarrow a$, $A \rightarrow e$, где A, B – нетерминалы, a – терминал, e – пустая цепочка. Автоматные грамматики также называют регулярными.