Язык и цель синтаксического анализа.

Машина при проверке текста, создает эталон этого текста, и сверяет его с введенным.

Основа понятия языка - L(N, T, P, S):

* N – мно-во нетерминальных символов
* T – мно-во терминальных символов
* P – мно-во правил
* S – начальный символ вывода

В этой структуре изначального символа с помощью правил можно построить все правильные цепочки. В дальнейшем под цепочкой будем понимать слово предложение или даже кусок текста. В записи цепочек участвуют исключительно члены мно-ва, а члены мно-ва N участвуют в записи правил языка. Правила в мно-ве P исключительно верны и достаточны.

Центральным аспектом теории является мно-во P. С точки зрения машинных алгоритмов мно-во P существует в двух вариантах:

* Порождающая грамматика
* Синтаксический граф

Между ними можно построить взаимно-однозначное соответствие.

Порождающая грамматика состоит из следующих элементов:

Каждое правило в ней представляет собой одну строку. У строки есть левая и правая части. В левой части всегда находятся символы из мно-ва N (заглавные латинские буквы). В правую часть могут входить элементы мно-ва N, T и два специальных символа - | и {}. Обединяет эти две части специальная операция отождествления l-> . Все они отделены “;”, а последнее правило заканчивается “.”

Первое правило S l-> A;

Грамматика такого вида называется Бэкус-Науровой формой (БНФ).

Операция отождествления заключается в следующем:

Левая часть правила должна заменятся на правую до тех пор, пока в ней не останется нетерминальных символов и специальных элементов.

Вертикальная черта означает альтернативу; при подставлении для отождествления можно взять любую из альтернатив.

Фигурные скобки обозначают цикл и записывается так, что при отождествлении тело цикла можно взять любое кол-во начиная от нуля.

Предположим у нас есть мно-во T = {‘(’, ‘)’; e} (e – пустой символ).

Пример:

P : S -> A; A -> (). => S -> A-> () – единственная правильная цепочка

P: S -> A; A -> ({A}){A}. – эти правила описывают все возможные правильные способы записать строку из скобок.

Синтаксический граф представляет собой картинку соответствующую БНФ. Элементы синтаксического графа:

1. ->( “(” ) -> - представление терминального символа
2. ->[ “A” ] -> - представление нетерминального символа
3. -> ( “(” ) -> [ “A” ] -> - отображение последовательности символов в правилах
4. -> -> < “T1” > -> -> - отображение альтернативы

-> < “T2” > ->

-> < “T3” > ->

5) -> -> -> -> -> -> -> ->

^ V

<- < “T1” > <-

Уточнение: никаких ромбов в графах не бывает, вместо них должны использоваться кружки, квадраты, стрелочки и тд, а ромб – это общий вид какого-то правила записанного в графе.

Уточнение: Синтаксически граф надо использовать так: поставить в начало графа указатель и двигаться по стрелке найдя кружок выписать его содержимое; найдя квадрат перепрыгнуть на тот символ, который он указывает; при окончательного выхода из графа будет выписана правильная цепочка в языке.

Уточнение: Существуют 6 правил, которые переводят синтаксический граф в текст программы.