ВСА Лекции

2.10 Язык и цель синтаксического анализа.

В основе синтаксического анализа лежит понятие языка, он состоит из 4х элементов: N, T, P, S.

N – множество нетерминальных символов

T – множество терминальных символов

P – множество правил

S – начальный символ вывода

В этой структуре из начального 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 множества T, а члены множества N участвуют в записи правил. Правила в множестве P исключительно верны и достаточны. Центральным аспектом теории является множество P. С точки зрения машинных алгоритмов множество P – существует в двух вариантах:

1. Порождающая грамматика
2. Синтаксический граф

Порождающая грамматика состоит из следующих элементов:

1. Каждое правило представляет собой строку
2. У строки есть левая и правая часть
3. В левой части – символы из множества N (в дальнейшем будем полагать что в множество N входят заглавные латинские буквы)
4. В правую часть могут входить элементы множества N, элементы множества T и два специальных символа (|, {})

Грамматика такого вида называется бекус-науровой формой.

Операция отождествления заключается в следующем – левая часть правила должна заменяться на правую до тех пор, пока в ней не останется ни терминальных символов, ни специальных элементов. Вертикальная черта означает альтернативу; При отождествлении можно взять любую часть альтернативы. При отождествлении тело цикла можно взять любое количество раз, начиная от нуля.