

Глава 4. Восходящий синтаксический анализ

4.6. *LR*-таблицы разбора

Возможны различные виды таблиц разбора, которые являются, по сути, различными формами представления магазинного автомата, принимающего данный контекстно-свободный язык. В качестве примера рассмотрим достаточно простой и понятный вид *LR*-таблицы разбора для *LR*-грамматик [18].

LR-анализатор использует два стека (стек символов и стек состояний) и специальную *LR*-таблицу разбора. При практической реализации стек символов не требуется, но мы будем его рассматривать для облегчения понимания принципов работы синтаксического анализатора.

LR-таблица разбора представляет собой прямоугольную матрицу, состоящую из столбцов для каждого символа грамматики (терминалов и нетерминалов), включая и маркер конца ввода, и строк, соответствующих каждому состоянию, в котором может находиться анализатор. Состояние обобщает информацию, содержащуюся в стеке ниже него. Комбинация состояния в вершине стека и текущего входного символа используется в качестве индекса таблицы разбора и по ее элементу определяется дальнейшее действие. Таблица разбора включает элементы четырех типов:

1. *Элементы переноса*. Обычно записываются в виде Si , что означает: поместить в стек символов входной символ, поместить в стек состояний состояние i и перейти в состояние i . Если входной символ является терминалом, принять его (перейти к следующему символу разбираемой входной строки).

2. *Элементы свертки*. Обычно записываются в виде Rj , что означает: выполнить свертку для продукции с номером j , т. е. допустив, что n есть число символов в правой части продукции с номером j , удалить n элементов из стека символов и n элементов из стека состояний и перейти к состоянию, находящемуся в вершине стека состояний. Нетерминал в левой части продукции с номером j считать входным символом на следующем шаге разбора.

3. *Элементы ошибок*. Это пустые ячейки в таблице разбора и соответствуют синтаксическим ошибкам.

4. *Элементы остановки (stop)*, которые указывают на успешное завершение разбора входной строки.

Следует заметить, что все *LR*-анализаторы (при $k \leq 1$) работают одинаково (независимо от того, построен анализатор для *LR*(1)-грамматики или какого-либо ее подкласса), разница между ними заключается только в построении таблиц разбора.