**Лабораторная работа №4**

**Восходящий синтаксический анализ**

***Цель работы*** – закрепление теоретических знаний по методам восходящего синтаксического анализа; приобретение умений и практических навыков построения LR-таблиц разбора, выбора и разработки структур данных и алгоритмов для реализации восходящего синтаксического анализа.

**Основные сведения**

При восходящем синтаксическом анализе дерево разбора для входной строки строится снизу вверх, начиная с листьев и завершая корнем дерева. Этот процесс можно рассматривать как свертку входной строки к начальному символу грамматики. Входная строка анализируется слева направо в поисках подстроки, которая может быть свернута, т. е. ищется правая часть продукции грамматики, совпадающая с подстрокой. Если такая продукция найдена, то соответствующая подстрока заменяется нетерминалом левой части продукции. В результате такой замены может быть получена сентенциальная форма грамматики, и процесс повторяется до тех пор, пока входная строка не преобразуется в начальный символ грамматики (корень дерева разбора). Действие замены подстроки нетерминалом левой части продукции называется *сверткой* или *приведением*, а свертываемая подстрока – *основой правосторонней сентенциальной формы*.

Процесс восходящего анализа реализуется с помощью стека. При этом анализатор выполняет две основные операции:

1) *сдвиг* (*перенос*), во время которого считывается символ входной строки и помещается в вершину стека (поэтому основа будет находиться в верхней части стека);

2) *приведение* (*свертка*) заключается в замене множества элементов в верхней части стека (основы) на один элемент – нетерминал из левой части соответствующей продукции.

При такой организации восходящего анализа проблема поиска основы и ее своевременного приведения сводится к проблеме разрешения конфликтов типа сдвиг/приведение (анализатор в конкретной ситуации должен знать, какое из этих действий необходимо выполнить) и приведение/приведение (если возможны разные приведения, необходимо знать, какое из них выполнять и выполнять ли их вообще). Очевидно, что необходимым условием для выполнения приведения является то, чтобы правая часть какой-либо продукции появилась в верхней части стека. Но это условие не является достаточным. Поэтому для разрешения конфликтов анализатор должен иметь дополнительную информацию. В зависимости от способов разрешения указанных конфликтов выделяют такие подклассы LR(1)-грамматик как LR(0), SLR(1), LALR(1) и наиболее общая LR(1)-грамматика.

Базовым понятием при реализации LR-разбора является понятие конфигурации. В общем случае LR(1)-конфигурация представляет собой пару. Первый элемент пары – это продукция грамматики, в правой части которой перед одним из символов стоит маркерная точка, он указывает, какая часть продукции уже исследована в данной точке синтаксическим анализатором. Второй элемент – терминальный символ, называемый *контекстом* конфигурации или *предпросмотром*. Контекст важен для LR(1)-конфигурации, соответствующей окончанию продукции, когда выполняется приведение. Это символ-следователь, который при разборе строки языка может оказаться предварительно просматриваемым символом в тех случаях, когда выполняется приведение в соответствии с порождающей продукцией.

Все возможные для заданной грамматики конфигурации по определенным правилам группируются в множества, соответствующие состояниям синтаксического анализатора. Конфигурации, которые неразличимы для анализатора, объединяются в одно состояние. Правила объединения конфигураций зависят от используемого подкласса LR(1)-грамматик. Чтобы найти все множество состояний анализатора, по грамматике строится детерминированный конечный автомат (*характеристический автомат*). В процессе его построения и производится объединение конфигураций в состояния. По характеристическому автомату затем тривиально строится соответствующая LR-таблица разбора.

Методы построения характеристических автоматов и соответствующих им таблиц разбора для различных подклассов LR(1)-грамматик подробно рассмотрены в [4].

**Подготовка к работе**

1. Изучить описание работы и другие теоретические сведения, касающиеся тематики выполняемой работы.

2. Построить характеристический автомат для заданной грамматики, определить к какому подклассу она принадлежит.

3. По характеристическому автомату построить LR-таблицу разбора.

4. Выбрать или разработать алгоритмы и структуры данных, используемых синтаксическим анализатором.

5. Разработать структуру синтаксического анализатора, определить его функции и алгоритмы и программно его реализовать.

**Выполнение работы**

1. Выполнить отладку всех программ, разработанных на этапе подготовки.

2. Исследовать работу синтаксического анализатора, внося в исходный текст транслируемой программы различные ошибки и рассматривая реакцию на них анализатора.

3. Продемонстрировать работу синтаксического анализатора преподавателю.

**Контрольные вопросы**

1. Основные задачи синтаксического анализа.
2. Построение дерева вывода при восходящем синтаксическом анализе.
3. Что такое основа?
4. Как строится характеристический автомат грамматики?
5. Что такое LR(1)-конфигурация?
6. Как выполняется операция замыкания?
7. Как определяется контекст конфигурации?
8. LR-таблица разбора. Алгоритм построения.
9. Алгоритм синтаксического анализа по LR-таблице разбора.
10. Достоинства и недостатки LR-разбора.