

### Лабораторная работа 3. Синтаксический анализатор.

В ходе синтаксического анализа исходный текст программы проверяется на соответствие синтаксическим нормам языка с построением дерева разбора (синтаксическое дерево), которое отражает синтаксическую структуру входной последовательности и удобно для дальнейшего использования, а также в случае несоответствия – позволяет вывести сообщения об ошибках.

Как правило, результатом синтаксического анализа является синтаксическое строение предложения, представленное либо в виде дерева зависимостей, либо в виде дерева составляющих, либо в виде некоторого сочетания первого и второго способов представления.

Таким образом на основе анализа выражений, состоящих из литералов, операторов и круглых скобок выполняется группирование токенов исходной программы в грамматические фразы, используемые для синтеза вывода.

Представление грамматических фраз исходной программы выполнить в виде дерева. Реализовать синтаксический анализатор с использованием одного из табличных методов (LL-, LR-метод, метод предшествования и пр.).

Разбор выражения  $COST = (PRICE + TAX) * 0.98$ .

Выходом анализатора служит дерево, которое представляет синтаксическую структуру, присущую исходной программе.

$$\langle \text{ИД1} \rangle = (\langle \text{ИД2} \rangle + \langle \text{ИД3} \rangle) * \langle \text{ИД4} \rangle.$$

По этой цепочке необходимо выполнить следующие действия:

- 1)  $\langle \text{ИД3} \rangle$  прибавить к  $\langle \text{ИД2} \rangle$ ;
- 2) результат (1) умножить на  $\langle \text{ИД4} \rangle$ ;
- 3) результат (2) поместить в ячейку, резервированную для  $\langle \text{ИД1} \rangle$ . Этой последовательности соответствует дерево, изображенное на рисунке 1.

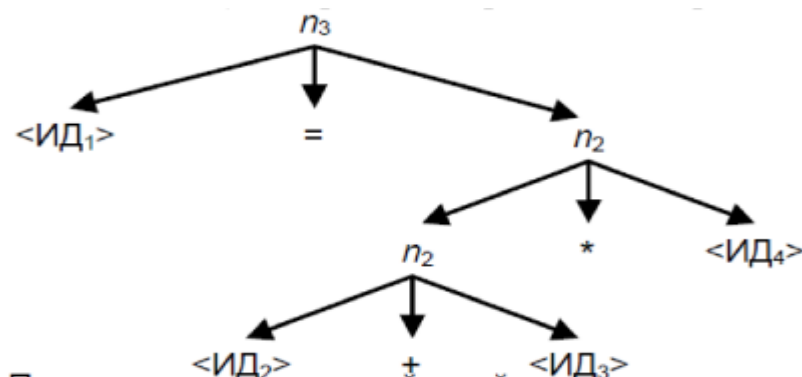


Рисунок 1. – Последовательность действий при вычислении выражения

Т.е. мы имеем последовательность шагов в виде помеченного дерева.

Внутренние вершины представляют те действия, которые можно выполнять. Прямые потомки каждой вершины либо представляют аргументы, к которым нужно применять действие (если соответствующая вершина помечена идентификатором или является внутренней), либо помогают определить, каким должно быть это действие, в частности, знаки «+», «\*» и «=». Скобки отсутствуют, т.к. они только определяют порядок действий.

LL- и LR- методы позволяют обнаружить ошибки на самых ранних стадиях, т.е. когда разбор потока токенов от лексического анализатора в соответствии с грамматикой языка становится невозможен.

Можно использовать нисходящий (англ. top-down parser) со стартового символа, до получения требуемой последовательности токенов. Для этих целей применим метод рекурсивного спуска либо LL-анализатор. Или использовать восходящий (англ. bottom-up parser) - продукции восстанавливаются из правых частей, начиная с токенов и кончая стартовым символом - LR-анализатор и проч.