

# УНИВЕРСАЛЬНЫЙ СИНТАКСИЧЕСКИЙ АНАЛИЗАТОР

А. И. Толкачёв

Научный руководитель - доцент М.С. Долинский  
(математический факультет)

Цель работы - получение универсального синтаксического анализатора (УСА), способного распознавать весь класс контекстно-свободных языков.

УСА состоит из компилятора ФБН и анализатора.

Для описания грамматики используются формы Бэкуса-Наура (ФБН) с некоторыми изменениями, облегчающими их восприятие:

- 1) отсутствие угловых скобок в описании нетерминалов;
- 2) двойные кавычки около терминальных строк;
- 3) разрешение комментариев;
- 4) наличие библиотеки макроподстановок и др.

Компилятор ФБН выполняет построение конечного детерминированного магазинного автомата, распознающего описанный язык.

При анализе текста может быть построено дерево, узлы которого - нетерминалы из описания языка.

В дальнейшем в описании языка будут поддерживаться атрибутные грамматики, что позволит использовать анализатор для построения компиляторов или для синтаксического анализа конструкций, которые не могут быть описаны с помощью ФБН (например, во многих языках переменные должны быть описаны перед использованием)

Для описания атрибутных грамматик планируется использование синтаксиса языка C++ и ФБН. При этом каждому нетерминалу из ФБН соот-

ветствует одна или несколько функций C++, которым передаются в качестве параметров синтезируемые и наследуемые атрибуты. Эти функции вызываются в порядке, соответствующем построенному дереву. При компиляции ФБН с описанием атрибутной грамматики получается текст программы на C++, в которую входят синтаксический анализатор, функции из описания атрибутной грамматики и функции, вызывающие их в нужном порядке. В эту программу можно внести изменения для работы с файлами, выдаче сообщений и т.д.

Для описанной грамматики может быть сгенерирована функция на языке C++, проводящая анализ строк. Весь код этой функции инкапсулирован в один класс, для возврата сообщений используется механизм обработки исключений. Эта возможность позволяет легко использовать синтаксический анализатор во многих программах.

Если текст не является цепочкой, принадлежащей грамматике, то выводится соответствующее сообщение, с указанием терминала или нетерминала, для которого построение поддерева невозможно, и последней позиции в тексте, принадлежащей дереву. В ближайшее время появится возможность задавать сообщения об ошибках при описании ФБН.

Анализатор реализован на языке C++ (версия Borland C++ 4.5). В настоящий момент имеется анализатор, работающий на основе метода рекурсивного спуска с медленным возвратом. Компилятор ФБН, выполняющий построение автомата, находится в стадии разработки.