

РЕАЛИЗАЦИЯ ВОСХОДЯЩЕГО СИНТАКСИЧЕСКОГО АНАЛИЗА С ПОМОЩЬЮ L-ГРАФОВ

Арсенин Никита Михайлович

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: veigh70@yandex.ru

Научный руководитель — Вылиток Алексей Александрович

Синтаксический анализ является одним из важнейших приложений теории формальных языков. Применительно к языкам программирования наибольший интерес представляет класс контекстно-свободных (бесконтекстных) грамматик. Существуют универсальные алгоритмы синтаксического анализа, основывающиеся на понятии бесконтекстной грамматики [1].

Для практических приложений используются подклассы бесконтекстных грамматик, к которым применимы более быстрые (по сравнению с универсальными) алгоритмы синтаксического анализа. Одним из таких классов является класс LR(1)-грамматик. Известно, что по любой грамматике из данного класса можно построить LR(1)-анализатор, позволяющий проводить восходящий синтаксический анализ за линейное время от длины входной строки [2]. Тем не менее внутреннее устройство такого анализатора оказывается достаточно сложным и требует оперирования большим числом дополнительных понятий и конструкций. Поэтому актуальной является задача построения аналога восходящего анализатора с более простой внутренней структурой.

Мы предлагаем подход, основанный на бесконтекстных L-графах с регулярным скобочным следом. Данные графы являются развитием понятия бесконтекстных L-графов, введенных в [3], и позволяют детерминизировать L-графы по аналогии с конечными автоматами.

Нами разработан алгоритм восходящего синтаксического анализа с помощью бесконтекстного L-графа с регулярным скобочным следом. Сначала строится L-граф по заданной LR(1)-грамматике с сохранением возможности восстановления синтаксических деревьев разбора. Построенный L-граф детерминизируется посредством операций конкатенации, объединения и звезды Клини над регулярными выражениями над скобочными следами. По детерминизированному L-графу строится алгоритм восходящего синтаксического анализа.

Все алгоритмические преобразования реализованы в виде прикладной программы на языке C# на платформе .NET. Программа

позволяет проводить полный цикл восходящего синтаксического анализа по входной LR(1)-грамматике, а также наглядно демонстрировать процесс детерминизации L-графа.

Литература

1. А. Ахо., Дж. Ульман. Теория синтаксического анализа, перевода и компиляции, том 1 М. : Мир, 1978 С. 352–369.
2. Серебряков В. А. Теория и реализация языков программирования. М. : Издательский отдел факультета ВМК МГУ имени М.В.Ломоносова, 2011 С. 352–369.
3. Вылиток А А, Сутырин П Г. Характеризация формальных языков графами // Сборник тезисов научной конференции "Тихоновские чтения", Москва, МГУ имени М.В.Ломоносова, факультет ВМК, 25-29 октября 2010 г. - 2010