

## ПРЕДИСЛОВИЕ

В современных информационных технологиях синтаксические методы играют существенную роль. Трансляторы языков программирования (компиляторы, интерпретаторы, конверторы), синтаксические редакторы, машинный перевод, различные средства обработки текстовой информации основаны на использовании синтаксических методов. Теория формальных языков и трансляций составляет теоретический фундамент этих методов. Основы этой теории были заложены Н. Хомским в 40–50-е годы XX столетия в связи с его лингвистическими работами, посвященными изучению естественных языков, но уже в следующем десятилетии синтаксические методы нашли широкое практическое применение в области разработки и реализации языков программирования. В настоящее время применение технологических средств, основанных на синтаксических методах, стало обыденным делом. Однако их грамотное и эффективное применение требует от пользователя знания, по крайней мере, основ математической теории, на которой они базируются.

Данное пособие написано на основе лекций, читавшихся автором в Ленинградском — Санкт-Петербургском государственном университете в разные годы для студентов математико-механического факультета. На отделении информатики этот курс читался совместно с С. Я. Фитиаловым — доцентом кафедры информатики, блестящим лектором, к сожалению, недавно ушедшим из жизни. Отчасти этим печальным обстоятельством вызвана подготовка данного пособия. Другая причина в том, что эта тематика, по-прежнему актуальная для специалистов и студентов, “вышла из моды”: фундаментальные монографии, изданные четверть века назад, стали библиографической редкостью. Автор надеется, что предлагаемое читателю пособие хотя бы в малой степени компенсирует существующую нехватку учебной литературы по изучаемому предмету.

Разбиение текста пособия на две части: I. Языки, грамматики, автоматы и II. Трансляции и синтаксические методы их реализации — соответствует былому распределению ролей двух лекторов, совместно читавших один курс. Часть I представляет собой авторизованный перевод отдельных глав монографии [2], с некоторыми дополнениями. Часть II является переложением нескольких тем из двухтомной монографии [6], отличающимся от оригинала, главным образом, тем, что в нем восстановлены полные доказательства многих утверждений, оставленных в первоисточнике в качестве упражнений. В Указателе литературы эти два источника следует считать основными.

Методические установки данного учебного пособия сводятся к тому, чтобы познакомить читателя с фундаментальными фактами классической теории языков и трансляций, дать ему представление о приемах доказательства утверждений в изучаемой области, а также снабдить его некоторым запасом практически полезных алгоритмов. Почти все доказательства носят конструктивный характер и потому дают полезный материал для практики.

Схематически содержание первой части следует в основном классификации языков по Н. Хомскому: в части I рассматриваются четыре типа грамматик и

соответствующих им распознавателей. Главу 6 о машинах Тьюринга в данном пособии можно рассматривать как факультативную, поскольку эта тема традиционно излагается в курсе по теории алгоритмов. Однако ее материал необходим для понимания последующих глав 7 и 8, в которых излагается связь языков типа 0 с машинами Тьюринга, и языков типа 1 — с линейно ограниченными автоматами. Несколько особняком стоит глава 9, в которой рассматриваются свойства замкнутости языков относительно элементарных операций и отображений. Она дополняет главу 3, в которой рассматриваются свойства замкнутости регулярных множеств.

Часть II посвящена описанию двух классов контекстно-свободных языков:  $LL(k)$  и  $LR(k)$  ввиду их практической важности. Первый является наибольшим естественным классом языков, допускающих детерминированный нисходящий разбор посредством  $k$ -предсказывающего алгоритма анализа. Второй представляет столь же естественный класс языков, допускающих детерминированный восходящий разбор посредством канонического анализатора Д. Кнута. Тот и другой имеют линейную сложность по времени. Как известно, метод Кнута используется в инструменте YACC — популярном строителе синтаксических анализаторов, а также в более развитых модификациях этого технологического средства. Фактически в них используются  $LR(1)$ -грамматики с дополнительными оптимизирующими ограничениями — так называемые  $LALR(1)$ -грамматики. Последние обсуждаются в конце этой части. В части II рассматриваются также некоторые классы синтаксически-управляемых трансляций с входными языками упомянутых классов и соответствующие алгоритмы их реализации. Последние описываются с достаточной для практики полнотой и обоснованием корректности и оценок сложности.

В Приложении приводится краткая сводка алгоритмически разрешимых и неразрешимых проблем, касающихся формальных языков. Она позволит практикующему читателю сориентироваться, что возможно или невозможно в мире формальных языков.

В пособии имеется достаточное число примеров, но нет задач и упражнений. Автор надеется исправить этот недостаток, подготавливая к изданию «Сборник задач и упражнений по теории формальных языков и трансляций».

Автор признателен рецензентам: профессору И. Л. Братчикову из Санкт-Петербургского государственного университета и доценту Ю. Л. Костюку из Томского государственного университета — за критические замечания, которые по возможности были приняты во внимание.

Особую благодарность хочется выразить заведующему лабораторией Java-технологии Научно-исследовательского института математики и механики СПбГУ им. академика В. И. Смирнова профессору В. О. Сафонову, предоставившему технические средства для подготовки оригинал-макета данного пособия.

*Б. К. Мартыненко*  
Старый Петергоф,  
Июнь 2002 г.