

ВВЕДЕНИЕ

Эта книга задумана как обзор наиболее важных компьютерных алгоритмов, которые применяются в настоящее время, а также как учебник фундаментальных технологий для постоянно растущего количества разработчиков, испытывающих потребность в такой информации. Ее можно использовать в качестве учебника для студентов второго, третьего и четвертого курсов, изучающих информатику, после овладения основными навыками программирования и знакомства с компьютерными системами, но перед спецкурсами по самым современным областям информатики или компьютерных приложений. Книга содержит реализации полезных алгоритмов и подробную информацию по характеристикам производительности этих алгоритмов, поэтому она может пригодиться и тем, кто занимается самообразованием, или послужить справочником для интересующихся разработкой компьютерных систем или приложений. Широкий круг рассматриваемых вопросов делает ее подходящим введением в данную область.

В новом издании текст был полностью переработан, и в него было включено более тысячи новых упражнений, более ста новых рисунков и десятки новых программ. Кроме того, ко всем рисункам и программам были добавлены подробные комментарии. Этот новый материал охватывает новые темы, а также содержит более полные пояснения ко многим классическим алгоритмам. Большой акцент на абстрактные типы данных расширяет сферу применения приведенных программ и делает их более пригодными для современных сред программирования. Читатели, знакомые с предыдущими изданиями книги, найдут множество новой информации, и все читатели найдут в книге большой объем учебного материала, который эффективно объясняет важные концепции.

Книга разбита на пять частей. В частях I–IV освещены фундаментальные концепции, структуры данных, алгоритмы сортировки и поиска; часть V посвящена более сложным алгоритмам и приложениям, построенным на базе абстракций и методов, которые разработаны в частях I–IV. Почти весь включенный в это издание материал по основным принципам и структурам данных написан заново. Следующий том (еще не готовый, части VI–VIII) охватывает строки (часть VI), вычислительную геометрию (часть VII) и современные алгоритмы и приложения (часть VIII).

Эта книга предназначена не только для программистов и студентов, изучающих информатику. Подавляющее большинство пользователей компьютеров желает работать быстрее либо решать более сложные задачи. Приведенные в книге алгоритмы представляют собой квинтэссенцию знаний, накопленных за последние более чем 50 лет, которые необходимы для эффективного использования компьютера в широком множестве приложений. Описанные здесь базовые методы стали важной составной частью научных исследований — начиная с задач моделирования систем из N тел в физике и завершая задачами анализа генетического кода в молекулярной биологии. Они являются также важными составными частями современных программных систем — от систем управления базами данных до механизмов поиска в интернете. По мере расширения сферы применения компьютерных приложений возрастает и значение описанных здесь базовых методов. Эта книга задумана как источник информации для студентов и профессионалов, которым важно понимать и эффективно использовать описанные фун-

даментальные алгоритмы в качестве базовых инструментов для любого компьютерного приложения, которым они займутся.

КРУГ РАССМАТРИВАЕМЫХ ВОПРОСОВ

Данная книга содержит 22 главы, сгруппированных в пять больших частей: анализ, структуры данных, сортировка, поиск и алгоритмы на графах. Приведенные описания призваны познакомить читателей с основными свойствами обширного круга основных алгоритмов. Описанные здесь алгоритмы находят широкое применение в течение многих лет и весьма важны как для профессиональных программистов, так и для изучающих информатику. Все описанные в книге хитроумные методы, от биномиальных очередей до патricia-деревьев, относятся к базовым концепциям, лежащим в основе компьютерных наук. Основная цель написания книги — собрать воедино фундаментальные методы из этих различных областей для ознакомления с лучшими методами решения задач с помощью компьютера.

Собранный в этой книге материал наиболее полезен для тех читателей, которые знакомы с одним-двумя курсами по информатике или обладают аналогичным опытом в программировании: курс по программированию на языках высокого уровня наподобие C++, Java или C и, возможно, курс по основным концепциям систем программирования. В тексте также имеются ссылки на работы, которые содержат дополнительный материал.

Большинство математических выкладок, сопровождающих аналитические результаты, приведены полностью (либо помечены как выходящие за рамки этой книги), поэтому для понимания материала требуется лишь минимальная математическая подготовка, хотя хорошая подготовка, конечно, не повредит. Основная цель здесь — связать важные алгоритмы с другими методами, а не изучать весь математический материал. Однако строгий подход, естественный для настоящих математиков, часто приводит к созданию хороших программ, поэтому я старался выдержать баланс между формальным подходом, на котором настаивают теоретики, и изложением материала, необходимого практикам, не жертвуя при этом строгостью.

ИСПОЛЬЗОВАНИЕ В УЧЕБНЫХ КУРСАХ

Изучение изложенного в книге материала может существенно варьироваться в зависимости от предпочтений преподавателя и подготовки студентов. Книга содержит достаточный объем базового материала для изучения структур данных начинающими, и в ней приведено достаточно дополнительного материала, чтобы использовать ее в качестве учебника по разработке и анализу алгоритмов для более подготовленных студентов. Одни преподаватели могут уделять больше внимания реализациям и практическим вопросам, а другие — анализу и теоретическим концепциям.

Автор разрабатывает различные учебные материалы для использования в сочетании с этой книгой: слайдовые презентации для использования на лекциях, задания по программированию, домашние задания и примеры экзаменационных билетов, а также интерактивные упражнения для студентов. Эти материалы будут доступны на веб-странице книги по адресу <http://www.awl.com/cseng/titles/0-201-35088-2>.

В элементарном курсе по структурам данных и алгоритмам основное внимание может уделяться базовым структурам данных, описанным в части II, и их использованию в реализациях, приведенных частях III и IV. В курсе по разработке и анализу алгоритмов основное внимание можно уделить фундаментальному материалу, изложенному в

части I и главе 5, а после этого рассмотреть способы достижения хорошей асимптотической производительности алгоритмов из частей III и IV. В курсе по разработке программного обеспечения можно опустить математические выкладки и дополнительные материалы по алгоритмам и сконцентрироваться на интеграции приведенных здесь реализаций в большие программы или системы. В курсе по алгоритмам может использоваться обзорный подход с кратким введением в концепции из всех этих областей.

В течение нескольких последних лет предыдущие издания книги использовались во множестве колледжей и университетов по всему миру в качестве учебника для студентов второго или третьего курсов компьютерных специальностей, а также дополнительного учебного пособия для других курсов. Опыт Принстонского университета позволяет утверждать, что охват материала в этой книге предоставляет вводные сведения по информатике, которые затем можно расширять в последующих курсах по анализу алгоритмов, программированию систем и теории компьютеров; все возрастающую группу студентов других специальностей этот материал снабжает большим набором технологий, которые они смогут немедленно и успешно использовать.

Упражнения, большинство из которых впервые включены в это издание, разделяются на несколько типов. Некоторые предназначены для проверки усвоения изложенного материала, и в них читателям предлагается просто проработать пример или применить описанные в тексте концепции. Другие же требуют реализации и сборки алгоритмов или экспериментального сравнения вариантов алгоритмов и изучения их свойств. Есть и упражнения, содержащие важную информацию, уровень детализации которой выходит за рамки основного материала книги. Обдумывание и выполнение упражнений будет полезно абсолютно всем читателям.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ АЛГОРИТМОВ

Каждый желающий как можно более эффективно использовать компьютер может воспользоваться этой книгой как справочным пособием или пособием для самостоятельного обучения. Читатели, имеющие опыт программирования, могут найти в книге информацию по конкретным темам. В большинстве случаев отдельные главы книги не зависят от других, хотя иногда в алгоритмах из одной главы используются методы, описанные в предыдущих главах.

Книга ориентирована на изучение алгоритмов, которые чаще всего применяются на практике. В ней приводится информация о средствах, с помощью которых читатели могут уверенно реализовывать, отлаживать и настраивать алгоритмы для решения конкретных задач или в качестве составной части некоторого приложения. В книгу включены полные реализации рассматриваемых методов, а также описания работы этих программ на взаимодополняющих примерах.

Поскольку алгоритмы записываются на конкретном языке, а не псевдокодом, программы можно быстро задействовать в практических целях. Их листинги доступны на веб-странице книги. При изучении алгоритмов эти работающие программы можно использовать различными способами. Их можно прочесть для проверки своего понимания особенностей алгоритма или для ознакомления с одним из возможных способов выполнения инициализации, задания граничных условий и других неочевидных ситуаций, которые зачастую вызывают затруднения при программировании. Читатели могут запустить их, чтобы увидеть алгоритмы в действии, эмпирически оценить производительность и сравнить результаты с данными, приведенными в таблицах, либо чтобы опробовать собственные модификации.

В книге подробно рассматриваются рабочие характеристики алгоритмов и ситуации, в которых эти алгоритмы могут оказаться полезными. Описываются интересные случаи взаимосвязи между рассматриваемыми практическими алгоритмами и чисто теоретическими результатами. Хотя это и не подчеркивается, в тексте книги устанавливается взаимосвязь между анализом алгоритмов и теорией информатики. В книге собрана и проанализирована конкретная информация по характеристикам производительности различных алгоритмов и реализаций.

ЯЗЫК ПРОГРАММИРОВАНИЯ

Для всех реализаций используется язык программирования C++. В программах применяется множество стандартных идиом C++, а в тексте приводятся краткие описания каждой из конструкций.

Вместе с Крисом Ван Виком мы разработали стиль программирования на C++, основанный на классах, шаблонах и перегруженных операциях, который, как мы считаем, позволяет эффективно представлять алгоритмы и структуры данных в виде реальных программ. Мы старались получать элегантные, компактные, эффективные и переносимые реализации. Везде, где это возможно, мы стремились сохранить единство стиля, чтобы аналогичные программы были похожи друг на друга.

Алгоритмы представлены в книге в максимально простой и непосредственной форме. Для множества приведенных в этой книге алгоритмов схожесть не зависит от языка: быстрая сортировка остается быстрой сортировкой (это лишь один яркий пример) независимо от того, выражена ли она на языке Ada, Algol-60, Basic, C, C++, Fortran, Java, Mesa, Modula-3, Pascal, PostScript, Smalltalk или на одном из других бесчисленных языков программирования, или в другой среде, где она зарекомендовала себя как эффективный метод сортировки. С одной стороны, представленный нами код основан на опыте реализации алгоритмов на этих и множестве других языков (доступна также C-версия этой книги, а Java-версия будет вскоре издана). С другой стороны, отдельные особенности некоторых из этих языков продиктованы опытом работы разработчиков с некоторыми алгоритмами и структурами данных, которые рассматриваются в книге.

Глава 1 является характерным примером такого подхода к разработке эффективных реализаций рассматриваемых алгоритмов на C++, а в главе 2 описывается используемый нами способ их анализа. Главы 3 и 4 посвящены описанию и обоснованию основных механизмов, используемых для реализаций типов данных и абстрактных типов данных. Эти четыре главы закладывают фундамент для понимания остальной части книги.

ОБ УПРАЖНЕНИЯХ

Классификация упражнений — занятие трудное, поскольку читатели подобных книг обладают различным уровнем знаний и опыта. Но все-таки небольшие указания не мешают, поэтому многие упражнения помечены одним из четырех маркеров.

Упражнения, предназначенные для *проверки понимания* материала, помечены белым треугольничком:

- ▷ **9.56.** Приведите биномиальную очередь, которая получится, если вставить ключи E A S Y Q U E S T I O N в первоначально пустую биномиальную очередь.

Чаще всего такие упражнения непосредственно связаны с примерами в тексте. Они не должны вызывать особых трудностей, но их выполнение может прояснить факт или концепцию, которые, возможно, были упущены при прочтении текста.

Упражнения, которые *дополняют текст новой и требующей размышлений* информацией, помечены белым кружком:

- **14.19.** Напишите программу, которая вставляет N случайных целых чисел в таблицу размером $N/100$ с цепочками переполнения, а затем определяет длину самого короткого и самого длинного списков, при $N = 10^3, 10^4, 10^5$ и 10^6 .

Такие упражнения заставляют подумать о важных концепциях, связанных с изложенным в тексте материалом, или ответить на вопрос, который мог возникнуть при чтении. Пожалуй, эти упражнения полезно прочесть даже при отсутствии времени на их выполнение.

Упражнения, которые ставят *задачу*, помечены черным кружком:

- **8.45.** Допустим, сортировка слиянием реализована таким образом, что разбиение файла выполняется в *произвольном* месте, а не точно в середине файла. Сколько в среднем сравнений выполнит этот метод для упорядочения N элементов?

Для выполнения таких упражнений может потребоваться значительное время, в зависимости от опыта читателя. Обычно их удобнее выполнять в несколько подходов.

Некоторые *особо трудные* (по сравнению с большинством других) упражнения помечены двумя черными кружками:

- **15.29.** Докажите, что высота *trie*-дерева, построенного из N случайных битовых строк, приблизительно равна $2 \lg N$.

Эти упражнения похожи на вопросы, рассматриваемые в научной литературе, однако некоторым читателям материала книги может хватить, чтобы попытаться ответить на них (возможно, успешно).

Мы старались, чтобы маркеры имели как можно меньше отношения к программной и математической подготовке читателей. Упражнения, которые требуют наличия опыта в программировании или математическом анализе, очевидны. Мы предлагаем всем читателям проверить свое понимание алгоритмов, реализовав их. И все же упражнения наподобие приведенного ниже не составят никакого труда для профессиональных программистов или студентов, изучающих программирование, но могут потребовать значительных усилий от тех, у кого такой практики нет:

- **1.22.** Измените программу 1.4, чтобы она генерировала случайные пары целых чисел в диапазоне от 0 до $N-1$ вместо того, чтобы считывать их из стандартного ввода, и выполняла цикл до тех пор, пока не будет выполнено $N-1$ операций *объединение*. Выполните программу для значений $N = 10^3, 10^4, 10^5$ и 10^6 и выведите общее количество ребер, генерируемых для каждого значения N .

Мы предлагаем читателям разбирать приводимые аналитические обоснования свойств изучаемых алгоритмов. Хотя упражнения вроде приведенного ниже не составят сложности для ученого или изучающего дискретную математику, они могут потребовать значительных усилий от тех, кто давно не занимался математическим анализом:

- 1.12.** Вычислите *среднее* расстояние от узла до корня в худшем случае в дереве из 2^n узлов, построенном алгоритмом взвешенного быстрого объединения.

Книга содержит слишком много упражнений, чтобы прочесть и усвоить все; но я надеюсь, что количество упражнений достаточно, чтобы позволить читателям глубже разобраться в интересующих их темах, чем при простом чтении текста.