Simon
Plouffe
1985

THE ART OF COMPUTER PROGRAMMING

SECOND EDITION

DONALD E. KNUTH *Stanford University*

ADDISON-WESLEY PUBLISHING COMPANY

Volume 2 / **Seminumerical Algorithms**

THE ART OF COMPUTER PROGRAMMING

SECOND EDITION

Reading, Massachusetts
Menlo Park, California · London · Amsterdam · Don Mills, Ontario · Sydney

This book is in the

ADDISON-WESLEY SERIES IN COMPUTER SCIENCE AND INFORMATION PROCESSING

MICHAEL A. HARRISON, Consulting Editor

Library of Congress Cataloging in Publication Data

Knuth, Donald Ervin (1938–
The Art of Computer Programming. 2d ed.

(Addison–Wesley series in computer science and information processing)
Includes index.
Contents: v. 1. Fundamental algorithms.—v. 2. Seminumerical algorithms.
1. Electronic digital computers—Programming. I. Title.
QA76.6.K64 001.6'42 73-1830
ISBN 0-201-03822-6 (v. 2)

The quotation on page 60 is reprinted by permission of Grove Press, Inc.

ISBN 0-201-03822-6
BCDEFGHIJ-MA-898765432

PREFACE

O dear Ophelia!
I am ill at these numbers:
I have not art to reckon my groans.

—Hamlet (Act II, Sc. 2, Line 120)

THE ALGORITHMS discussed in this book deal directly with numbers; yet I believe they are properly called *seminumerical*, because they lie on the borderline between numeric and symbolic calculation. Each algorithm not only computes the desired answers to a problem, it also is intended to blend well with the internal operations of a digital computer. In many cases a person will not be able to appreciate the beauty of such an algorithm unless he or she also has some knowledge of a computer's machine language; the efficiency of the corresponding machine program is a vital factor that cannot be divorced from the algorithm itself. The problem is to find the best ways to make computers deal with numbers, and this involves tactical as well as numerical considerations. Therefore the subject matter of this book is unmistakably a part of computer science, as well as of numerical mathematics.

Some people working in "higher levels" of numerical analysis will regard the topics treated here as the domain of system programmers. Other people working in "higher levels" of system programming will regard the topics treated here as the domain of numerical analysts. But I hope that there are a few people left who will want to look carefully at these basic methods; although the methods reside perhaps on a low level, they underlie all of the more grandiose applications of computers to numerical problems, so it is important to know them well. We are concerned here with the interface between numerical mathematics and computer programming, and it is the mating of both types of skills that makes the subject so interesting.

There is a noticeably higher percentage of mathematical material in this book than in other volumes of this series, because of the nature of the subjects treated. In most cases the necessary mathematical topics are developed here starting almost from scratch (or from results proved in Volume 1), but in some easily recognizable sections a knowledge of calculus has been assumed.

This volume comprises Chapters 3 and 4 of the complete series. Chapter 3 is concerned with "random numbers": it is not only a study of various methods for generating random sequences, it also investigates statistical tests for randomness, as well as the transformation of uniform random numbers into other types of random quantities; the latter subject illustrates how random numbers are used in practice. I have also included a section about the nature of randomness itself. Chapter 4 is my attempt to tell the fascinating story of what mankind has been able to learn about the processes of arithmetic, after centuries of progress. It discusses various systems for representing numbers, and how to convert between them; and it treats arithmetic on floating point numbers, high-precision integers, rational fractions, polynomials, and power series, including the questions of factoring and finding greatest common divisors.

Each of Chapters 3 and 4 can be used as the basis of a one-semester college course at the junior to graduate level. Although courses on "Random Numbers" and on "Arithmetic" are not presently a part of many college curricula, I believe the reader will find that the subject matter of these chapters lends itself nicely to a unified treatment of material that has real educational value. My own experience has been that these courses are a good means of introducing elementary probability theory and number theory to college students; nearly all of the topics usually treated in such introductory courses arise naturally in connection with applications, and the presence of these applications can be an important motivation that helps the student to learn and to appreciate the theory. Furthermore, each chapter gives a few hints of more advanced topics that will whet the appetite of many students for further mathematical study.

For the most part this book is self-contained, except for occasional discussions relating to the MIX computer explained in Volume 1. Appendix B contains a summary of the mathematical notations used, some of which are a little different from those found in traditional mathematics books.

In addition to the acknowledgments made in the preface to Volume 1, I would like to express deep appreciation to Elwyn R. Berlekamp, John Brillhart, George E. Collins, Stephen A. Cook, D. H. Lehmer, M. Donald MacLaren, Mervin E. Muller, Kenneth B. Stolarsky, and H. Zassenhaus, who have generously devoted considerable time to reading portions of the preliminary manuscript, and who have suggested many valuable improvements.

Princeton, New Jersey D. E. K.
October 1968

Preface to the Second Edition

My first plan, when beginning to prepare this new edition, was to make it like the second edition of Volume 1: I went through the entire book and tried to improve every page without greatly perturbing the page numbering. But the number of improvements turned out to be so great that the entire book needed to be typeset again. As a result, I decided to make this book the first test case for a new computer typesetting system I have been developing. I hope that most readers will like the slight changes in format, since my aim has been to produce a book whose typography is of the highest possible quality—superior even to the fine appearance of the previous editions, in spite of the fact that a computer is now involved. If all goes well, the third edition of Volume 1 and the second edition of Volume 3, and all editions of Volumes 4 through 7, will be published in the present style.

The decision to reset this entire book has freed me from the shackles of the previous page numbering, so I have been able to make major refinements and to insert a lot of new material. I estimate that about 45 percent of the book has changed. I did try, however, to keep the exercise numbers from being substantially altered; although many of the old exercises have been replaced by new and better ones, the new exercises tend to relate to the same idea as before. The explosive growth of seminumerical research in recent years has of course made it impossible for me to insert all of the beautiful ideas in this field that have been discovered since 1968; but I think that this edition does contain an up-to-date survey of all the major paradigms and basic theory of the subject, and it seems reasonable to believe that very few of the topics discussed here will ever become obsolete.

The National Science Foundation and the Office of Naval Research have been particularly generous in their support of my research as I work on these books. I am also deeply grateful for the advice and unselfish assistance of many readers, too numerous to mention. In this regard I want to acknowledge especially the help of several people whose contributions have been really major: B. I. Aspvall, R. P. Brent, U. Dieter, M. J. Fischer, R. W. Gosper, D. C. Hoaglin, W. M. Kahan, F. M. Liang, J. F. Reiser, A. G. Waterman, S. Winograd, and M. C. Wunderlich. Furthermore Marion Howe and other people in the Addison–Wesley production department have been enormously helpful in untangling literally thousands of hand-written inserts so that a very chaotic manuscript has come out looking reasonably well-organized. I suppose some mistakes still remain, or have crept in, and I would like to fix them; therefore I will cheerfully pay $2.00 reward to the first finder of each technical, typographical, or historical error.

Stanford, California D. E. K.
July 1980

'Defendit numerus,' [there is safety in numbers]
is the maxim of the foolish;
'Deperdit numerus,' [there is ruin in numbers]
of the wise.

—C. C. COLTON (1820)

NOTES ON THE EXERCISES

THE EXERCISES in this set of books have been designed for self-study as well as classroom study. It is difficult, if not impossible, for anyone to learn a subject purely by reading about it, without applying the information to specific problems and thereby being encouraged to think about what has been read. Furthermore, we all learn best the things that we have discovered for ourselves. Therefore the exercises form a major part of this work; a definite attempt has been made to keep them as informative as possible and to select problems that are enjoyable to solve.

In many books, easy exercises are found mixed randomly among extremely difficult ones. This is sometimes unfortunate because readers like to know in advance how long a problem ought to take—otherwise they may just skip over all the problems. A classic example of such a situation is the book *Dynamic Programming* by Richard Bellman; this is an important, pioneering work in which a group of problems is collected together at the end of some chapters under the heading "Exercises and Research Problems," with extremely trivial questions appearing in the midst of deep, unsolved problems. It is rumored that someone once asked Dr. Bellman how to tell the exercises apart from the research problems, and he replied, "If you can solve it, it is an exercise; otherwise it's a research problem."

Good arguments can be made for including both research problems and very easy exercises in a book of this kind; therefore, to save the reader from the possible dilemma of determining which are which, *rating numbers* have been provided to indicate the level of difficulty. These numbers have the following general significance:

Rating	*Interpretation*
00	An extremely easy exercise that can be answered immediately if the material of the text has been understood; such an exercise can almost always be worked "in your head."
10	A simple problem that makes you think over the material just read, but it is by no means difficult. It should be possible to do this in one minute at most; pencil and paper may be useful in obtaining the solution.
20	An average problem that tests basic understanding of the text material, but you may need about fifteen or twenty minutes to answer it completely.

30 A problem of moderate difficulty and/or complexity; this one may involve over two hours' work to solve satisfactorily.

40 Quite a difficult or lengthy problem that would be suitable for a term project in classroom situations. It is expected that a student will be able to solve the problem in a reasonable amount of time, but the solution is not trivial.

50 A research problem that has not yet been solved satisfactorily, as far as the author knew at the time of writing, although many people have tried. If you have found an answer to such a problem, you ought to write it up for publication; furthermore, the author of this book would appreciate hearing about the solution as soon as possible (provided that it is correct).

By interpolation in this "logarithmic" scale, the significance of other rating numbers becomes clear. For example, a rating of *17* would indicate an exercise that is a bit simpler than average. Problems with a rating of *50* that are subsequently solved by some reader may appear with a *45* rating in later editions of the book.

The author has earnestly tried to assign accurate rating numbers, but it is difficult for the person who makes up a problem to know just how formidable it will be for someone else to find a solution; and everyone has more aptitude for certain types of problems than for others. It is hoped that the rating numbers represent a good guess as to the level of difficulty, but they should be taken as general guidelines, not as absolute indicators.

This book has been written for readers with varying degrees of mathematical training and sophistication; as a result, some of the exercises are intended only for the use of more mathematically inclined readers. The rating is preceded by an *M* if the exercise involves mathematical concepts or motivation to a greater extent than necessary for someone who is primarily interested only in programming the algorithms themselves. An exercise is marked with the letters "*HM*" if its solution necessarily involves a knowledge of calculus or other higher mathematics not developed in this book. An "*HM*" designation does *not* necessarily imply difficulty.

Some exercises are preceded by an arrowhead, "►"; this designates problems that are especially instructive and that are especially recommended. Of course, no reader/student is expected to work *all* of the exercises, so those that seem to be the most valuable have been singled out. (This is not meant to detract from the other exercises!) Each reader should at least make an attempt to solve all of the problems whose rating is *10* or less; and the arrows may help to indicate which of the problems with a higher rating should be given priority.

Solutions to most of the exercises appear in the answer section. Please use them wisely; do not turn to the answer until you have made a genuine effort to solve the problem by yourself, or unless you do not have time to work this particular problem. *After* getting your own solution or giving the problem a decent try, you may find the answer instructive and helpful. The solution given will often be quite short, and it will sketch the details under the assumption that you have earnestly tried to solve it by your own means first. Sometimes the solution gives less information than was asked; often it gives more. It is quite possible that you may have a better answer than the one published here, or you may have found an error in the published solution; in such a case, the author will be pleased to know the details. Later editions of this book will give the improved solutions together with the solver's name where appropriate.

When working an exercise you may generally use the answers to previous exercises, unless specifically forbidden from doing so. The rating numbers have been assigned with this in mind; thus it is possible for exercise $n + 1$ to have a lower rating than exercise n, even though it includes the result of exercise n as a special case.

Summary of codes:		*00*	Immediate
		10	Simple (one minute)
		20	Medium (quarter hour)
▶	Recommended	*30*	Moderately hard
M	Mathematically oriented	*40*	Term project
HM	Requiring "higher math"	*50*	Research problem

EXERCISES

▶ **1.** [*00*] What does the rating "*M20*" mean?

2. [*10*] Of what value can the exercises in a textbook be to the reader?

3. [*M50*] Prove that when n is an integer, $n > 2$, the equation $x^n + y^n = z^n$ has no solution in positive integers x, y, z.

Exercise is the beste intrument in learnyng.

—ROBERT RECORDE (*The Whetstone of Witte*, 1557)

CONTENTS