☆ THE ART OF COMPUTER PROGRAMMING ☆

☆☆☆ ERRATA TO VOLUME 1 (2nd edition) ☆☆☆

This document is a transcript of the notes that I made in my personal copy of *The Art of Computer Programming*, Volume 1 (second edition) between 1975 and 1996. When I prepared the third edition in 1997, I made thousands of additional changes; but all of the major corrections and most of the major amendments are reported here.

Three levels of updates — "errors," "amendments," and "plans" — appear, indicated by three different typographic conventions:

▸**Page 666** line 1 ______ 04 Jul 1776

Technical or typographical errors (aka bugs) are the most critical items, so they are flagged with a '▸' preceding the page number. The date on which I first was told about the bug is shown; this is the effective date on which I paid the finder's fee. The necessary corrections are indicated in a straightforward way. If, for example, the book says 'n' where it should have said '$n+1$', the change is shown thus:

$n \rightsquigarrow n+1$

Page 666 line 2 ______ 14 Jul 1789

Amendments to the text appear in the same format as bugs, but without the '▸'. These are things I wish I had known about or thought of when I wrote the original text, so I added them later. The date is the date I drafted the new text.

Page 666 line 3 . . . . 20 Nov 1917

Plans for the future represent a third kind of item. In such notes I sketched my intentions about things that I wasn't ready to flesh out further when I wrote them down. You can identify these items because they're written in slanted type, and preceded by a bunch of dots '. . . . .' leading to the date on which I recorded the plan in my files.

My shelves at home are bursting with preprints and reprints of significant research results that I want to digest and summarize, where appropriate, in the ultimate edition of Volume 1. I didn't do that in the third edition because I would surely have to do it over again later: New results continue to pour forth at a great rate, and I will have time to rewrite that volume only once. Volumes 4 and 5 need to be finished first. So I've put most of my effort so far into writing up those parts of the the total picture that seem to have converged to their near-final form. It follows, somewhat paradoxically, that the updates in this document are most current in the areas where there has been least activity.

On the other hand I do believe that the changes listed here bring Volume 1 completely up to date in two respects: (1) All of the research problems in the previous edition — i.e., all exercises that were rated 46 and above — have received new ratings of 45 or less whenever I learned of a solution; and in such cases, the answer now refers to that solution. Several of the research problems have also been rescinded or restated. (2) All of the historical information about pioneering developments has been amended whenever new details have come to my attention.

The ultimate, glorious, future editions of Volumes 1–3 are works in progress. Please let me know of any improvements that you think I ought to make. Send your comments either by snail mail to D. E. Knuth, Computer Science, Gates Building 4B, Stanford University, Stanford CA 94305-9045, or by email to `taocp@cs.stanford.edu`*. (Use email for book suggestions only, please — all other correspondence is returned unread to the sender, because I have no time to read ordinary email.) Although I'm working full time on Volume 4 these days, I will try to reply to all such messages within a year of receipt. Current news about The Art of Computer Programming is posted on*

`http://www-cs-faculty.stanford.edu/~knuth/taocp.html`

and updated regularly.

— Don Knuth, April 1997

In addition to the errors listed here,
about half of the occurrences of 'which' in volumes one and three
should be changed to 'that'.

— DONALD E. KNUTH (*The Art of Computer Programming Errata et Addenda*, 1981)

Writing a series like The Art of Computer Programming
is similar to painting the Forth Rail Bridge.
No sooner is it finished than the job
must be started again.

— MALCOLM CLARK (1992)

☆☆☆☆☆ FUNDAMENTAL ALGORITHMS ☆☆☆☆☆

Page v and the following pages . . . . 09 Feb 1995

(In the third edition I will revise this preface extensively, so that it will reflect the actual scope of these books instead of the state they were in when Volume 1 was first published.)

Page xii additional abbreviations for frequently quoted journals ―― 09 Feb 1995

SICOMP = SIAM Journal on Computing

FOCS = IEEE Symposium on Foundations of Computer Science

SODA = ACM-SIAM Symposium on Discrete Algorithms

STOC = ACM Symposium on Theory of Computing

Crelle = Journal für die reine und angewandte Mathematik

Page xviii line 16 ―― 28 Feb 1997

book. $\leadsto$ book. If the rating number is divided by 5, the remainder indicates the amount of detailed work required. Thus, an exercise rated *24* may take longer to solve than an exercise that is rated *25*, but the latter will require more creativity.

Page xix new rating for exercise 4 ―― 25 Mar 1995

M50 $\leadsto$ *HM45*

Page xix new quote for bottom of page ―― 14 May 1976

We can face our problem.
We can arrange such facts as we have
with order and method.
— HERCULE POIROT, in *Murder on the Orient Express* (1934)

Page xxi line 9 ―― 14 Mar 1995

*2.3.4.2. $\leadsto$ 2.3.4.2.

▸**Page 1** line 11 ―― 23 Sep 1983

ADA AUGUSTA $\leadsto$ AUGUSTA ADA

Page 1 replacement for bottom 7 lines ―― 12 Aug 1995

It comes from the name of a famous Persian textbook author, Abū ‘Abd Allāh Muḥammad ibn Mūsā al-Khwārizmī (c. 825) — literally, “Father of Abdullah, Mohammed, son of Moses, native of Khwārizm.” The Aral Sea in Central Asia was once known as Lake Khwārizm, and the Khwārizm region is located in the Amu River basin just south of that sea. Al-Khwārizmī wrote the celebrated book *Kitab al jabr wa’l-muqābala* (“Rules of restoring and equating”); another word, “algebra,” stems from the title of his book, which was a systematic study of the solution of linear and quadratic equations. [For notes on al-Khwārizmī’s life and work, see H. Zemanek, *Lecture Notes in Computer Science* **122** (1981), 1–81.]

Page 2 last line of Algorithm E, also Fig. 1 ——— 02 Apr 1993

[Interchange.] $\leadsto$ [Reduce.]

Page 4 line −10 ——— 09 Aug 1995

routine, $\leadsto$ *routine, rigmarole,*

Page 5 line 9 ——— 28 Dec 1996

“incommensurate.” $\leadsto$ “incommensurable.” Another example of a nonterminating computational method is a *reactive process*, which continually interacts with its environment.

Page 5 line −12 ——— 28 Dec 1996

begins. $\leadsto$ begins, or dynamically as the algorithm runs.

Page 8 and throughout the books . . . . 19 Mar 1995

Equation numbers like (1), (2), and (3) will henceforth be typeset with “oldstyle” digits like (1), (2), *and* (3). *All corrections in this list tacitly assume that such a change has already been made.*

▸**Page 8** replacement for (2) ——— 25 Jun 1993

$$
\begin{aligned}
f\big((m,n)\big) &= (m,n,0,1); \qquad f\big((n)\big) = (n);\\
f\big((m,n,r,1)\big) &= (m,n,\text{remainder of } m \text{ divided by } n,2);\\
f\big((m,n,r,2)\big) &= (n) \quad \text{if} \quad r=0, \quad (m,n,r,3) \quad \text{otherwise};\\
f\big((m,n,r,3)\big) &= (n,p,p,1).
\end{aligned}
\tag{2}
$$

Page 9 lines 3–5 ——— 30 Jan 1997

Markov in 1951, . . . 51085). $\leadsto$ Markov in his book *The Theory of Algorithms* [*Trudy Mat. Inst. Akad. Nauk* **42** (1954), 1–376], later revised and enlarged by N. M. Nagorny (Moscow: Nauka, 1984; English edition, Dordrecht: Kluwer, 1988).

Page 9 line 2 of exercise 3 ——— 02 Apr 1993

interchange values $\leadsto$ rename the values

Page 9 line −5 ——— 15 Jan 1996

of C_1”: $\leadsto$ of C_1” or “C_2 simulates C_1”:

Page 11 new paragraph to follow line 5 ——— 25 Mar 1995

An expansion and more leisurely presentation of much of the following material can be found in the book *Concrete Mathematics* by Graham, Knuth, and Patashnik, second edition (Reading, Mass.: Addison-Wesley, 1994). That book will be called simply *CMath* when we need to refer to it later.

▸**Page 12** line 11 after Fig. 2 ——— 06 May 1996

denote the numbers $\leadsto$ denote the number

▸**Page 17** line 16 ——— 06 May 1996

tions is $\rightsquigarrow$ tion is

Page 18 line 7 ——— 31 Jan 1997

39–45. The idea $\rightsquigarrow$ 39–45. Later authors found it advantageous to reverse Floyd's direction, going from an assertion that should hold *after* an operation to the "weakest precondition" that must hold *before* the operation is done; such an approach makes it possible to discover new algorithms that are guaranteed to be correct, if we start from the specifications of the desired output and work backwards. [See E. W. Dijkstra, *CACM* **18** (1975), 453–457; *A Discipline of Programming* (Prentice–Hall, 1976).]

The idea

Page 18 line 12 ——— 11 Aug 1981

99.] $\rightsquigarrow$ 99. See also A. M. Turing's early comments about verification in *Report of a Conference on High Speed Automatic Calculating Machines* (Cambridge Univ., 1949), 67–68 and figs.; reprinted with commentary by F. L. Morris and C. B. Jones in *Ann. Hist. Computing* **6** (1984), 139–143.]

Page 18 insert quotation before the exercises ——— 13 Jul 1995

The understanding of the theory of a routine
may be greatly aided by providing, at the time of construction
one or two statements concerning the state of the machine
at well chosen points. ...
In the extreme form of the theoretical method
a watertight mathematical proof is provided for the assertions.
In the extreme form of the experimental method
the routine is tried out on the machine with a variety of initial
conditions and is pronounced fit if the assertions hold in each case.
Both methods have their weaknesses.

— A. M. TURING, Ferranti Mark I Programming Manual (1950)

Page 19 bottom three lines ——— 08 Apr 1975

$T \le 3n_0$, where n_0 is the original value of n, $\rightsquigarrow$ $T \le n$,
$3n_0$ $\rightsquigarrow$ $3n$

Page 21 new paragraph to precede line -11 ——— 09 Jun 1995

Difficult problems about integers are often solved by working with real numbers, and difficult problems about real numbers are often solved by working with a still more general class of values called complex numbers. A *complex number* is a quantity z of the form $z = x + iy$, where x and y are real and i is a special quantity that satisfies the equation $i^2 = -1$. We call x and y the "real part" and the "imaginary part" of z, and we define the absolute value $|z|$ to be $\sqrt{x^2 + y^2}$. The theory of complex numbers is in many ways simpler and more beautiful than the theory of real numbers, but it is usually considered to be an advanced topic. Therefore we shall concentrate on real numbers in this book, except when real numbers turn out to be unnecessarily complicated.

▸**Page 22** replacement for (7) 01 Feb 1997

$$10^{0.30102999} = 1.9999999739\ldots, \qquad 10^{0.30103000} = 2.0000000199\ldots, \qquad (7)$$

Page 22 to fill a lacuna 25 Mar 1995

I forgot to define the notations used for intervals (except when Appendix B of Volume 2 mentions half-open intervals). In the next editions I'll adopt the conventions $(a\,.\,.\,b)$, $[a\,.\,.\,b)$, etc., of Ramshaw and Hoare, as used in Concrete Mathematics.

Page 25 line 1 of exercise 11 06 May 1996

$x = \log_{10} 2 \leadsto x \approx \log_{10} 2$

Page 26 line 15 14 May 1976

hardware. $\leadsto$ hardware. The idea goes back in essence to Henry Briggs, who used it (in decimal rather than binary form) to compute logarithm tables, published in 1624.

Page 27 additional conventions 11 Feb 1995

I will be using the notation $\sum_{j=1}^{n}$ as well as $\sum_{1\le j\le n}$ in the next editions, in cases where it makes a more compact formula.

Page 27 lines 9 and 10 25 Mar 1995

was introduced by J. Lagrange in 1772. $\leadsto$ with definite limits was introduced by J. Fourier in 1820.

Page 28 lines 4–5 after (5) 22 Jul 1996

i.e., for each ... $p(j) = i$, and conversely. $\leadsto$
more precisely, for each ... $p(j) = i$.

Page 29 line -5 01 Jul 1987

$m \le n \leadsto 1 \le m \le n$.

Page 34 replacement for exercise 30 26 Mar 1995

30. [*M23*] (J. Binet, 1812.) Without using induction, prove the identity

$$\left(\sum_{j=1}^{n} a_j x_j\right)\left(\sum_{j=1}^{n} b_j y_j\right) = \left(\sum_{j=1}^{n} a_j y_j\right)\left(\sum_{j=1}^{n} b_j x_j\right) + \sum_{1\le j<k\le n} (a_j b_k - a_k b_j)(x_j y_k - x_k y_j)\,.$$

[An important special case arises when $w_1, \ldots, w_n, z_1, \ldots, z_n$ are arbitrary complex numbers and we set $a_j = \overline{x}_j = w_j$, $b_j = \overline{y}_j = \overline{z}_j$:

$$\left(\sum_{j=1}^{n} |w_j|^2\right)\left(\sum_{j=1}^{n} |z_j|^2\right) = \left|\sum_{j=1}^{n} w_j z_j\right|^2 + \sum_{1\le j<k\le n} |w_j\overline{z}_k - w_k\overline{z}_j|^2\,.$$

The terms $|w_j\overline{z}_k - w_k\overline{z}_j|^2$ are nonnegative, so the famous "Cauchy-Schwarz inequality"

$$\left(\sum_{j=1}^{n} |w_j|^2\right)\left(\sum_{j=1}^{n} |z_j|^2\right) \ge \left|\sum_{j=1}^{n} w_j z_j\right|^2$$

is a consequence of Binet's formula.]

Page 37 new exercise ______ 08 Jun 1995

47. [*M27*] (C. Krattenthaler.) Prove that

$$\det\begin{pmatrix} (x+q_2)(x+q_3) & (x+p_1)(x+q_3) & (x+p_1)(x+p_2) \\ (y+q_2)(y+q_3) & (y+p_1)(y+q_3) & (y+p_1)(y+p_2) \\ (z+q_2)(z+q_3) & (z+p_1)(z+q_3) & (z+p_1)(z+p_2) \end{pmatrix} = (x-y)(x-z)(y-z)(p_1-q_2)(p_1-q_3)(p_2-q_3)\,,$$

and generalize this equation to an identity for an $n\times n$ determinant in $3n-2$ variables $x_1, \ldots, x_n, p_1, \ldots, p_{n-1}, q_2, \ldots, q_n$. Compare your formula to the result of exercise 38.

Page 39 and throughout the books . . . . 01 Nov 1995

I will adopt the notation '$r \perp s$' for the relation "r and s are relatively prime." This will replace '$\gcd(r,s)=1$' in many places.

▸**Page 40** line 2 of exercise 19 ______ 10 Nov 1986

$nn' \bmod m = 1 \leadsto nn' \equiv 1 \pmod{m}$

▸**Page 41** line 2 of exercise 28 ______ 22 Oct 1985

$a^{\varphi(m)} \bmod m = 1 \leadsto a^{\varphi(m)} \equiv 1 \pmod{m}$

▸**Page 42** line 4 ______ 23 Feb 1981

$\sum_{1\le k<n}\lceil k/2\rceil \leadsto \sum_{1\le k\le n}\lceil k/2\rceil$

Page 42 replacement for exercise 38 ______ 08 Jun 1995

38. [*M26*] (E. Busche, 1909.) Prove that, for all real x and y with $y>0$,

$$\sum_{0\le k<y}\left\lfloor x+\frac{k}{y}\right\rfloor = \bigl\lfloor xy + \lfloor x+1\rfloor(\lceil y\rceil - y)\bigr\rfloor\,.$$

In particular, when y is a positive integer n, we have the important formula

$$\lfloor x\rfloor + \left\lfloor x+\frac{1}{n}\right\rfloor + \cdots + \left\lfloor x+\frac{n-1}{n}\right\rfloor = \lfloor nx\rfloor\,.$$

Page 44 new exercise ______ 28 Sep 1996

49. [*M30*] Suppose the function $f(x)$ satisfies the laws (i) $f(n+x)=n+f(x)$ for all integers n; (ii) $f(x)=f\bigl(f(nx)/n\bigr)$ for all positive integers n. Prove that either $f(x)=\lfloor x\rfloor$ for all rational x, or $f(x)=\lceil x\rceil$ for all rational x.

Page 44 add a footnote to line 3 after (1) ______ 19 Apr 1977

book. $\leadsto$ book.*

---

* In fact, permutations are so important, Vaughan Pratt has suggested calling them "perms." As soon as Pratt's convention is established, textbooks of computer science will be somewhat shorter, and perhaps less expensive.

Page 46 line 3 after (7) ______ 26 Mar 1995

1.2.11.2. $\leadsto$ 1.2.11.2. Exercise 24 gives a simple proof of a less precise result.

Page 48 line -3 ______ 08 Jun 1995

very seldom used, and $\leadsto$ less common;

Page 49 line 3 ——— 01 Jun 1990

$$\Gamma(x) = \frac{x!}{x} = \lim_{m\to\infty} \frac{m^x m!}{x(x+1)(x+2)\ldots(x+m)}. \qquad (15)$$

Page 49 new paragraph to follow line 4 ——— 09 Jun 1995

Equations (13) and (15) define factorials and the Gamma function for complex values as well as real values; but we generally use the letter z instead of n or x when thinking of a variable that has both real and imaginary parts. The factorial and Gamma functions are related not only by the rule $z! = \Gamma(z+1)$ but also by

$$(-z)!\,\Gamma(z) = \frac{\pi}{\sin \pi z}\,, \qquad (16)$$

which holds whenever z is not an integer. (See exercise 23.)

Page 49 another new paragraph to follow line 4 ——— 19 Dec 1996

Although $\Gamma(z)$ is infinite when z is zero or a negative integer, the function $1/\Gamma(z)$ is well defined for all complex z. (See exercise 1.2.7–24.) Advanced applications of the Gamma function often make use of an important contour integral formula due to Hermann Hankel:

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i}\oint \frac{e^t\,dt}{t^z}; \qquad (17)$$

the path of complex integration starts at $-\infty$, then circles the origin in a counterclockwise direction and returns to $-\infty$. [*Zeitschrift für Math. und Physik* **9** (1864), 1–21.]

Page 49 line 7 ——— 09 Oct 1995

869. $\leadsto$ 869; see also J. Dutka, *Archive for History of Exact Sciences* **31** (1984), 15–34.

Page 50 line 2 ——— 01 Oct 1984

any integer $\leadsto$ *any* positive integer

Page 50 line 2 of exercise 18 ——— 01 Jun 1995

1656 $\leadsto$ 1655

Page 50 exercise 21 ——— 03 May 1995

(Faa di Bruno's formula.) $\leadsto$ (L. F. A. Arbogast, 1800.)

Page 51 new exercise ——— 09 Jun 1995

23. [*HM20*] Prove (16), given that $\pi z \prod_{n=1}^{\infty}(1 - z^2/n^2) = \sin \pi z$.

Page 51 new exercise ——— 03 Jun 1995

▸**24.** [*HM21*] Prove the handy inequalities

$$\frac{n^n}{e^{n-1}} \le n! \le \frac{n^{n+1}}{e^{n-1}}\,, \qquad \text{integer } n \ge 1.$$

[*Hint:* $1 + x \le e^x$ for all real x.]

▸**Page 51** line 13 — 28 Feb 1978

manner $\leadsto$ matter

Page 52 line 6 after Table 1 — 25 Aug 1976

Szu-yuen $\leadsto$ *Szu-yüan*

Page 52 lower half — 09 Jun 1995

line -17: earliest known appearance $\leadsto$ earliest known detailed discussion
line -14: the Chandaḥ-Sûtra. $\leadsto$ Piṅgala's *Chandaḥ-Sûtra*. [See G. Chakravarti, *Bull. Calcutta Math. Soc.* **24** (1932), 79–88.]

Page 57 equation (18) — 31 Jul 1976

$n \geq 0 \leadsto n$

Page 58 top half . . . . . 26 Mar 1995

I'll state these identities in more general form, as on page 169 of Concrete Math; then, for example, Eq. (22) *will apply on page 60, line* -7.

▸**Page 59** line 14 — 08 Dec 1986

if $n \geq 0$? $\leadsto$ if n is a nonnegative integer?

Page 60 notations like $[a \leq 3]$ . . . . . 11 Feb 1995

I'll introduce Iverson's bracket notation in section 1.2.3, and use it as I do in Concrete Mathematics, except that I'll also retain the Kronecker delta.

Page 64 new material to follow line 14 — 26 Mar 1995

On the other hand this sum does have a closed form as a function of n when m is an explicit negative integer; for example,

$$\sum_{k=0}^{n} \binom{-2}{k} = (-1)^n \left\lceil \frac{n+1}{2} \right\rceil .$$

There is also a simple formula

$$\sum_{k=0}^{n} \binom{m}{k} \left(k - \frac{m}{2}\right) = -\frac{m}{2} \binom{m-1}{n}$$

for a sum that looks as though it should be harder, not easier.

How can we decide when to stop working on a sum that resists simplification? Fortunately, there is now a good way to answer that question in many important cases: An algorithm due to R. W. Gosper and D. Zeilberger will discover closed forms in binomial coefficients when they exist, and will prove the impossibility when they do not exist. The Gosper-Zeilberger algorithm is beyond the scope of this book, but it is explained in *CMath* §5.8.

Page 65 line 2 — 09 Jun 1995

Note that (20) is a trinomial coefficient. $\leadsto$ Both sides of Eq. (20) are the trinomial coefficient $\binom{r}{k,\, m-k,\, r-m}$.

Page 65 more to come . . . . 11 Feb 1995

Here I will introduce basic material about hypergeometric functions, although I'll refer to Concrete Mathematics for most of the details. The article by Dutka, cited at the end of Section 1.2.5, contains much historical information about hypergeometrics. I'll also update the material on Stirling numbers, based on what I've learned since the 60s; they will be called Stirling cycle numbers and Stirling subset numbers, not Stirling numbers of the first and second kind. I'll also point out that the notations $\left\{ {n \atop k} \right\}$ *and* $\left[{n \atop k} \right]$ *were introduced by Karamata in 1935 [Mathematica (Cluj)* **9** *(1935), 164–178]. I'll leave out some of the "expansion formulas" on page 67, referring instead to CMath for a detailed treatment.*

Page 66 replacement for third line of the caption ________ 21 Sep 1976

see D. E. Barton, F. N. David, and M. Merrington, *Biometrika* **47** (1960), 439–445; **50** (1963), 169–176.

Page 71 in 14 places ________ 08 Apr 1975

$B \rightsquigarrow \mathrm{B}$ (roman, not italic) in notation for the Beta function; namely in line 1, line 2, line 3 (twice), line 4 (thrice), line 5(twice), line 7, line 10 (twice), line 12, line 15

▸**Page 72** replacement for line 4 of exercise 53 ________ 29 Nov 1984

$$\binom{-1/2}{n} = \frac{(-1)^n}{2^{2n}}\binom{2n}{n}, \quad \binom{1/2}{n} = \frac{(-1)^{n-1}}{2^{2n}(2n-1)}\binom{2n}{n} = \frac{(-1)^{n-1}}{2^{2n-1}(2n-1)}\binom{2n-1}{n} - \delta_{n0}$$

Page 72 exercise 56 ________ 26 Sep 1996

binomial number system $\rightsquigarrow$ *combinatorial number system*

Page 72 exercise 58 ________ 26 Mar 1995

line 1: [*M21*] $\rightsquigarrow$ [*M23*]

new line to be inserted after line 2: Find q-nomial generalizations of the fundamental identities (17) and (21).

▸**Page 72** exercise 59 ________ 15 Nov 1978

$A_{n(k-1)} + \binom{n}{k}$. $\rightsquigarrow$ $A_{(n-1)(k-1)} + \binom{n}{k}$ for $nk > 0$.

Page 73 replacement for exercise 63 ________ 26 Mar 1995

63. [*M30*] If l, m, and n are integers and $n \ge 0$, prove that

$$\sum_{j,k} (-1)^{j+k}\binom{j+k}{k+l}\binom{r}{j}\binom{n}{k}\binom{s+n-j-k}{m-j} = (-1)^l\binom{n+r}{n+l}\binom{s-r}{m-n-l}.$$

Page 73 new exercises ________ 09 Jun 1995

65. [*M29*] Let n be a positive integer, and suppose x and y are real numbers satisfying $n \le y \le x \le y+1$. Then $\binom{y}{n+1} \le \binom{x}{n+1} \le \binom{y+1}{n+1} = \binom{y}{n+1} + \binom{y}{n}$, so there is a unique real number z such that

$$\binom{x}{n+1} = \binom{y}{n+1} + \binom{z}{n}, \qquad n-1 \le z \le y.$$

Prove that

$$\binom{x}{n} \le \binom{y}{n} + \binom{z}{n-1}.$$

[*Hint:* Consider the expansion $\binom{x}{n+1} = \binom{z+1}{n+1} + \sum_{k\ge 0}\binom{z-k}{n-k}\binom{x-z-1+k}{k+1}$.]

▸**66.** [*M20*] We often need to know that binomial coefficients aren't too large. Prove the easy-to-remember upper bound

$$\binom{n}{k} \le \left(\frac{ne}{k}\right)^k, \qquad \text{when } n \ge k \ge 0.$$

Page 73 new exercise ———— 08 May 1996

67. [*M25*] (A. De Moivre.) Prove that, if n is a nonnegative integer,

$$\sum_k \binom{n}{k} p^k (1-p)^{n-k} |k - np| = 2\lceil np\rceil \binom{n}{\lceil np\rceil} p^{\lceil np\rceil}(1-p)^{n+1-\lceil np\rceil}.$$

Page 75 replacement for lines 3 and 4 ———— 09 Jun 1995

These results are due to Euler; for discussion and proof, see *CMath*, §6.5.

Page 77 new rating for exercise 13 ———— 26 Sep 1987

[*M22*] $\rightsquigarrow$ [*HM22*]

▸**Page 78** bottom half ———— 16 Oct 1985

line −13: originated $\rightsquigarrow$ published
line −11: *Liber Abbaci* $\rightsquigarrow$ *Liber Abaci*

Page 79 new paragraph to follow line 7 ———— 10 Jun 1995

Before Fibonacci wrote his work, the sequence F_n had already been discussed by Indian scholars, who had long been interested in rhythmic patterns that are formed from one-beat and two-beat notes. The number of such rhythms having n beats altogether is F_{n+1}; therefore both Gopāla (before 1135) and Hemachandra (c. 1150) mentioned the numbers 1, 2, 3, 5, 8, 13, 21, … explicitly. [See P. Singh, *Historia Math.* **12** (1985), 229–244; see also exercise 4.5.3–32.]

Page 79 lines 8–10 ———— 25 Jun 1976

Kepler … plant life. $\rightsquigarrow$ Kepler, 1611, who was musing about the numbers he saw around him [J. Kepler, *The Six-Cornered Snowflake* (Oxford: Clarendon Press, 1966), p 21].

Page 79 line −6 ———— 22 Aug 1996

1961). $\rightsquigarrow$ 1961). The fact that the ratio F_{n+1}/F_n approaches ϕ was known to the early European reckoning master Simon Jacob, who died in 1564 [see P. Schreiber, *Historia Math.* **22** (1995), 422–424].

Page 80 lines 5–6 ———— 30 Oct 1995

(New York: Chelsea, 1952) $\rightsquigarrow$ (Carnegie Inst. of Washington, 1919)

Page 80 line 9 ——— 16 Feb 1995

quoted ... Cassini $\leadsto$ discovered relations, mentioned by Kepler in a letter he wrote in 1608 but first published by J. D. Cassini

Page 84 and 85, new notation in exercise 29 ——— 27 Mar 1995

▸**29.** [*M23*] (*Fibonomial coefficients.*) Édouard Lucas defined the quantities

$$\binom{n}{k}_{\mathcal{F}} = \frac{F_n F_{n-1} \ldots F_{n-k+1}}{F_k F_{k-1} \ldots F_1} = \prod_{j=1}^{k} \left(\frac{F_{n-k+j}}{F_j} \right)$$

in a manner analogous to binomial coefficients. (a) Make a table of $\binom{n}{k}_{\mathcal{F}}$ for $0 \le k \le n \le 6$. (b) Show that $\binom{n}{k}_{\mathcal{F}}$ is always an integer because we have

$$\binom{n}{k}_{\mathcal{F}} = F_{k-1}\binom{n-1}{k}_{\mathcal{F}} + F_{n-k+1}\binom{n-1}{k-1}_{\mathcal{F}}.$$

[The same notational change should be made in exercise 30 and in the answers on page 492.]

Page 85 line 4 ——— 27 Mar 1995

D. Jarden $\leadsto$ D. Jarden, T. Motzkin

Page 85 exercise 35 ——— 17 Aug 1996

line 1: [*M24*] $\leadsto$ [*M26*]
line 5: every number has a unique representation. $\leadsto$ every nonnegative number has a unique representation. What are the representations of integers?

Page 86 new exercises ——— 10 Jun 1995

40. [*M25*] Solve the recurrence

$$f(1) = 0; \quad f(n) = \min_{0<k<n} \max\bigl(1 + f(k), 2 + f(n-k)\bigr), \ n > 1.$$

▸**41.** [*M25*] (Yuri Matiyasevich, 1990.) Let $f(x) = \lfloor x + \phi^{-1} \rfloor$. Prove that if $n = F_{k_1} + \cdots + F_{k_r}$ is the representation of n in the Fibonacci number system of exercise 34, then $F_{k_1+1} + \cdots + F_{k_r+1} = f(\phi n)$. Find a similar formula for $F_{k_1-1} + \cdots + F_{k_r-1}$.

42. [*M26*] (D. A. Klarner.) Show that if m and n are nonnegative integers, there is a unique sequence of indices $k_1 \ll k_2 \ll \cdots \ll k_r$ such that

$$m = F_{k_1} + F_{k_2} + \cdots + F_{k_r}, \qquad n = F_{k_1+1} + F_{k_2+1} + \cdots + F_{k_r+1}.$$

(The k's may be negative, and r may be zero.)

▸**Page 86** line −12 ——— 16 Dec 1979

$|z| < z_0 \leadsto |z| < |z_0|$

Page 86 line −11 ——— 16 Oct 1985

The sequence converges $\leadsto$ The series converges

▸**Page 88** bottom line ——— 01 Apr 1979

$$\sum_{1 \le j < m} \leadsto \sum_{j=0}^{m-1}$$

Page 90 on the line between (23) and (24) — 14 May 1976

series $\leadsto$ *series* (cf. (17) and (18))

Page 90 generalize equation (25) — 14 May 1976

$$\frac{1}{(1-z)^{m+1}}\ln\left(\frac{1}{1-z}\right)=\sum_{k\ge 1}(H_{m+k}-H_m)\binom{m+k}{k}z^k\,. \tag{25}$$

Page 90 insert new sentence just after (26) — 20 Aug 1996

For further generalizations, including many sums of harmonic numbers, see D. A. Zave, *Inf. Proc. Letters* **5** (1976), 75–77; J. Spieß, *Math. Comp.* **55** (1990), 839–863.

Page 91 after (30) . . . . 27 Mar 1995

I'll say more about convolution polynomials here and how they generalize this equation and (21)*; see the new exercises in Section 4.7.*

Page 92 extension . . . . 28 Mar 1995

I will introduce the notation $[z^n]\,G(z)$ *here (and in Appendix A) for extracting the coefficient of* z^n *in a power series* $G(z)$*. And I can refer to my paper in A Classical Mind (Prentice-Hall, 1994), 247–258, for examples of how to use it fluently.*

Page 92 lines −10 to −6 — 10 Feb 1995

expression for h_{12}. ... Table 24.2. $\leadsto$ expression for h_{12}. By differentiating (34) it is not difficult to derive the recurrence

$$h_n=\frac{1}{n}(S_1h_{n-1}+S_2h_{n-2}+\cdots+S_nh_0),\qquad n\ge 1\,. \tag{35}$$

Page 92 line −4 — 26 Mar 1995

697. $\leadsto$ 697; his approach is continued in *CMath*, Chapter 7. See also the book *generatingfunctionology* by H. S. Wilf, second edition (Academic Press, 1994).

Page 92 quotation to precede the exercises — 03 Oct 1996

A generating function is a clothesline
on which we hang up a sequence of numbers for display.
— H. S. WILF (1989)

Page 93 replacement for exercise 11 — 12 Jul 1995

▸**11.** [*M25*] Equation (35) can also be used to express the S's in terms of the h's: We find $S_1=h_1$, $S_2=2h_2-h_1^2$, $S_3=3h_3-3h_1h_2+h_1^3$, etc. What is the coefficient of $h_1^{k_1}h_2^{k_2}\ldots h_m^{k_m}$ in this representation of S_m, when $k_1+2k_2+\cdots+mk_m=m$?

Page 94 new exercise — 27 Mar 1995

20. [*M21*] For what coefficients c_{mk} is $\sum_{n\ge 0}n^mz^n=\sum_{k=0}^{m}c_{mk}z^k/(1-z)^{k+1}$?

Page 94 new exercise — 12 Jul 1995

The former exercise 11 becomes exercise 21.

Page 94 new exercise ——— 10 Jun 1995

22. [*M33*] (L. Carlitz.) (a) Prove that for all integers $m \geq 1$ there are polynomials $f_m(z_1, \ldots, z_m)$ and $g_m(z_1, \ldots, z_m)$ such that the formula

$$\sum_{k_1,\ldots,k_m \geq 0} \binom{r}{n-k_1}\binom{k_1}{n-k_2}\ldots\binom{k_{m-1}}{n-k_m} z_1^{k_1}\ldots z_m^{k_m} = f_m(z_1,\ldots,z_m)^{n-r} g_m(z_1,\ldots,z_m)^r$$

is an identity for all integers $n \geq r \geq 0$.

(b) Generalizing exercise 15, find a closed form for the sum

$$S_n(z_1,\ldots,z_m) = \sum_{k_1,\ldots,k_m \geq 0} \binom{k_1}{n-k_2}\binom{k_2}{n-k_3}\ldots\binom{k_m}{n-k_1} z_1^{k_1}\ldots z_m^{k_m}$$

in terms of the functions f_m and g_m in part (a).

(c) Find a simple expression for $S_n(z_1, \ldots, z_m)$ when $z_1 = \cdots = z_m = z$.

Page 94 new exercises ——— 09 Feb 1997

23. [*M21*] Find a generating function $G(z)$ for which

$$[z^n]\,G(z) = \sum_{k_0+2k_1+4k_2+8k_3+\cdots=n} \binom{r}{k_0}\binom{r}{k_1}\binom{r}{k_2}\binom{r}{k_3}\ldots.$$

24. [*M22*] Prove that, if $G(z)$ is any generating function, we have

$$\sum_k \binom{m}{k} [z^{n-k}]\,G(z)^k = [z^n]\,(1+zG(z))^m.$$

Evaluate both sides of this identity when $G(z)$ is (a) $1/(1-z)$; (b) $(e^z-1)/z$.

▸**25.** [*M23*] Evaluate the sum $\sum_k \binom{n}{k}\binom{2n-2k}{n-k}(-2)^k$ by simplifying the equivalent formula $\sum_k [w^k]\,(1-2w)^n\,[z^{n-k}]\,(1+z)^{2n-2k}$.

26. [*M40*] Explore a generalization of the notation (31), according to which we might write, for example, $[z^2-2z^5]\,G(z) = a_2 - 2a_5$ when $G(z)$ is given by (1).

Page 95 replacement for lines 2–8 ——— 10 Jun 1995

will find m and j such that $m = X[j] = \max_{1\leq i\leq n} X[i]$, where j is the largest index that satisfies this relation.

M1. [Initialize.] Set $j \leftarrow n$, $k \leftarrow n-1$, $m \leftarrow X[n]$. (During this algorithm we will have $m = X[j] = \max_{k<i\leq n} X[i]$.)

M2. [All tested?] If $k = 0$, the algorithm terminates.

M3. [Compare.] If $X[k] \leq m$, go to M5.

M4. [Change m.] Set $j \leftarrow k$, $m \leftarrow X[k]$. (This value of m is a new current maximum.)

M5. [Decrease k.] Decrease k by one and return to M2. ▮

Page 97 clarifying remarks ——— 10 Mar 1981

line 10: $A = k$. $\leadsto$ $A = k$. Let this number be $P_{nk} = n!\,p_{nk}$.
line 14: that $\leadsto$ that $P_{nk} = P_{(n-1)(k-1)} + (n-1)P_{(n-1)k}$, i.e.,

Page 98 caption for Fig. 10 ——— 31 Jul 1976

2.11 $\leadsto$ 2.10

▸**Page 99** line 2 after (14) ——— 10 Feb 1997

rth moment $\leadsto$ rth moment of $A+1$, namely $\sum_k (k+1)^r p_{nk}$,

Page 100 line 2 ——— 13 May 1996

suffice. $\leadsto$ suffice. More complicated algorithms will help us develop an ability to reason fluently about probabilities.

Page 100 line −3 ——— 25 Jun 1996

Omit the equation number (21); we're going to use it for a better equation.

Page 101 lines 2–4 ——— 25 Jun 1996

nontrivial task ... in this case are $\leadsto$ nontrivial task.

Fortunately there is a much simpler way to proceed. By Taylor's theorem we have

$$G(1+z) = G(1) + G'(1)z + \frac{G''(1)}{2!}z^2 + \cdots; \tag{21}$$

therefore we merely have to replace z by $z+1$ in (20) and read off the coefficients:

$$G(1+z) = \frac{1}{n}\frac{(1+z)^{n+1} - 1 - z}{z} = 1 + \frac{n+1}{2}z + \frac{(n+1)(n-1)}{6}z^2 + \cdots.$$

It follows that $G'(1) = \frac{1}{2}(n+1)$, $G''(1) = \frac{1}{3}(n+1)(n-1)$, and the statistics for the uniform distribution are

Page 101 new paragraph, three lines before (23) ——— 03 Aug 1995

The mean or average value of a random quantity X is often called its *expected value*, and denoted by $\mathrm{E}\,X$. The variance of X is then $\mathrm{E}\,X^2 - (\mathrm{E}\,X)^2$. In this notation, the probability generating function for X is $G(z) = \mathrm{E}\,z^X$, the expected value of z^X, in cases when X takes only nonnegative integer values. Similarly, when X is a statement that is either true or false, the probability that X is true is $\Pr(X) = E[X]$, using Iverson's bracket notation.

Page 102 line 5 ——— 14 May 1976

distribution, the $\leadsto$ distribution, we can improve significantly on Chebyshev's inequality: The

Page 102 new material following line 12 ——— 16 May 1996

We often need to know that a random variable is unlikely to be much larger or smaller than its mean value. Two extremely simple yet powerful formulas, called the *tail inequalities*, provide convenient estimates of such probabilities. If X has probability generating function $G(z)$, then

$$\Pr(X \le r) \;\le\; x^{-r}G(x) \qquad \text{for } 0 < x \le 1; \tag{24}$$

$$\Pr(X \ge r) \;\le\; x^{-r}G(x) \qquad \text{for } x \ge 1. \tag{25}$$

The proofs are easy: If $G(z) = p_0 + p_1 z + p_2 z^2 + \cdots$, we have

$$\Pr(X \le r) = p_0 + p_1 + \cdots + p_{\lfloor r \rfloor} \le x^{-r} p_0 + x^{1-r} p_1 + \cdots + x^{\lfloor r \rfloor - r} p_{\lfloor r \rfloor} \le x^{-r} G(x)$$

when $0 < x \le 1$, and

$$\Pr(X \ge r) = p_{\lceil r \rceil} + p_{\lceil r \rceil + 1} + \cdots \le x^{\lceil r \rceil - r} p_{\lceil r \rceil} + x^{\lceil r \rceil + 1 - r} p_{\lceil r \rceil + 1} + \cdots \le x^{-r} G(x)$$

when $x \ge 1$. By choosing a value of x that minimizes or approximately minimizes the right-hand sides of (24) and (25), we often obtain upper bounds that are fairly close to the true tail probabilities on the left-hand sides.

Exercises 22–24 illustrate the tail inequalities in several important cases. These inequalities are special cases of a more general principle first pointed out by A. N. Kolmogorov in his book *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Springer, 1933): If $f(t) \ge s > 0$ for all $t \ge r$, then $\Pr(X \ge r) \le s^{-1} \operatorname{E} f(X)$ whenever $\operatorname{E} f(X)$ exists. We obtain (25) when $f(t) = x^t$ and $s = x^r$.

Page 102 replacement for exercise 6 ———— 25 Jun 1996

6. [*HM27*] We have computed the mean and the variance of the important probability distributions (8), (18), (20). What is the *third* semi-invariant, κ_3, in each of those cases?

Page 103 top ———— 27 Mar 1995

Put exercise 12 first, and renumber it exercise 11. The old exercise 11 becomes:

12. [*HM21*] When $G(z) = p_0 + p_1 z + p_2 z^2 + \cdots$ represents a probability distribution, the quantities $M_n = \sum_k k^n p_k$ and $m_n = \sum_k (k - M_1)^n p_k$ are called the "nth moment" and "nth central moment," respectively. Show that $G(e^t) = 1 + M_1 t + M_2 t^2/2! + \cdots$; then use Arbogast's formula (exercise 1.2.5–21) to show that

$$\kappa_n = \sum_{\substack{k_1, k_2, \ldots, k_n \ge 0 \\ k_1 + 2k_2 + \cdots = n}} \frac{(-1)^{k_1 + k_2 + \cdots + k_n - 1} n!\,(k_1 + k_2 + \cdots + k_n - 1)!}{k_1!\,1!^{k_1}\,k_2!\,2!^{k_2} \ldots k_n!\,n!^{k_n}} M_1^{k_1} M_2^{k_2} \ldots M_n^{k_n}.$$

In particular, $\kappa_1 = M_1$, $\kappa_2 = M_2 - M_1^2$ (as we already knew), $\kappa_3 = M_3 - 3M_1M_2 + 2M_1^3$, $\kappa_4 = M_4 - 4M_1M_3 + 12M_1^2M_2 - 3M_2^2 - 6M_1^4$. What are the analogous expressions for κ_n in terms of the central moments m_2, m_3, …, when $n \ge 2$?

Page 103 exercise 13 ———— 27 Mar 1995

line 1: characteristic functions $\rightsquigarrow$ probability generating functions

replacement for line 3: $\lim_{n \to \infty} e^{-it\mu_n/\sigma_n} G_n(e^{it/\sigma_n}) = e^{-t^2/2}$

line 4: for all … u. $\rightsquigarrow$ for all real values of t. line 6: (1944) $\rightsquigarrow$ (1944), 3–48

line 11: Levy $\rightsquigarrow$ Lévy

Page 104 new exercises ———— 12 Aug 1996

19. [*M21*] If $a_k > a_j$ for $1 \le j < k$, we say that a_k is a *left-to-right maximum* of the sequence $a_1\,a_2 \ldots a_n$. Suppose $a_1\,a_2 \ldots a_n$ is a permutation of $\{1, 2, \ldots, n\}$, and let $b_1\,b_2 \ldots b_n$ be the inverse permutation, so that $a_k = l$ if and only if $b_l = k$. Show that a_k is a left-to-right maximum of $a_1\,a_2 \ldots a_n$ if and only if k is a right-to-left minimum of $b_1\,b_2 \ldots b_n$.

▸**20.** [*M22*] Suppose we want to calculate $\max\{|a_1 - b_1|, |a_2 - b_2|, \ldots, |a_n - b_n|\}$ when $b_1 \le b_2 \le \cdots \le b_n$. Show that it is sufficient to calculate $\max\{m_L, m_R\}$, where

$$m_L = \max\{a_k - b_k \mid a_k \text{ is a left-to-right maximum of } a_1,\, a_2 \ldots a_n\},$$
$$m_R = \max\{b_k - a_k \mid a_k \text{ is a right-to-left minimum of } a_1,\, a_2 \ldots a_n\}.$$

[Thus, if the a's are in random order, the number of k's for which a subtraction must be performed is only $O(\log n)$.]

Page 104 new exercise ______ 13 May 1996

21. [*M25*] Suppose X is a random variable whose values are a *mixture* of the probability distributions generated by $g_1(z)$, $g_2(z)$, …, $g_r(z)$; it uses $g_k(z)$ with probability p_k, where $p_1 + p_2 + \cdots + p_r = 1$. What is the generating function for X? Express the mean and variance of X in terms of the means and variances of g_1, g_2, …, g_r.

Page 104 new exercises ______ 16 May 1996

▸**22.** [*HM21*] Let X be the number of "heads" that occur when a random coin is flipped n times, with generating function (18). Use (25) to prove that

$$\Pr\bigl(X \ge n(p+\epsilon)\bigr) \le e^{-\epsilon^2 n/(2q)}$$

when $\epsilon \ge 0$, and obtain a similar estimate for $\Pr\bigl(X \le n(p-\epsilon)\bigr)$.

▸**23.** [*HM22*] Suppose X has the generating function $(q_1 + p_1 z)(q_2 + p_2 z)\ldots(q_n + p_n z)$, where $p_k + q_k = 1$ for $1 \le k \le n$. Let $\mu = \mathrm{E}\,X = p_1 + p_2 + \cdots + p_n$. (a) Prove that

$$\Pr(X \le \mu r) \le (r^{-r}e^{r-1})^{\mu}, \text{ when } 0 < r \le 1;$$
$$\Pr(X \ge \mu r) \le (r^{-r}e^{r-1})^{\mu}, \text{ when } r \ge 1.$$

(b) Express the right-hand sides of these estimates in convenient form when $r \approx 1$. (c) Show that if r is sufficiently large we have $\Pr(X \ge \mu r) \le 2^{-\mu r}$.

24. [*HM23*] Estimate the tail probabilities for a random variable that has the *negative binomial distribution* generated by $(q - pz)^{-n}$, where $q = p + 1$.

Page 104 line −15 ______ 27 Mar 1995

1892 ⇝ 1894

Page 107 new paragraph to insert after line 2 ______ 28 Mar 1995

Sometimes an asymptotic series is correct although it does not correspond to a generating function, because the infinite series would not be convergent. For example, the basic formulas that express factorial powers in terms of ordinary powers,

$$n^{\overline{r}} = \sum_{k=0}^{m} {r \brack r-k} n^{r-k} + O(n^{r-m-1}), \tag{14}$$

$$n^{\underline{r}} = \sum_{k=0}^{m} (-1)^k {r \brack r-k} n^{r-k} + O(n^{r-m-1}), \tag{15}$$

are asymptotically valid for any real r and any fixed integer $m \ge 0$, yet the sum $\sum_{k\ge 0} {1/2 \brack 1/2-k} n^{1/2-k}$ diverges for all n. [See D. E. Knuth, *AMM* **99** (1992),

415–419.] Of course, when r is a nonnegative integer, $n^{\overline{r}}$ and $n^{\underline{r}}$ are simply polynomials of degree r, and (15) is essentially the same as 1.2.6–(40). When r is a negative integer and $|n| > |r|$, the infinite sum $\sum_{k\ge 0}\left[{r \atop r-k}\right] n^{r-k}$ converges to $n^{\overline{r}} = 1/(n-1)^{\underline{-r}}$. This sum can also be written in the form $\sum_{k\ge 0}\left\{{k-r \atop -r}\right\} n^{r-k}$.
[Now change the number of existing equation (14) to (16).]

Page 107 middle of page . . . . 11 Feb 1995

I'll introduce $\Omega\big(f(n)\big)$ and $\Theta\big(f(n)\big)$ notations here — "order at least" and "order exactly," to go with "order at most" — and put them on page 611 as well.

Page 108 exercise 11 ——— 17 Jul 1996

Eq. (14) $\leadsto$ Eq. (16)

Page 108 new exercise ——— 28 Mar 1995

12. [*HM30*] (B. F. Logan.) By adding together appropriate multiples of (15) we can show that $n^r = \sum_{k=0}^{m}\left\{{r \atop r-k}\right\} n^{\underline{r-k}} + O(n^{r-m-1})$ for any integer $m \ge 0$. Show that, in fact,

$$z^r = \sum_{k\ge 0}\left\{{r \atop r-k}\right\} z^{\underline{r-k}}$$

for arbitrary r and for any complex number $z = x + iy$ with $x > 0$.

▸**Page 108** line 7 ——— 26 Sep 1980

Academæ $\leadsto$ *Academiæ Scientiarum*

Page 109 line 4 ——— 13 Apr 1995

James Bernoulli $\leadsto$ Jacques Bernoulli in his *Ars Conjectandi*, published posthumously

▸**Page 110** the first two lines after (13) ——— 25 Oct 1979

$f^{(2k+1)}(z)$ tends ... from 1 to n $\leadsto$ $f^{(2k+2)}(x)\, f^{(2k+4)}(x) > 0$ for $1 < x < n$.

Page 111 replacement for line 3 ——— 30 Jun 1995

$$= \ln n + \gamma + \sum_{1\le k<m} \frac{(-1)^{k-1}B_k}{kn^k} + O\left(\frac{1}{n^m}\right).$$

Replacing m by $m+1$ yields

$$H_{n-1} = \ln n + \gamma + \sum_{1\le k\le m} \frac{(-1)^{k-1}B_k}{kn^k} + O\left(\frac{1}{n^{m+1}}\right). \qquad (16)$$

▸**Page 111** line 8 ——— 15 Nov 1978

$2\big(k!/(2\pi)^k\big) \leadsto (-1)^{1+k/2}2\big(k!/(2\pi)^k\big)$

Page 112 new wording for exercise 3 ——— 25 Oct 1979

3. [*HM20*] Let $C_m = \big((-1)^m B_m/m!\big)\big(f^{(m-1)}(n) - f^{(m-1)}(1)\big)$ be the mth correction term in Euler's summation formula. If $f^{(2k)}(x)$ has a constant sign for $1 \le x \le n$, show that $|R_{2k}| \le |C_{2k}|$ when $k > 0$; in other words, the remainder is not larger in absolute value than the last term computed.

Page 112 new exercise ———— 28 Mar 1995

▸**8.** [*M25*] Find the asymptotic value of $\binom{2n}{n}$ with a relative error of $O(1/n^3)$, in two ways: (a) via Stirling's approximation; (b) via exercise 1.2.6–47 and Eq. 1.2.11.1–(14).

▸**Page 116** lines −6 and −4 ———— 29 Nov 1977

Analysis $\leadsto$ A crude analysis
$O(n^{n-1/2}e^{-n}) \leadsto O(n^n e^{-n})$ (twice)

▸**Page 117** line 5 ———— 29 Nov 1977

three $\leadsto$ two

▸**Page 117** line −7 ———— 27 Mar 1995

$n^{-k/2} \leadsto x^{-k/2}$

Page 117 line −5 ———— 19 Aug 1996

The derivation $\leadsto$ An even simpler approach to the asymptotics of $Q(n)$, perhaps the simplest possible, appears in exercise 20. The derivation

Page 118 lines 2–5 ———— 08 Jul 1995

["Asymptotische . . . given above use $\leadsto$ [*Math. Zeitschrift* **53** (1950), 136–148]. See also N. M. Temme, *Math. Comp.* **29** (1975), 1109–1114; *SIAM J. Math. Anal.* **10** (1979), 757–766.

Our derivations of the asymptotic series for $P(n)$, $Q(n)$, and $R(n)$ use

Page 118 lines 11 and 12 ———— 16 Oct 1985

North Holland $\leadsto$ North-Holland

▸**Page 118** exercise 5 ———— 29 Nov 1977

$O(n^{n-1/2}e^{-n}) \leadsto O(n^n e^{-n})$

Page 118 replacement for line 1 of exercise 8 ———— 12 Dec 1985

8. [*HM30*] If $f(x) = O(x^r)$ as $x \to \infty$ and $0 \le r < 1$, show that

Page 119 new rating for exercise 13 ———— 30 Jun 1995

[*HM46*] $\leadsto$ [*HM42*]

Page 119 new exercise ———— 16 Mar 1981

18. [*M25*] Show that the sums $\sum \binom{n}{k} k^k (n-k)^{n-k}$ and $\sum \binom{n}{k} (k+1)^k (n-k)^{n-k}$ can be expressed very simply in terms of the Q function.

Page 119 new exercises ———— 19 Aug 1996

19. [*HM30*] (*Watson's lemma.*) Show that if the integral $C_n = \int_0^\infty e^{-nx} f(x)\,dx$ exists for all large n, and if $f(x) = O(x^\alpha)$ for $0 \le x \le r$, where $r > 0$ and $\alpha > -1$, then $C_n = O(n^{-1-\alpha})$.

▸**20.** [*HM30*] Let $w = z + \frac{1}{3}z^2 + \frac{1}{36}z^3 - \frac{1}{270}z^4 + \cdots = \sum_{k=1}^\infty c_k z^k$ be the power series solution to the equation $z = (w^2 - \frac{2}{3}w^3 + \frac{2}{4}w^4 - \frac{2}{5}w^5 + \cdots)^{1/2}$. Show that

$$Q(n) + 1 = \sum_{k=1}^{m-1} k c_k \Gamma(k/2) \left(\frac{n}{2}\right)^{1-k/2} + O(n^{1-m/2})$$

for all $m \ge 1$. [*Hint:* Apply Watson's lemma to the identity of exercise 15.]

Page 120 major changes coming . . . . 27 Mar 1995

The `MIX` computer will eventually be replaced by a 64-bit RISC machine called `MMIX`, which will be considerably nicer and more relevant to computers of the new era. The third edition will be `MIX`'s last hurrah.

▸**Page 131** line −3 ——— 30 Aug 1996

`SRC` ⇝ `SRC 4`
(I believe this error occurs only in the "Indian Student Edition")

Page 133 lines 5, 18, 20 ——— 14 May 1976

record ⇝ block
records ⇝ blocks

Page 134 line −12 ——— 19 Apr 1977

`MUL` requires ⇝ `MUL`, `NUM`, and `CHAR` each require

Page 136 and 137 ——— 01 Jan 1980

Replace Table 1 by new table from endpapers of Volume 2, but retain the caption '**Table 1**'. The changes to the table ars: Character code 10 is `Δ`; character code 20 is `Σ`; timing for opcode 05 is *10*, for opcode 07 is $1 + 2F$, and for opcode 35 is $1 + T$; rAX = registers A and X as one; rIi = index register i; `JGE` and `JLE` stand respectvely for $\geq$ and $\leq$.

Page 150 replacement for line 15 ——— 18 Dec 1996

rule 13), or
d) a literal constant (denoting a reference to an internally created symbol, see rule 12).

Page 150 line 22 ——— 01 Jun 1981

context ⇝ `OP` field, as shown in Table 1.3.1–1

▸**Page 152** line 16 ——— 29 Nov 1977

facilate ⇝ facilitate

Page 159 exercise 19 . . . . 17 Feb 1995

in the next edition, the crossword puzzle program will generate PostScript output, using the default Courier font (I will explain the small subset of PostScript needed for this)

Page 161 lines 2 and 3 ——— 23 Feb 1997

and it is . . . reason. ⇝ and we shall denote it by "()".

▸**Page 163** line 4 of the table ——— 21 Feb 1997

[After step A3, the ↑ should be one position further to the right, i.e., between c and f, not between a and c.]

Page 165 replacement for program lines 57–58 ——— 23 Feb 1997

```
57  *
58  LPREN ALF      (       Constants used in the program
```

Page 172 replacement for lines -12 through -10 ——— 26 Sep 1996

Inverses. The inverse π^- of a permutation π is the rearrangement that undoes the effect of π; if i goes to j under π, then j goes to i under π^-. Thus the product $\pi\pi^-$ equals the identity permutation, and so does the product $\pi^-\pi$. People often denote the inverse by π^{-1} instead of π^-, but the superscript 1 is redundant (for the same reason that $x^1 = x$).

Page 173 new material for this page and the following one ——— 11 Nov 1980

Here is a new Algorithm I together with a new Program I:

Algorithm I (*Inverse in place*). Replace $X[1]X[2]\ldots X[n]$, a permutation on $\{1,2,\ldots,n\}$, by its inverse. This algorithm is due to Bing-Chao Huang [*Inf. Proc. Letters* **12** (1981), 237–238].

I1. [Initialize.] Set $m \leftarrow n$, $j \leftarrow -1$.

I2. [Next element.] Set $i \leftarrow X[m]$. If $i < 0$, go to step I5 (the element has already been processed).

I3. [Invert one.] (At this point $j < 0$ and $i = X[m]$. If m is not the largest element of its cycle, the original permutation had $X[-j] = m$.) Set $X[m] \leftarrow j$, $j \leftarrow -m$, $m \leftarrow i$, $i \leftarrow X[m]$.

I4. [End of cycle?] If $i > 0$, go back to I3 (the cycle has not ended); otherwise set $i \leftarrow j$. (In the latter case, the original permutation had $X[-j] = m$, and m is largest in its cycle.)

I5. [Store final value.] Set $X[m] \leftarrow -i$. (Originally $X[i]$ was equal to m.)

I6. [Loop on m.] Decrease m by 1. If $m > 0$, go back to I2; otherwise the algorithm terminates. ▌

See Table 2 for an example of this algorithm. The method is based on inversion of successive cycles of the permutation, tagging the inverted elements by making them negative, afterwards restoring the correct sign.

Table 2

COMPUTING THE INVERSE OF 6 2 1 5 4 3 BY ALGORITHM I

(Read the columns from left to right.) At point *, the cycle (1 6 3) has been inverted.

After step:	I2	I3	I3	I3	I5*	I2	I3	I3	I5	I2	I5	I5	I3	I5	I5
$X[1]$	6	6	6	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	3
$X[2]$	2	2	2	2	2	2	2	2	2	2	2	2	−4	2	2
$X[3]$	1	1	−6	−6	−6	−6	−6	−6	−6	−6	−6	6	6	6	6
$X[4]$	5	5	5	5	5	5	5	−5	−5	−5	5	5	5	5	5
$X[5]$	4	4	4	4	4	4	−1	−1	4	4	4	4	4	4	4
$X[6]$	3	−1	−1	−1	1	1	1	1	1	1	1	1	1	1	1
m	6	3	1	6	6	5	4	5	5	4	4	3	2	2	1
j	−1	−6	−3	−1	−1	−1	−5	−4	−4	−4	−4	−4	−2	−2	−2
i	3	1	6	−1	−1	4	5	−1	−4	−5	−5	−6	−4	−2	−3

Algorithm I resembles parts of Algorithm A, and it very strongly resembles the cycle-finding algorithm in Program B (lines 50–64). Thus it is typical of a number of algorithms involving rearrangements. When preparing a `MIX` implementation, we find that it is most convenient to keep the value of $-i$ in a register instead of i itself:

Program I (*Inverse in place*). rI1 $\equiv m$; rI2 $\equiv -i$; rI3 $\equiv j$; and $n =$ `N`, a symbol to be defined when this program is assembled as part of a larger routine.

```
INVERT ENT1 N     1  I1. Initialize. m ← n.
       ENT3 -1    1  j ← −1.
2H     LD2N X,1   N  I2. Next element. i ← X[m].
       J2P  5F    N  To I5 if i < 0.
3H     ST3  X,1   N  I3. Invert one. X[m] ← j.
       ENN3 0,1   N  j ← −m.
       ENN1 0,2   N  m ← i.
       LD2N X,1   N  i ← X[m].
4H     J2N  3B    N  I4. End of cycle? To I3 if i > 0.
       ENN2 0,3   C  Otherwise set i ← j.
5H     ST2  X,1   N  I5. Store final value. X[m] ← −i.
6H     DEC1 1     N  I6. Loop on m.
       J1P  2B    N  To I2 if m > 0. ▌
```

The timing for this program is easily worked out in the manner shown earlier; every element $X[m]$ is set first to a negative value in step I3 and later to a positive value in step I5. The total time comes to $(14N + C + 2)u$, where N is the size of the array and C is the total number of cycles. The behavior of C in a random permutation is analyzed below.

There is almost always more than one algorithm ...

Page 175 following the program ______ 22 Feb 1995

This program ... how fast it $\rightsquigarrow$ To decide how fast this program
we must reluctantly ... show that $\rightsquigarrow$ analysis shows that
it turns out ... personal taste. $\rightsquigarrow$ the running time of Algorithm J turns out to be essentially proportional to $n \ln n$, while that of Algorithm I is essentially proportional to n.

Page 177 line 17 ______ 11 Nov 1980

A, B, and I, $\rightsquigarrow$ A and B,

Page 180 line 8 ______ 14 May 1976

itself " $\rightsquigarrow$ itself."

Page 181 new exercise ______ 30 Jun 1995

32. [*M25*] (a) Prove that any permutation $\pi = \pi_1\pi_2\ldots\pi_{2m+1}$ of the form

$$\pi = (2\,3)^{e_2}(4\,5)^{e_4}\ldots(2m\;2m{+}1)^{e_{2m}}(1\,2)^{e_1}(3\,4)^{e_3}\ldots(2m{-}1\;2m)^{e_{2m-1}},$$

where each e_k is 0 or 1, has $|\pi_k - k| \le 2$ for $1 \le k \le 2m+1$.

(b) Given any permutation ρ of $\{1, 2, \ldots, n\}$, construct a permutation π of the stated form such that $\rho\pi$ is a single cycle. Thus every permutation is "near" a cycle.

Page 181 new exercise ——— 26 Sep 1996

33. [*M33*] If $m = 2^{2^l}$ and $n = 2^{2l+1}$, show how to construct sequences of permutations $(\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jn}; \beta_{j1}, \beta_{j2}, \ldots, \beta_{jn})$ for $0 \le j < m$ with the following "orthogonality" property:

$$\alpha_{i1}\beta_{j1}\alpha_{i2}\beta_{j2}\ldots\alpha_{in}\beta_{jn} = \begin{cases} (1\,2\,3\,4\,5), & \text{if } i = j; \\ (1), & \text{if } i \ne j. \end{cases}$$

Each α_{jk} and β_{jk} should be a permutation of $\{1, 2, 3, 4, 5\}$.

Page 181 new exercises ——— 22 Feb 1997

▸**34.** [*M25*] (*Transposing blocks of data.*) One of the most common permutations needed in practice is the change from $\alpha\beta$ to $\beta\alpha$, where α and β are substrings of an array. In other words, if $x_0x_1\ldots x_{m-1} = \alpha$ and $x_mx_{m+1}\ldots x_{m+n-1} = \beta$, we want to change the array $x_0x_1\ldots x_{m+n-1} = \alpha\beta$ to the array $x_mx_{m+1}\ldots x_{m+n-1}x_0x_1\ldots x_{m-1} = \beta\alpha$; this is the permutation on $\{0, 1, \ldots, m+n-1\}$ that takes k into $(k+m) \bmod (m+n)$. Show that this "cyclic-shift" permutation has a simple cycle structure, and exploit that structure to devise a simple algorithm for the desired rearrangement.

35. [*M30*] Continuing the previous exercise, let $x_0x_1\ldots x_{l+m+n-1} = \alpha\beta\gamma$ where α, β, and γ are strings of respective lengths l, m, and n, and suppose that we want to change $\alpha\beta\gamma$ to $\gamma\beta\alpha$. Show that the corresponding permutation has a simple cycle structure that leads to an efficient algorithm. [Exercise 34 considered the special case $m = 0$.] *Hint:* Consider changing $(\alpha\beta)(\gamma\beta)$ to $(\gamma\beta)(\alpha\beta)$.

36. [*27*] Write a `MIX` subroutine for the algorithm in the answer to exercise 35, and analyze its running time. Compare it with the simpler method that goes from $\alpha\beta\gamma$ to $(\alpha\beta\gamma)^R = \gamma^R\beta^R\alpha^R$ to $\gamma\beta\alpha$, where θ^R denotes the left-right reversal of the string θ.

Page 189 insert quotation before the exercises ——— 13 Jul 1995

Up to a point it is better to let the snags [bugs] be there
than to spend such time in design that there are none
(how many decades would this course take?).
— A. M. TURING, Proposals for ACE (1945)

Page 191 clarifications ——— 03 Dec 1983

line −16: each (where ... characters). $\leadsto$ each, until a period appears; the last group may have fewer than three characters.

line −12: Our program is to $\leadsto$ If the input sequence is '`1.`', the output is simply '`.`', not '`..`', because the first period terminates the output. Our program should

line −11: each card. $\leadsto$ each card except possibly the last.

Page 197 lines −12 and −11 ——— 30 Jun 1995

book, ... Chapter 10, and the `XMIX` interpreter in Chapter 12. $\leadsto$
series of books, e.g., the recursive interpreter in Chapter 8 and the "Parsing Machine" in Chapter 10.

Page 198 lines 19–21 ——— 14 Jun 1977

An illustration ... See also the book $\leadsto$ See, for example, the book

▸**Page 209** program line *21* ________ 04 Apr 1980

`LDA` ⇝ `ENTA`

Page 211 following line 17 . . . . 20 Sep 1996

Once again the reader is asked to be indulgent about the anachronistic `MIX` *computer with its punched cards, etc. Although such old-fashioned devices are now quite obsolete, they still can teach important lessons. The* `MMIX` *computer, when it comes, will of course teach those lessons even better.*

Page 220 line −16 ________ 20 Sep 1996

times. Larger values of N are therefore ⇝ times (except when the device has special timing constraints as in exercise 19). Larger values of N are therefore primarily

Page 224 line −11 ________ 14 Jun 1977

F = 7 ⇝ F = 9

Page 225 replacement for exercise 19 ________ 20 Sep 1996

▸**19.** [*M28*] Special considerations arise when input or output involves short blocks on a rotating device like a magnetic disk. Suppose a program works with $n \ge 2$ consecutive blocks of information in the following way: Block k begins to be input at time t_k, where $t_1 = 0$. It is assigned for processing at time $u_k \ge t_k + T$ and released from its buffer at time $v_k = u_k + C$. The disk rotates once every P units of time, and its reading head passes the start of a new block every L units; so we must have $t_k \equiv (k-1)L$ (modulo P). Since the processing is sequential, we must also have $u_k \ge v_{k-1}$ for $1 < k \le n$. There are N buffers, hence $t_k \ge v_{k-N}$ for $N < k \le n$.

How large does N have to be so that the finishing time v_n has its minimum possible value, $T + C + (n-1)\max(L, C)$? Give a general rule for determining the smallest such N. Illustrate your rule when $L = 1$, $P = 100$, $T = .5$, $n = 100$, and (a) $C = .5$; (b) $C = 1.0$; (c) $C = 1.01$; (d) $C = 1.5$; (e) $C = 2.0$; (f) $C = 2.5$; (g) $C = 10.0$; (h) $C = 50.0$; (i) $C = 200.0$.

Page 225 line −17 ________ 29 Mar 1995

Harvard Mark I calculator ⇝ Harvard Mark I calculator [cf. *Mechanization of Thought Processes* (London: Nat. Phys. Lab., 1959), 164]

Page 225 line −9 ________ 25 Jun 1976

about 1946 ⇝ during 1946 and 1947

Page 225 lines −6 through −3 ________ 29 Mar 1995

designed computer hardware to facilitate subroutine linkage ⇝ had designed hardware and software for subroutine linkage as early as 1945

89–90. ⇝ 87–90; *Comp. J.* **20** (1977), 271–272.

Page 235 line 3 ________ 14 Feb 1997

before ⇝ before or after

Page 239 lines −3 and −2 ______ 14 May 1976

Is there ... which decides $\rightsquigarrow$ [See Rosenstiehl and Tarjan, *J. Algorithms* **5** (1984), 389–390, for an algorithm that decides

Page 239 new exercise ______ 30 Jun 1995

14. [*21*] Suppose you are allowed to use only stacks as data structures. How can you implement a queue efficiently with two stacks?

▸**Page 244** line 3 ______ 14 Mar 1976

$\mathtt{L} < \mathtt{TOP}[i] \rightsquigarrow \mathtt{L} \le \mathtt{TOP}[i]$

Page 246 improved overlap ______ 04 Feb 1979

replacement for display on line −10: $\mathtt{OLDTOP[}j\mathtt{]} \equiv \mathtt{D[}j\mathtt{]} \equiv \mathtt{NEWBASE[}j+1\mathtt{]}$
line −9: $n+1$, $\rightsquigarrow$ n;
lines −8 and −7: It will ... overlap. $\rightsquigarrow$

Page 248 new paragraph after line 4 ______ 07 Feb 1979

Further study of Algorithm G has been made by D. S. Wise and D. C. Watson, *BIT* **16** (1976), 442–450. See also A. S. Fraenkel, *Inf. Proc. Letters* **8** (1979), 9–10, who suggests working with pairs of stacks that grow towards each other.

▸**Page 248** bottom line ______ 01 Jan 1976

$\mathtt{BASE}, \mathtt{BASE}+1, \mathtt{BASE}+2 \rightsquigarrow \mathtt{BASE}+1, \mathtt{BASE}+2, \mathtt{BASE}+3$

Page 250 rating for exercise 13 ______ 01 Mar 1979

[*M47*] $\rightsquigarrow$ [*HM42*]

Page 252 lines −12 and −11 ______ 18 Aug 1980

together or to break one apart. $\rightsquigarrow$ together, or to break one apart into two that will grow independently.

Page 254 line −4 ______ 30 Jun 1995

A stack is a simplest kind of linked list. $\rightsquigarrow$ The simplest kind of linked list is a stack.

Page 258 line 4 ______ 21 Sep 1976

we $\rightsquigarrow$ exercise 30 describes a somewhat more natural alternative, and we

▸**Page 264** line *18* of the program ______ 14 Mar 1976

step T3 $\rightsquigarrow$ step T4

Page 265 new paragraph before the exercises ______ 14 May 1976

In spite of the fact that Algorithm T is so efficient, we will study an even better algorithm for topological sorting in Section 7.4.1.

Page 265 line −5 ______ 30 Aug 1996

he mentioned $\rightsquigarrow$ he proved it for infinite sets as well as finite sets, and mentioned

Page 268 lines 17–18 ——— 11 Feb 1995

if possible ... data tape. $\leadsto$ exercise 24 at the beginning of Chapter 5 explains how to handle this case with $O(\log n)^2$ passes over the data.

Page 270 new exercise ——— 21 Sep 1976

30. [*17*] Suppose that queues are represented as in (12), but with an empty queue represented by $F = \Lambda$ and R *undefined*. What insertion and deletion procedures should replace (14) and (17)?

▸**Page 272** line 1 after (4), in certain printings only ——— 25 Oct 1984

list head, w often $\leadsto$ list head, which is often

▸**Page 275** corrections to Program A ——— 14 May 1976

line *04*: `6H` $\leadsto$ `1H`
line *17*: `6B` $\leadsto$ `1B`
replace lines *05–07* by

```
05       LD2  1,3(LINK)   1+m''   Q ← LINK(Q1).
06  0H   LD1  1,1(LINK)   1+p     P ← LINK(P).
07  SW1  LDA  1,1         1+p     rA(0:3) ← ABC(P).
```

now renumber lines *13–35* as lines *14–36*, and replace line *12* by

```
12       LD2  1,3(LINK)   q'      Q ← LINK(Q1).
13       JMP  2B          q'      Repeat.
```

▸**Page 276** line 12 ——— 14 May 1976

$29m' + 18m'' + 29p' \leadsto 27m' + 18m'' + 27p'$

▸**Page 276** exercise 10 ——— 01 Jan 1976

never exceed b $\leadsto$ never exceed $b - 1$
up to b^3. $\leadsto$ up to $b^3 - 1$.

▸**Page 284** Table 1, left column ——— 14 May 1976

the line for time 0200 is out of place; it belongs just before the line for time 0256
in the line for time 0693, M1 $\leadsto$ M5

▸**Page 286** Fig. 12 ——— 14 May 1976

the shading in this figure mysteriously disappeared from the third column of nodes, in the second edition; the first edition was OK!

▸**Page 293** replacement for line 7 ——— 10 Mar 1997

```
POOLMAX  CON  POOL
POOL     END  BEGIN   Storage pool follows literals, temp storage ▌
```

Page 302 new paragraph before the exercises ——— 29 Mar 1995

The representation of sparse matrices via orthogonal circular lists is instructive, but numerical analysts have developed better methods. See Fred G. Gustavson, *ACM Trans. on Math. Software* **4** (1978), 250–269; see also the graph and network algorithms in Chapter 7.

▸**Page 304** line 3 of exercise 20 ——— 14 May 1976

`A(I,J)` $\leadsto$ `A[I,J]`

Page 304 new exercise ——— 14 May 1976

21. [*20*] Suggest a storage allocation function for $n \times n$ matrices where n is variable. The elements `A[I,J]` for $1 \le$ `I`, `J` $\le n$ should occupy n^2 consecutive locations, regardless of the value of n.

Page 304 new exercise ——— 30 Mar 1995

22. [*M25*] (P. Chowla, 1961.) Find a polynomial $p(i_1, \ldots, i_k)$ that assumes each nonnegative integer value exactly once as the indices $(i_1, \ldots, i_k)$ run through all k-dimensional nonnegative integer vectors, with the additional property that $i_1+\cdots+i_k < j_1 + \cdots + j_k$ implies $p(i_1, \ldots, i_k) < p(j_1, \ldots, j_k)$.

Page 304 new exercise ——— 01 Jul 1995

23. [*22*] An *extendible matrix* is initially 1×1, then it grows from size $m \times n$ either to size $(m+1) \times n$ or to size $m \times (n+1)$ by adding either a new row or a new column. Show that such a matrix can be given a simple allocation function in which the elements `A[I,J]` occupy mn consecutive locations, for $0 \le$ `I` $< m$ and $0 \le$ `J` $< n$; no elements change location when the matrix grows.

Page 304 new exercise ——— 01 Sep 1996

▸**24.** [*25*] (*The sparse array trick.*) Suppose you want to use a large array for random access, although you won't actually be referring to very many of its entries. You want `A`$[k]$ to be zero the first time you access it, yet you don't want to spend the time to set every location to zero. Explain how it is possible to read and write any desired elements `A`$[k]$ reliably, given k, without assuming anything about the actual initial memory contents, by doing only a small fixed number of additional operations per array access.

Page 307 future shock . . . . 01 Jan 1977

"parent-child-sibling" nomenclature will replace "father-son-brother" terminology throughout the next edition, but the corrections in this errata list still use the old conventions

Page 308 the name below Nicholas I ——— 05 Jan 1983

Louise $\leadsto$ Charlotte

Page 315 exercise 20 ——— 30 Mar 1995

b-tree $\leadsto$ 0-2 tree (five times)

Page 315 replacement for exercise 22 ——— 31 Mar 1995

22. [*21*] Find a simple one-to-one correspondence between binary trees with n nodes and dissections of an $(n+2)$-sided convex polygon into n triangles, assuming that sides of the polygon are distinct.

Page 315 new exercise ——— 01 Apr 1995

▸**23.** [*M35*] (Conway and Coxeter.) A *frieze pattern* is an infinite array such as

$$\begin{array}{l}
1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ \ldots\\
\ \ 3\ \ 1\ \ 3\ \ 1\ \ 4\ \ 1\ \ 2\ \ 3\ \ 1\ \ 3\ \ 1\ \ 4\ \ 1\ \ 2\ \ 3\ \ 1\ \ 3\ \ 1\ \ 4\ \ \ldots\\
5\ \ 2\ \ 2\ \ 2\ \ 3\ \ 3\ \ 1\ \ 5\ \ 2\ \ 2\ \ 2\ \ 3\ \ 3\ \ 1\ \ 5\ \ 2\ \ 2\ \ 2\ \ 3\ \ \ldots\\
\ \ 3\ \ 3\ \ 1\ \ 5\ \ 2\ \ 2\ \ 2\ \ 3\ \ 3\ \ 1\ \ 5\ \ 2\ \ 2\ \ 2\ \ 3\ \ 3\ \ 1\ \ 5\ \ 2\ \ \ldots\\
1\ \ 4\ \ 1\ \ 2\ \ 3\ \ 1\ \ 3\ \ 1\ \ 4\ \ 1\ \ 2\ \ 3\ \ 1\ \ 3\ \ 1\ \ 4\ \ 1\ \ 2\ \ 3\ \ \ldots\\
\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ \ldots
\end{array}$$

in which the top and bottom rows consist entirely of 1s, and each diamond of adjacent values $a\,{}^{b}_{c}\,d$ satisfies $ad - bc = 1$. Find a one-to-one correspondence between n-node binary trees and $(n+1)$-rowed frieze patterns of positive integers.

Page 315 new exercise ——— 24 Aug 1996

▸**24.** [*21*] Standard European paper sizes A0, A1, A2, ... are rectangles whose sides are in the ratio $\sqrt{2}$ to 1. Therefore if we cut a sheet of An paper in half, we get two sheets of A$(n+1)$ paper. Use this principle to design a graphic representation of binary trees, and illustrate your idea by drawing the representation of 2.3.1–(1) below.

▸**Page 320** line −16 ——— 29 Jun 1992

`RLINK(P)` $=$ `Q` $\leadsto$ `RLINK(P)` $=$ `Q` $\neq \Lambda$

▸**Page 322** tree illustration near bottom of page ——— 14 May 1976

the equation number ‘(9)’ must be inserted at the right of this diagram

Page 323 trivial improvements to Program S ——— 17 Oct 1979

line *03*: `ENT6` $\leadsto$ `ENT5`
line *03*: `Q` $\leadsto$ `P`
line *04*: `S2` $\leadsto$ `2F`
line *09*: $n+1 \leadsto n$
line *09*: Set $\leadsto$ <u>*S2. Search to left.*</u> Set
line *10*, first column: $\leadsto$ `2H`
line *11*: `*-2` $\leadsto$ `S2`
line 5 on page 324: 8 $\leadsto$ 7

Page 325 line 8 ——— 19 Apr 1977

otherwise. $\leadsto$ otherwise, making the latter node the right child of `NODE(Q)`.

Page 327 line 17 ——— 24 May 1996

Chapter 11. $\leadsto$ Chapter 10.

Page 331 lines 1–2 ——— 23 Feb 1995

You may also ... storage. $\leadsto$ But no other bits in the tree nodes are available for temporary storage.

Page 332 new quote to insert just before Section 2.3.2 ——— 16 Nov 1977

Binary or dichotomous systems, although regulated by a principle,
are among the most artificial arrangements
that have ever been invented.
— WILLIAM SWAINSON (*A Treatise on the Geography and*
Classification of Animals, 1835)

▸**Page 345** line −3 ——— 16 Sep 1996

$u'_{2'}$ ⤳ u'_2
$S(u)$ ⤳ $s(u)$

Page 346 replacement for exercise 18 ——— 01 Jul 1995

▸**18.** [*25*] An oriented tree specified by n links FATHER[j] for $1 \le j \le n$ implicitly defines an ordered tree if the nodes in each family are ordered by their location. Design an efficient algorithm that constructs a doubly linked circular list containing the nodes of this ordered tree in preorder. For example, given

$$\begin{aligned} j &= 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8 \\ \mathtt{FATHER}[j] &= 3\ 8\ 4\ 0\ 4\ 8\ 3\ 4 \end{aligned}$$

your algorithm should produce

$$\begin{aligned} \mathtt{LLINK}[j] &= 3\ 8\ 4\ 6\ 7\ 2\ 1\ 5 \\ \mathtt{rLINK}[j] &= 7\ 6\ 1\ 3\ 8\ 4\ 5\ 2 \end{aligned}$$

and it should also report that the root node is 4.

▸**Page 351** line 4 before (10) ——— 01 Jul 1995

takes (5) into (6) ⤳ takes (3) into (6)

▸**Page 359** line 1 of exercise 2 ——— 20 Apr 1997

46–50 ⤳ 53–57

Page 362 exercise 19 . . . . 12 Apr 1995

This (solved) exercise will make more sense in Volume 4. I'll replace it by one or two new exercises about the "Tamari lattice," using SCOPE *links as in* (5) *but with nodes in symmetric order; cf. the construction of Pallo, Comp. J.* **29** *(1986), 171–175.*

Page 367 line −6 ——— 03 Jul 1995

second law: the ⤳ second law [*Ann. Physik und Chemie* **64** (1845), 497–514]: The

Page 368 new paragraph for bottom of page ——— 28 Jul 1996

Applications of Theorem K to software for measuring the performance of programs in high-level languages are discussed by Thomas Ball and James R. Larus in *ACM Trans. Prog. Languages and Systems* **16** (1994), 1319–1360.

Page 371 line 22 ——— 14 Mar 1995

***2.3.4.2.** ⤳ **2.3.4.2.**

Page 377 line 1 of exercise 18 ——— 03 Jul 1995

(Okada and Onodera, . . . 117.) ⤳

Page 378 lines 1 and 2 of exercise 19 ——— 03 Jul 1995

(C. W. Borchardt, . . . 121.) ⤳ (*The matrix tree theorem.*)

Page 381 new exercise ———— 19 May 1981

27. [*M30*] (*Steady states.*) Let G be a directed graph on vertices $V_1, \ldots, V_n$, whose arcs have been assigned probabilities $p(e)$ as in exercise 26. Instead of having "start" and "stop" vertices, however, assume that G is strongly connected; thus each vertex V_j is a root, and we assume that the probabilities $p(e)$ are positive and satisfy $\sum_{\mathrm{init}(e)=V_j} p(e) = 1$ for all j. A random process of the kind described in exercise 26 is said to have a "steady state" $(x_1, \ldots, x_n)$ if

$$x_j = \sum_{\mathrm{fin}(e)=V_j} p(e)\, x_{\mathrm{init}(e)}, \qquad 1 \le j \le n.$$

Let t_j be the sum, over all oriented subtrees T_j of G that are rooted at V_j, of the products $\prod_{e \in T_j} p(e)$. Prove that $(t_1, \ldots, t_n)$ is a steady state of the random process.

Page 381 new exercise ———— 09 Dec 1991

▸**28.** [*M35*] Consider the $(m+n) \times (m+n)$ determinant illustrated here for $m = 2$ and $n = 3$:

$$\det\begin{pmatrix} a_{10}+a_{11}+a_{12}+a_{13} & 0 & a_{11} & a_{12} & a_{13} \\ 0 & a_{20}+a_{21}+a_{22}+a_{23} & a_{21} & a_{22} & a_{23} \\ b_{11} & b_{12} & b_{10}+b_{11}+b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & b_{20}+b_{21}+b_{22} & 0 \\ b_{31} & b_{32} & 0 & 0 & b_{30}+b_{31}+b_{32} \end{pmatrix}.$$

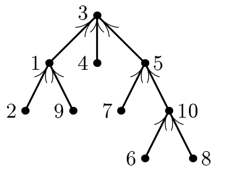
Show that when this determinant is expanded as a polynomial in the a's and b's, each nonzero term has coefficient $+1$. How many terms appear in the expansion? Give a rule, related to oriented trees, that characterizes exactly which terms are present.

Page 382 lines -9, -8, -7 ———— 10 Feb 1995

has recently ... *domino type* $\leadsto$ arises in connection with a family of interesting tiling problems introduced by Hao Wang. A *tetrad type*

(Now change domino $\leadsto$ tetrad throughout the following discussion, including the exercises and their answers.)

Page 390 replacement for (16) and the line following ———— 26 Apr 1995



(16)

with 10 vertices has the canonical representation 1, 3, 10, 5, 10, 1, 3, 5, 3.

Page 395 new paragraph to follow line 4 ———— 11 Jun 1996

The special case where $s = 1$ and $z_1 = 1$ in (25) and (26) is especially important in applications, so it has become known as the "tree function"

$$T(y) = \sum_{n \ge 1} \frac{n^{n-1}}{n!} y^n = y e^{T(y)}. \tag{28}$$

See Corless, Gonnet, Hare, Jeffrey, and Knuth, *Advances in Computational Math.* **5** (1996), 329–359, for a discussion of some of this function's remarkable properties and applications.

Page 395 lines 7 and 8 ______ 09 Feb 1997

This important . . . above. ⇝ More recently, a mathematical *theory of species* developed by André Joyal [*Advances in Math.* **42** (1981), 1–82] has led to a high-level viewpoint in which algebraic operations on generating functions correspond directly to combinatorial properties of structures. The book *Combinatorial Species and Tree-like Structions* by F. Bergeron, G. Labelle, and P. Leroux, (Cambridge Univ. Press, 1997), presents numerous examples of this beautiful and instructive theory, generalizing many of the formulas derived above.

Page 397 line 9 of exercise 18 ______ 26 Apr 1995

10, 3, 3, 10, 10, 9, 9, 8, 8 ⇝ 3, 1, 3, 3, 5, 10, 5, 10, 1

▸**Page 398** exercise 26 ______ 23 Mar 1997

line 2: $(2,3,2)$ ⇝ $(1,4,2)$
line 4: red, yellow, blue, yellow ⇝ yellow, yellow, blue, yellow
line 5: 3, 2; 1, 2, 1; 2, 4. ⇝ 3; 1, 2, 2, 1; 2, 4.

Page 398 exercise 28 ______ 12 Apr 1995

Use ⇝ (*Bipartite trees.*) Use

Page 398 line 3 of exercise 29 ______ 04 Jul 1995

for sufficiently small ⇝ for fixed t and for sufficiently small

▸**Page 398** bottom line ______ 27 May 1978

with ⇝ than

Page 402 three lines before (9) ______ 19 Mar 1981

Huffman: ⇝ Huffman [*Proc. IRE* **40** (1951), 1098–1101]:

Page 404 replacement for lines 1–7 ______ 04 Jul 1995

Every time this construction combines two weights, they are at least as big as the weights previously combined, if the given w_i were nonnegative. This means that there is a neat way to find Huffman's tree, provided that the given weights have been sorted into nondecreasing order: We simply maintain two queues, one containing the original weights and the other containing the combined weights. At each step the smallest unused weight will appear at the front of one of the queues, so we never have to search for it. See exercise 13, which shows that the same idea works even when the weights may be negative.

In general, there are many trees that minimize $\sum w_j l_j$. If the algorithm sketched in the preceding paragraph always uses an original weight instead of a combined weight in case of ties, then the tree it constructs has the smallest value of $\max l_j$ and of $\sum l_j$ among all trees that minimize $\sum w_j l_j$. If the weights are positive, this tree actually minimizes $\sum w_j f(l_j)$ for *any* convex function f, over all such trees. [See E. S. Schwartz, *Information and Control* **7** (1964), 37–44; G. Markowsky, *Acta Informatica* **16** (1981), 363–370.]

Page 405 second line of exercise 10 —— 15 Mar 1981

given weights $\rightsquigarrow$ given nonnegative weights

Page 405 exercise 12 —— 15 Mar 1981

Suppose $\rightsquigarrow$ [*M20*] Suppose

Page 405 new exercises —— 15 Mar 1981

13. [*22*] Design an algorithm that begins with m weights $w_1 \le w_2 \le \cdots \le w_m$ and constructs an extended binary tree having minimum weighted path length. Represent the final tree in three arrays

$$A[1] \ldots A[2m-1], \quad L[1] \ldots L[m-1], \quad R[1] \ldots R[m-1];$$

here $L[i]$ and $R[i]$ point to the left and right children of internal node i, the root is node 1, and $A[i]$ is the weight of node i. The original weights should appear as the external node weights $A[m]$, $\ldots$, $A[2m-1]$. Your algorithm should make fewer than $2m$ weight-comparisons. *Caution:* Some or all of the given weights may be negative!

14. [*25*] (T. C. Hu and A. C. Tucker.) After k steps of Huffman's algorithm, the nodes combined so far form a forest of $m-k$ extended binary trees. Prove that this forest has the smallest total weighted path length, among all forests of $m-k$ extended binary trees that have the given weights.

15. [*M25*] Show that a Huffman-like algorithm will find an extended binary tree that minimizes (a) $\max(w_1 + l_1, \ldots, w_m + l_m)$; (b) $w_1 x^{l_1} + \cdots + w_m x^{l_m}$, given $x > 1$.

16. [*M25*] (F. K. Hwang.) Let $w_1 \le \cdots \le w_m$ and $w'_1 \le \cdots \le w'_m$ be two sets of weights with

$$\sum_{j=1}^{k} w_j \le \sum_{j=1}^{k} w'_j \qquad \text{for } 1 \le k \le m.$$

Prove that the minimum weighted path lengths satisfy $\sum_{j=1}^{m} w_j l_j \le \sum_{j=1}^{m} w'_j l'_j$.

17. [*HM30*] (C. R. Glassey and R. M. Karp.) Let s_1, $\ldots$, s_{m-1} be the numbers inside the internal (circular) nodes of an extended binary tree formed by Huffman's algorithm, in the order of construction. Let s'_1, $\ldots$, s'_{m-1} be the internal node weights of any extended binary tree on the same set of weights $\{w_1, \ldots, w_m\}$, listed in any order such that each nonroot internal node appears before its parent. (a) Prove that $\sum_{j=1}^{k} s_j \le \sum_{j=1}^{k} s'_j$ for $1 \le k < m$. (b) The result of (a) is equivalent to

$$\sum_{j=1}^{m-1} f(s_j) \le \sum_{j=1}^{m-1} f(s'_j)$$

for every nondecreasing concave function f, i.e., every function f with $f'(x) \ge 0$ and $f''(x) \le 0$. [Cf. Hardy, Littlewood, and Pólya, *Messenger of Math.* **58** (1939), 145–152.] Use this fact to show that the minimum value in the recurrence

$$F(n) = f(n) + \min_{1 \le k < n} \bigl(F(k) + F(n-k)\bigr), \qquad F(1) = 0$$

always occurs when $k = 2^{\lceil \lg(n/3) \rceil}$, given any function $f(n)$ with the property that $\Delta f(n) = f(n+1) - f(n) \ge 0$ and $\Delta^2 f(n) = \Delta f(n+1) - \Delta f(n) \le 0$.

Page 405 line 19 ——— 04 Jul 1995

497–508], who $\rightsquigarrow$ 497–508, English translation in *IRE Transactions* **CT5** (1958), 4–7]; Kirchhoff

▸**Page 405** line −18 ——— 18 Aug 1987

(1859) 114–115 $\rightsquigarrow$ (1859), 112–115

▸**Page 406** line 3 ——— 14 Mar 1987

(1760), 13–15, 203–209 $\rightsquigarrow$ (1758–1759), 13–15, 203–210

Page 406 line 6 ——— 28 Jan 1997

"Catalan numbers." $\rightsquigarrow$ "Catalan numbers." A Mongolian Chinese mathematician, An-T'u Ming, had encountered the Catalan numbers before 1750 in his study of infinite series, but he did not relate them to trees or other combinatorial objects [see J. Luo, *Acta Scientiarum Naturalium Universitatis Intramongolicæ* **19** (1988), 239–245; *Combinatorics and Graph Theory* (World Scientific Publishing, 1993), 68–70].

Page 406 line 8 ——— 29 Oct 1996

C. W. Borchardt in 1860 $\rightsquigarrow$ J. J. Sylvester [*Quart. J. Pure and Applied Math.* **1** (1857), 55–56]

▸**Page 406** line 11 ——— 27 May 1978

-types $\rightsquigarrow$ -tuples

▸**Page 406** line −14 ——— 14 May 1976

partic lar $\rightsquigarrow$ particular

Page 416 line 2 ——— 04 Jul 1995

atom; such $\rightsquigarrow$ atom; moreover, steps B4 and C4 need not put items on the stack when they know that the items will immediately be removed. Such

Page 420 new paragraph to follow line 8 ——— 13 Jul 1996

Many programs have the property that the ratio $\rho = \mathtt{N}/\mathtt{M}$ of good nodes to garbage is quite small. When the pool of memory becomes full in such cases, it might be best to move all the active List data to another memory pool of equal size, using a copying technique (see exercise 10) but without bothering to preserve the contents of the nodes being copied. Then when the second memory pool fills up, we can move the data back to the first one again. With this method more data can be kept in high-speed memory at once, because link fields tend to point to nearby nodes. Moreover, there's no need for a marking phase, and storage allocation is simply sequential.

Page 420 lines 14–15 ——— 21 Sep 1976

[See the ... 372.] $\rightsquigarrow$ An elaborate system, which implements this idea and also includes a mechanism for postponing operations on reference counts in order to achieve further efficiency, has been described by L. P. Deutsch and D. G. Bobrow in *CACM* **19** (1976), 522–526.

Page 420 line 17 — 29 Nov 1977

possible; see ⇝ possible. See N. E. Wiseman and J. O. Hiles, *Comp. J.* **10** (1968), 338–343;
506, and ⇝ 506; and

Page 420 new paragraph before the exercises — 07 Feb 1979

Daniel P. Friedman and David S. Wise have observed that the reference counter method can be employed satisfactorily in many cases even when Lists point to themselves, if certain link fields are not included in the counts [*Inf. Proc. Letters* **8** (1979), 41–45].

Page 420 another new paragraph before the exercises — 04 Jul 1995

A great many variants and refinements of garbage collection algorithms have been proposed. Jacques Cohen, in *Computing Surveys* **13** (1981), 341–367, presents a detailed review of the literature prior to 1981, with important comments about the extra cost of memory accesses when pages of data are shuttled between slow memory and fast memory.

Page 435 new paragraph to precede subsection A — 13 Sep 1996

Several authors began about 1975 to call the pool of available memory a "heap." But in the present series of books, we will use that word only in its more traditional sense related to priority queues (see Section 5.2.3).

Page 437 following (1) — 25 Jun 1976

For these reasons the ⇝ (A contrary example appears in exercise 7.) The point is that neither method clearly dominates the other, hence the simple

▸**Page 439** illustration (3) — 04 Jul 1995

The shaded segment is too dark; it should match (2), and both shadings should in fact be lighter.

Page 445 lines 10 and 11 — 16 Jan 1977

on the average, each with a random lifetime, ⇝ *each equally likely to be the next one deleted,*

Page 445 replacement for line −6 through page 446 line 6 — 04 Jul 1995

Let us assume that N is essentially constant, but that A, B, C, and ϵ are random quantities that reach a stationary distribution after a block is freed and a (slightly different) stationary distribution after a block is allocated. The average change in M when a block is freed is the average value of $(C-A)/N$; the average change in M when a block is allocated is $1-p$. So the equilibrium assumption tells us that the average value of $C-A-N+pN$ is zero. But then the average value of $2M$ is pN plus the average value of ϵ, since $2M = N+A-C+\epsilon$ by (15). The fifty-percent rule follows.

Our assumption that each deletion applies to a random reserved block will be valid if the lifetime of a block is an exponentially distributed random variable. On the other hand, if all blocks have roughly the same lifetime, this assumption is false; John E. Shore has pointed out that type A blocks tend to be "older" than type C blocks when allocations and liberations tend to have a somewhat first-in-first-out character, since a sequence of adjacent reserved blocks tends to be in order from youngest to oldest and since the most recently allocated block is almost never type A. This tends to produce a smaller number of available blocks, giving even better performance than the fifty-percent rule would predict. [See *CACM* **20** (1977), 812–820.]

For more detailed information about the fifty-percent rule, see D. J. M. Davies, *BIT 20* (1980), 279–288; C. M. Reeves, *Comp. J.* **26** (1983), 25–35; G. Ch. Pflug, *Comp. J.* **27** (1984), 328–333.

▸**Page 447** lines 4, 6, 7 ——— 14 May 1976

audition $\leadsto$ condition
emergencies hence $\leadsto$ emergencies. Hence

Page 448 line −9 ——— 15 Nov 1978

areas $\leadsto$ areas of the same size

Page 451 line 7 ——— 28 Feb 1978

304. $\leadsto$ 304; J. A. Campbell, *Comp. J.* **14** (1971), 7–9; John E. Shore, *CACM* **18** (1975), 433–440; Norman R. Nielsen, *CACM* **20** (1977), 864–873.

Page 451 new paragraph to follow line 7 ——— 13 Jul 1996

Sometimes it is better to divide the pool of memory into two halves and to do all allocation sequentially within one half. Instead of freeing blocks as they become available, we simply wait until the current active half of memory is full; then we can move all active data to the other half, simultaneously removing all holes between blocks, with a method like that of exercise 33. The size of each half pool might also be adjusted as we switch from one half to the other.

Page 451 line 8 ——— 06 Jul 1995

E. Overflow. $\leadsto$ ***E. Distributed fit.** If the distribution of block sizes is known in advance, and if each block present is equally likely to be the next one freed regardless of when it was allocated, we can use a technique that has substantially better memory utilization than the general-purpose techniques described so far, by following the suggestions of E. G. Coffman, Jr., and F. T. Leighton [*J. Computer and System Sci.* **38**, (1989), 2–35]. Their "distributed fit" method works by partitioning memory into $N + \sqrt{N}\lg N$ slots, where N is the desired maximum number of blocks to be handled in steady state. Each slot has a fixed size, hence it has fixed boundaries, and it will either be empty or contain a single allocated block.

The first N slots in Coffman and Leighton's scheme are laid out according to the assumed distribution of sizes, while the last $\sqrt{N}\lg N$ slots all have the maximum size. For example, if we assume that the block sizes will be uniformly distributed between 1 and 256, and if we expect to handle $N = 2^{14}$ such blocks, we would divide the memory into $N/256 = 2^6$ slots of each size 1, 2, ..., 256, followed by an "overflow area" that contains $\sqrt{N}\lg N = 2^7 \cdot 14 = 1792$ blocks of size 256. When the system is operating at full capacity, we expect it to handle N blocks of average size $\frac{257}{2}$, occupying $\frac{257}{2}N = 2^{21} + 2^{13} = 2{,}105{,}344$ locations; this is the amount of space we have allocated to the first N slots. We have also set aside an additional $1792 \cdot 256 = 458{,}752$ locations to handle the effects of random variations; this additional overhead amounts to $O(N^{-1/2}\log N)$ of the total space, rather than a constant multiple of N as in the buddy system, so it becomes a negligible fraction when $N \to \infty$. In our example, however, it still amounts to about 18% of the total allocation.

The slots should be arranged in order so that the smaller slots precede the larger ones. Given this arrangement, we can allocate blocks by using either the first fit or the best fit technique. (Both methods are equivalent in this case, because the slot sizes are ordered.) The effect, under our assumptions, is to start searching at an essentially random place among the first N slots whenever a new allocation request comes in, and to continue until we find an empty slot.

If the starting slot for each search is truly random between 1 and N, we will not have to invade the overflow area very often. Indeed, if we insert exactly N items starting at random slots, overflow will occur only $O(\sqrt{N}\,)$ times, on the average. The reason is that we can compare this algorithm to hashing with linear probing (Algorithm 6.4L), which has the same behavior except that the search for an empty cell wraps around from N to 1 instead of going into an overflow area. The analysis of Algorithm 6.4L in Theorem 6.4K shows that, when N items have been inserted, the average displacement of each item from its hash address is $\frac{1}{2}(Q(N) - 1) \sim \sqrt{\pi N/8}$; by circular symmetry this average is easily seen to be the same as the average number of times a search goes from slot k to slot $k+1$, for each k. Overflows in the distributed-fit method correspond to searches that go from slot N to slot 1, except that our situation is even better because we avoid some congestion by not wrapping around. Therefore fewer than $\sqrt{\pi N/8}$ overflows will occur, on the average. This analysis does not take account of deletions, which preserve the assumptions of Algorithm 6.4L only if we move blocks back when deleting another block that intervened between their starting slots and their allocated slots (see Algorithm 6.4R); again, however, moving them back would only increase the chance of overflow. Our analysis also fails to account for the effect of having more than N blocks present at once; this will happen if we assume only that the arrival time between blocks is about one Nth of the residence time. For that we need to extend the analysis of Algorithm 6.4L, but Coffman and Leighton proved that the overflow area will almost never need more than $\sqrt{N}\lg N$ slots; the probability of running off the end is less than $O(N^{-M})$ for all M.

In our example, the starting slot for the search during an allocation is not uniform among slots 1, 2, ..., N; it is, instead, uniform among slots 1, 65, ..., $N-63$, because there are $N/256 = 64$ slots of each size. But this deviation from the random model considered in the previous paragraph makes overflow even less likely than predicted. All bets are off, of course, if the assumptions about block size distribution and occupancy time are violated.

F. Overflow.

Page 451 bottom line ________ 20 Mar 1981

36–40. $\leadsto$ 36–40, and in exercises 42–43 where he has shown that the best-fit method has a very bad worst case by comparison with first-fit.

Page 451 new paragraph for bottom of page ________ 22 Jul 1996

G. For further reading. A comprehensive survey and critical review of dynamic storage allocation techniques, based on many more years of experience than were available to the author when the material above was written, has been compiled by Paul R. Wilson, Mark S. Johnstone, Michael Neely, and David Boles, *Lecture Notes in Computer Science* **986** (1995), 1–116.

Page 452 line 1 of exercise 6 ________ 12 Apr 1995

After $\leadsto$ (*Next fit.*) After

Page 454 exercise 28 ________ 19 Apr 1977

line 2: F = 5; for $\leadsto$ F = 5. For
line 4: ." $\leadsto$; the sign of rA is unaffected. Execution time is $2u$."

▸**Page 455** line 3 of exercise 39 ________ 05 Jun 1975

$N(n,m) \leadsto N(n,m)/n$

▸**Page 455** line −7 ________ 19 Dec 1976

divisible by 2^k $\leadsto$ not divisible by 2^k

Page 455 new exercises for bottom of page ________ 01 Apr 1981

42. [*M40*] (J. M. Robson, 1975.) Let $N_{\mathrm{BF}}(n,m)$ be the amount of memory needed to guarantee non-overflow when the best-fit method is used for allocation (cf. exercise 38). Find an attacking strategy to show that $N_{\mathrm{BF}}(n,m) \ge mn - O(n+m^2)$.

43. [*HM35*] Continuing exercise 42, let $N_{\mathrm{FF}}(n,m)$ be the memory needed when the first-fit method is used. Find a defensive strategy to show that $N_{\mathrm{FF}}(n,m) \le H_m n/\ln 2$. (Hence the worst case of first-fit is not far from the best possible worst case.)

Page 455 new exercise ________ 06 Jul 1995

44. [*M21*] Suppose the distribution function $F(x) =$ (probability that a block has size $\le x$) is continuous. For example, $F(x)$ is $(x-a)/(b-a)$ for $a \le x \le b$ if the sizes are uniformly distributed between a and b. Give a formula that expresses the sizes of the first N slots that should be set up when we use the distributed-fit method.

Page 456 line 8 ________ 25 Jun 1976

V-1.] $\leadsto$ V-1; and see especially also the work of Konrad Zuse, *Berichte der Gesellschaft für Math. und Datenv.* **63** (Bonn, 1972), written in 1945. Zuse was the first to develop nontrivial algorithms that worked with lists of dynamically varying lengths.]

Page 458 lines 11 and 12 — 29 Mar 1995

In 1947 ... 1.4.5). $\rightsquigarrow$ In the mid-1940s, A. M. Turing developed a stack mechanism called Reversion Storage for subroutine linkage, local variables, and parameters. His names for "push" and "pop" were "bury" and "disinter/unbury." (See the references in Section 1.4.5.)

Page 458 line 24 — 06 Jul 1995

Weizenbaum]. $\rightsquigarrow$ Weizenbaum]. Ivan Sutherland introduced the use of independent doubly linked lists within larger nodes, in his Sketchpad system (Ph.D. thesis, Mass. Inst. of Technology, 1963).

Page 458 near the bottom — 25 Jun 1976

lines -15 thru -13: The A-1 ... code; $\rightsquigarrow$ The machine language for several early computers used a three-address code to represent the computation of arithmetic expressions;
lines -11 and -10: the A-1 compiler language $\rightsquigarrow$ an extended three-address code

Page 460 lines 3 and 4 — 10 Feb 1995

In recent years ... written. $\rightsquigarrow$

Page 460 lines -9 and -8 — 30 Sep 1996

A. van Wijngaarden ... 218]. $\rightsquigarrow$ A. van Wijngaarden, B. J. Mailloux, J. E. L. Peck, and C. H. A. Koster [*Numerische Math.* **14** (1969), 79–218]; Dennis M. Ritchie [*History of Programming Languages — II* (ACM Press, 1996), 671–698].

Page 463 near the top — 14 Feb 1979

line 1: equality. $\rightsquigarrow$ equality,
line 4: older $\rightsquigarrow$ other
new paragraph to be inserted after line 4:

A related model of computation was proposed by A. N. Kolmogorov as early as 1952. His machine essentially operates on graphs G, having a specially designated starting vertex v_0. The action at each step depends only on the subgraph G' consisting of all vertices at distance $\leq n$ from v_0 in G, replacing G' in G by another graph $G'' = f(G')$, where G'' includes v_0 and the vertices at distance exactly n from v_0, and possibly other vertices (which are newly created); the remainder of graph G is left unaltered. Here n is a fixed number specified in advance for any particular algorithm, but it can be arbitrarily large. A symbol from a finite alphabet is attached to each vertex, and restrictions are made so that no two vertices with the same symbol can be adjacent to a common vertex. (See A. N. Kolmogorov, *Uspekhi Mat. Nauk* **8**, 4 (1953), 175–176; Kolmogorov and Uspenskiĭ, *Uspekhi Mat. Nauk* **13**, 4 (1958), 3–28; *Amer. Math. Soc. Translations*, series 2, **29** (1963), 217–245.) Linking automata can easily simulate graph machines, taking at most a bounded number of steps per graph step. Conversely, however, it is unlikely that graph machines can simulate arbitrary linking automata without unboundedly increasing the running time, unless the definition is changed from undirected to directed graphs, in view of the restriction to vertices of bounded degree. The linking model is, of course, quite close to the operations available to programmers on real machines, while the graph model is not.

Page 465 replacement for lines 6–9 ______ 25 Mar 1995

4. See W. J. LeVeque, *Topics in Number Theory* **2** (Reading, Mass.: Addison-Wesley, 1956), Chapter 3; P. Ribenboim, *13 Lectures on Fermat's Last Theorem* (Springer, 1979); Andrew Wiles, *Annals of Math.* **141** (1995), 443–551.

Page 465 better answer to exercise 1.1–3 ______ 14 May 1976

line 3: let r be $\leadsto$ let m be
line 4: If $r = 0$, $\leadsto$ If $m = 0$,
line 5: **F3.** [Remainder n/m.] Divide n by m and let n be the remainder.
lines 6 and 7 (steps F4 and F5) are deleted
line 8: **F6.** $\leadsto$ **F4.**

▸**Page 466** replacement for lines 10–17 of answer 9 ______ 07 Jul 1995

$\{(m,n,a,b,r,4)\} \cup \{(m,n,a,b,5)\}$. Let $f_2\bigl((m,n)\bigr) = (m,n,m,n,1)$; $f_2\bigl((m,n,d)\bigr) = (m,n,d)$; $f_2\bigl((m,n,a,b,1)\bigr) = (m,n,a,b,a \bmod b,2)$; $f_2\bigl((m,n,a,b,r,2)\bigr) = (m,n,b)$ if $r = 0$, otherwise $(m,n,a,b,r,3)$; $f_2\bigl((m,n,a,b,r,3)\bigr) = (m,n,b,b,r,4)$; $f_2\bigl((m,n,a,b,r,4)\bigr) = (m,n,a,r,5)$; $f_2\bigl((m,n,a,b,5)\bigr) = f_2\bigl((m,n,a,b,1)\bigr)$.

Now let $h\bigl((m,n)\bigr) = g\bigl((m,n)\bigr) = (m,n)$; $h\bigl((m,n,d)\bigr) = (d)$; $h\bigl((m,n,a,b,1)\bigr) = (a,b,0,1)$; $h\bigl((m,n,a,b,r,2)\bigr) = (a,b,r,2)$; $h\bigl((m,n,a,b,r,3)\bigr) = (a,b,r,3)$; $h\bigl((m,n,a,b,r,4)\bigr) = h\bigl(f_2((m,n,a,b,r,4))\bigr)$; $h\bigl((m,n,a,b,5)\bigr) = (a,b,b,1)$; $j\bigl((m,n,a,b,r,3)\bigr) = j\bigl((m,n,a,b,r,4)\bigr) = 2$, otherwise $j(q) = 1$. Then C_2 represents C_1.

Page 466 new paragraph at end of answer 9 ______ 15 Jan 1996

For an alternative approach to simulation, see R. W. Floyd and R. Beigel, *The Language of Machines* (Computer Science Press, 1994), Section 3.3.

Page 467 line −9 ______ 13 Apr 1995

integer. $\leadsto$ integer. The algorithm extended in this way computes the regular continued fraction of the ratio of its inputs; see Section 4.5.3.

Page 468 top two lines ______ 14 May 1976

Add ... *A6.* $\leadsto$ Add "$T \le 3(n-d)+k$" to assertions *A3*, *A4*, *A5*, *A6*, where k takes the respective values 2, 3, 3, 1.

Page 468 better answer to exercise 1.2.2–3 ______ 08 Apr 1975

3. $-1/27$, but the text hasn't defined it.

Page 469 line 2 of answer 11 ______ 06 May 1996

2.00000 .... $\leadsto$ 2.00000 ..., if x's digits agree with the digits of $\log_{10} 2$.

Page 469 new answer ______ 01 Feb 1997

25. Note that $z = 2^{-p}\lfloor 2^{p-k}x \rfloor$, when p is the precision (the number of binary digits after the radix point). The quantity $y + \log_b x$ stays approximately constant.

Page 469 bottom five lines ______ 13 Apr 1995

28. The following solution uses the same auxiliary table as before.

E1. [Initialize.] If $1-\epsilon$ is the largest possible value

E1. [Initialize.] If $1-\epsilon$ is the largest possible value of x, set $y \leftarrow$ (nearest approximation to $b^{1-\epsilon}$), $x \leftarrow 1-\epsilon$, $k \leftarrow 1$. (The quantity yb^{-x} will remain approximately constant in the following steps.)

E2. [Test for end.] If $x = 0$, stop.

E3. [Compare.] If $x < \log_b\bigl(2^k/(2^k-1)\bigr)$, increase k by 1 and repeat this step.

E4. [Reduce values.] Set $x \leftarrow x - \log_b\bigl(2^k/(2^k-1)\bigr)$, $y \leftarrow y - (y$ shifted right $k)$, and go to E2. ▮

Page 470 top ______ 13 Apr 1995

lines 1–3: The computational ... Further $\rightsquigarrow$ If y is set to $b^{1-\epsilon}(1+\epsilon_0)$ in step E1, the subsequent computational error arises when $x \leftarrow x + \log_b(1-2^{-k}) + \delta_j$ and $y \leftarrow y(1-2^{-k})(1+\epsilon_j)$ during the jth execution of step E4, for certain small errors δ_j and ϵ_j. When the algorithm terminates we have computed $y = b^{x-\Sigma\delta_j}\prod(1+\epsilon_j)$. Further

lines 6 and 7: entry ... $-\cdots$. $\rightsquigarrow$ entry $\ln\bigl(2^k/(2^k-1)\bigr)$ can be given with high accuracy: It equals $2^{-k} + \frac{1}{2}2^{-2k} + \frac{1}{3}2^{-3k} + \cdots$.

Page 470 line 10 ______ 18 Apr 1995

388. $\rightsquigarrow$ 388; V. S. Linskiĭ, *Vychisl. Mat.* **2** (1957), 90–119; D. E. Knuth, *METAFONT: The Program* (Reading, Mass.: Addison-Wesley, 1986), §120–§147.

Page 470 answer 3 ______ 08 Apr 1975

the value 3 is ... n^2. $\rightsquigarrow$ $n^2 = 3$ occurs for no n, and in the second place $n^2 = 4$ occurs for *two* n.

▸**Page 470** replacement for answer 10 ______ 27 Apr 1975

9, 10. No; the applications of rule (d) assume that $n \ge 0$. (The result is correct for $n = -1$ but the derivation isn't.)

Page 471 new answer ______ 26 Mar 1995

30. Write the left side as $\sum_{1\le j,k\le n} a_j b_k x_j y_k$, and do a similar thing on the right. (This identity is the special case $m = 2$ of exercise 46.)

Page 472 first two lines ______ 08 Jul 1981

Note: The formulas ... differences." $\rightsquigarrow$ *Notes:* Dr. Matrix was anticipated in this discovery by L. Euler, who wrote to Christian Goldbach about it on 9 November 1762. See Euler's *Institutionum Calculi Integralis* **2** (1769), §1169; and E. Waring, *Phil. Trans.* **69** (1779), 64–67.

Page 472 append to answer 33 ______ 25 Jun 1996

[J. J. Sylvester, *Quart. J. Math.* **1** (1857), 141–152.]

▸**Page 473** answer 44 line 2 ______ 12 Nov 1976

$x_k + y_i \rightsquigarrow x_j + y_k$

Page 474 line 7 ______ 25 Jun 1981

$i+n-1$, and $j+n-1$. $\rightsquigarrow$ $i+n-1$, $j+n-1$, $n-i+1$, and $n-j+1$.

Page 474 append to bottom line ______ 27 Mar 1995

[*J. de l'École Polytechnique* **9** (1813), 280–354; **10** (1815), 29–112. Binet and Cauchy presented their papers on the same day in 1812.]

Page 474 new answer ______ 08 Jun 1995

47. Let $a_{ij} = \bigl(\prod_{k=1}^{j-1}(x_i+p_k)\bigr)\bigl(\prod_{k=j+1}^{n}(x_i+q_k)\bigr)$. Subtract column $k-1$ from column k and factor out $p_{k-j} - q_k$, for $k = n,\ n-1,\ \ldots,\ j+1$ (in that order), for $j = 1,\ 2,\ \ldots,\ n-1$ (in that order). This leaves $\prod_{1\le i<j\le n}(p_i - q_j)$ times $\det(b_{ij})$ where $b_{ij} = \prod_{k=j+1}^{n}(x_i + q_k)$. Now subtract q_{k+j} times column $k+1$ from column k for $k = 1,\ \ldots,\ n-j$, and for $j = 1,\ \ldots,\ n-1$; this leaves $\det(c_{ij})$, where $c_{ij} = x_i^{n-j}$ essentially defines a Vandermonde matrix. We can now proceed as in exercise 37, operating on rows instead of columns, obtaining

$$\det(a_{ij}) = \prod_{1\le i<j\le n}(x_i - x_j)(p_i - q_j)\,.$$

When $p_j = q_j = y_j$ for $1 \le j \le n$, the matrix in this exercise is a Cauchy matrix with row i multiplied by $\prod_{j=1}^{n}(x_i + y_j)$, so this result generalizes exercise 38 by adding $n-1$ independent parameters. [*Manuscripta Math.* **69** (1990), 177–178.]

Page 475 answer 7 ______ 26 Mar 1995

Apply exercise 3 and Eq. (4). $\rightsquigarrow$ $\lfloor x+y\rfloor = \bigl\lfloor \lfloor x\rfloor + x \bmod 1 + \lfloor y\rfloor + y \bmod 1\bigr\rfloor = \lfloor x\rfloor + \lfloor y\rfloor + \lfloor x \bmod 1 + y \bmod 1\rfloor$.

Page 477 replacement for answer 38 ______ 08 Jun 1995

38. Both sides increase by $\lceil y\rceil$ when x increases by 1, so we can assume that $0 \le x < 1$. Then both sides are zero when $x = 0$, and both sides increase by 1 when x increases past the values $1 - k/y$ for $y > k \ge 0$. [*Crelle* **136** (1909), 42; the case $y = n$ is due to C. Hermite, *Acta Math.* **5** (1884), 315.]

Page 478 line 8 ______ 16 Jan 1977

(to appear) $\rightsquigarrow$ **13** (1975), 251–261

▸**Page 478** answer 41 line 4 ______ 04 Nov 1975

$\dfrac{1}{4} \rightsquigarrow \dfrac{1}{8}$ (twice)

Page 478 replacement for answer 41 line 6 ______ 15 Jan 1980

Other correct answers are $\lceil\sqrt{2n} - \frac12\rceil$, $\lceil(\sqrt{8n+1} - 1)/2\rceil$, $\lfloor(\sqrt{8n-7} + 1)/2\rfloor$, etc.

Page 479 new answer ______ 28 Sep 1996

49. Since $f(0) = f\bigl(f(0)\bigr) = f\bigl(f(0) + 0\bigr) = f(0) + f(0)$, we have $f(n) = n$ for all integers n. If $f(\frac12) = k \le 0$, we have $k = f\bigl(\frac{1}{1-2k} f(\frac12 - k)\bigr) = f\bigl(\frac{1}{1-2k}\bigl(f(\frac12) - k\bigr)\bigr) = f(0) = 0$. And if $f(\frac{1}{n-1}) = 0$ we have $f(\frac1n) = f\bigl(\frac1n f(1 + \frac{1}{n-1})\bigr) = f(\frac{1}{n-1}) = 0$; furthermore $1 \le m < n$ implies $f(\frac{m}{n}) = f\bigl(\frac1q f(\frac{qm}{n})\bigr) = f(\frac1q) = 0$, for $q = \lceil n/m\rceil$, by induction on m. Thus $f(\frac12) \le 0$ implies $f(x) = \lfloor x\rfloor$ for all rational x. On the other hand, if $f(\frac12) > 0$ the function $g(x) = -f(-x)$ satisfies (i) and (ii) and has $g(\frac12) = 1 - f(\frac12) \le 0$; hence $f(x) = -g(-x) = -\lfloor -x\rfloor = \lceil x\rceil$ for all rational x. [P. Eisele and K. P. Hadeler, *AMM* **97** (1990), 475–477.]

It does not follow, however, that $f(x) = \lfloor x\rfloor$ or $\lceil x\rceil$ for all *real* values of x. If, for example, $h(x)$ is any function with $h(1) = 1$ and $h(x+y) = h(x) + h(y)$ for all real x and y, then the function $f(x) = \bigl\lfloor h(x)\bigr\rfloor$ satisfies (i) and (ii); but $h(x)$ may be unbounded and highly erratic when $0 < x < 1$ [G. Hamel, *Math. Annalen* **60** (1905), 459–462].

Page 480 answer 18 ——— 01 Jun 1995

$2\Gamma(\frac{3}{2})^2$. $\rightsquigarrow$ $2\Gamma(\frac{3}{2})^2$. [Wallis's own heuristic "proof" can be found in D. J. Struik's *Source Book in Mathematics* (Harvard University Press, 1969), 244–253.]

Page 480 bottom line ——— 03 May 1995

For a table ... 1.2.9. $\rightsquigarrow$

Page 481 corrected kudology ——— 01 May 1995

line 4: argument. $\rightsquigarrow$ argument suggested by H. S. Wall [*Bull. Amer. Math. Soc.* **44** (1938), 395–398]:
line −1: (This solution is by R. McEliece.) $\rightsquigarrow$

Page 482 lines 1 and 2 ——— 03 May 1995

Faa di Bruno's ... formula, $\rightsquigarrow$ Arbogast's original proof [*Du Calcul des Dérivations* (Strasbourg, 1800), §52] was based on the fact that $D_x^k u/k!$ is the coefficient of z^k in $u(x+z)$ and $D_u^j w/j!$ is the coefficient of y^j in $w(u+y)$, hence the coefficient of z^n in $w(u(x+z))$ is

$$\frac{D_x^n w}{n!} = \sum_{0\le j\le n} \frac{D_u^j w}{j!} \sum_{\substack{k_1+k_2+\cdots+k_n=j\\ k_1+2k_2+\cdots+nk_n=n\\ k_1,k_2,\ldots,k_n\ge 0}} \frac{j!}{k_1!\,k_2!\ldots k_n!}\left(\frac{d_x^1 u}{1!}\right)^{k_1}\left(\frac{d_x^2 u}{2!}\right)^{k_2}\ldots\left(\frac{d_x^n u}{n!}\right)^{k_n}.$$

His formula was forgotten for many years, then rediscovered independently by F. Faà di Bruno [*Quarterly J. Math.* **1** (1857), 359–360], who observed that it can also be expressed as a determinant

$$D_x^n = \det\begin{pmatrix} \binom{n-1}{0}u_1 & \binom{n-1}{1}u_2 & \binom{n-1}{2}u_3 & \ldots & \binom{n-1}{n-2}u_{n-1} & \binom{n-1}{n-1}u_n\\ -1 & \binom{n-2}{0}u_1 & \binom{n-2}{1}u_2 & \ldots & \binom{n-2}{n-3}u_{n-2} & \binom{n-2}{n-2}u_{n-1}\\ 0 & -1 & \binom{n-3}{0}u_1 & \ldots & \binom{n-3}{n-4}u_{n-3} & \binom{n-3}{n-3}u_{n-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \ldots & -1 & \binom{0}{0}u_1 \end{pmatrix}$$

where $u_j = D_x^j u\, D_u$; both sides of this equation are differential operators to be applied to w. For a generalization of Arbogast's formula to functions of several variables,

Page 482 new answer ——— 09 Jun 1995

23. $z\,(-z)!\,\Gamma(z) = \lim_{m\to\infty}\prod_{n=1}^{m}(1-z/n)^{-1}(1+z/n)^{-1}$ by (13) and (15).

Page 482 new answer ——— 03 Jun 1995

24. $n^n/n! = \prod_{k=1}^{n-1}(k+1)^k/k^k \le \prod_{k=1}^{n-1} e$; $n!/n^{n+1} = \prod_{k=1}^{n-1} k^{k+1}/(k+1)^{k+1} \le \prod_{k=1}^{n-1} e^{-1}$.

Page 483 last line of answer 11 ——— 26 Mar 1995

formula. $\rightsquigarrow$ formula. [Similar results hold for q-nomial and Fibonomial coefficients; see Knuth and Wilf, *Crelle* **396** (1989), 212–219.]

Page 483 replacement for lines 2 and 3 of answer 17 ——— 01 Nov 1984

$$\sum_n \binom{r+s}{n} x^n = (1+x)^{r+s} = \sum_k \binom{r}{k} x^k \sum_m \binom{s}{m} x^m$$
$$= \sum_k \binom{r}{k} x^k \sum_n \binom{s}{n-k} x^{n-k} = \sum_n \biggl(\sum_k \binom{r}{k} \binom{s}{n-k} \biggr) x^n$$

Page 484 closing lines of answer 25 ——— 04 May 1995

the simple but less elementary . . . 2.3.4.4–33. $\leadsto$ the more general formulas in exercises 2.3.4.4–33 and 4.7–22.

Page 485 answer 31 ——— 14 May 1976

We have $\leadsto$ [This sum was first obtained in closed form by J. F. Pfaff, *Nova acta acad. scient. Petr.* **11** (1797), 38–57.] We have

Page 486 extending to page 487 ——— 14 May 1976

$B \leadsto \mathrm{B}$ (not italics) in the solutions to exercises 40, 41 (twice), 48 (twice)

Page 487 new answer ——— 19 Apr 1977

42. $1/\bigl((r+1)\mathrm{B}(k+1, r-k+1)\bigr)$, if this is defined according to exercise 41(b). In general when z and w are arbitrary complex numbers we define

$$\binom{z}{w} = \lim_{\zeta\to z} \lim_{\omega\to w} \frac{\zeta!}{\omega!\,(\zeta-\omega)!}, \qquad \text{where } \zeta! = \Gamma(\zeta+1);$$

the value is infinite when z is a negative integer and w is not an integer.

With this definition, Eq. (6) holds for all complex n and k, except when n is a negative integer and k is an integer; Eqs. (7), (9), and (20) are never false, although they may occasionally take indeterminate forms such as $0 \cdot \infty$ or $\infty + \infty$. Eq. (17) becomes

$$\binom{z}{w} = \frac{\sin \pi(w-z-1)}{\sin \pi z} \binom{w-z-1}{w}.$$

We can even extend the binomial theorem (13) and Vandermonde's convolution (21), obtaining $\sum_k \binom{r}{\alpha+k} z^{\alpha+k} = (1+z)^r$ and $\sum_k \binom{r}{\alpha+k}\binom{s}{\beta-k} = \binom{r+s}{\alpha+\beta}$; these formulas hold for all complex r, s, z, α, and β whenever the series converge, provided that complex powers are suitably defined. [See L. Ramshaw, *Inf. Proc. Letters* **6** (1977), 223–226.]

▸**Page 488** answer 52 line 1 ——— 10 Jan 1981

$\pi^2/6 - 1 \leadsto \pi^2/6$

Page 488 new answer ——— 10 Jun 1995

53. For (b), set $r = \frac{1}{2}$ and $s = -\frac{1}{2}$ in the result of (a).

Page 488 additional material for answer 56 ——— 26 Mar 1995

We can find this representation by a "greedy" method, first choosing the largest possible d, then the largest possible c for $n - \binom{d}{4}$, etc. [Section 7.2.1 discusses further properties of this important representation.]

▸**Page 488** answer 58 line 3 ——— 20 Oct 1979

$q^{(s-n-k)k} \leadsto q^{(s-n+k)k}$

Page 488 replacement for lines 4–7 of answer 58 ______ 26 Mar 1995

And the identity $1-q^t=-q^t(1-q^{-t})$ makes it easy to generalize (17) to

$$\binom{r}{k}_q=(-1)^k\binom{k-r-1}{k}_q q^{kr-k(k-1)/2}.$$

These coefficients were introduced by C. F. Gauss [*Commentationes societatis regiaæ scientiarum Gottingensis recentiores* **1** (1811); reprinted in his *Werke* **2** (1863), 9–45]. They arise in many diverse applications; see, for example, Sections 5.1.2, 7.2.2, and the author's note in *J. Combinatorial Theory* **A10** (1971), 178–180.

Useful facts: When n is a nonnegative integer, $\binom{n}{k}_q$ is a polynomial of degree $k(n-k)$ in q with nonnegative integer coefficients, and it satisfies the reflective laws

$$\binom{n}{k}_q=\binom{n}{n-k}_q=q^{k(n-k)}\binom{n}{k}_{q^{-1}}.$$

If $|q|<1$ and $|x|<1$, the q-nomial theorem holds when n is an arbitrary real number, if we replace the left-hand side by $\prod_{k\ge0}((1+q^kx)/(1+q^{n+k}x))$; properties of power series make it necessary to verify this only when n is a positive integer, because we can set $q^n=y$ and the identity has been verified for infinitely many values of y. Now we can negate the upper index in the q-nomial theorem, obtaining

$$\prod_{k\ge0}\frac{(1-q^{k+r+1}x)}{(1-q^kx)}=\sum_k\binom{-r-1}{k}_q q^{k(k-1)/2}(-q^{r+1}x)^k=\sum_k\binom{k+r}{k}_q x^k,$$

a formula due to Cauchy [*Comptes Rendus* **17** (Paris, 1843), 523] and Heine [*Crelle* **34** (1847), 285–328]. For further information, see G. Gasper and M. Rahman, *Basic Hypergeometric Series* (Cambridge Univ. Press, 1990).

▸**Page 488** new answer to exercise 59 ______ 30 Aug 1980

59. $(n+1)\binom{n}{k}-\binom{n}{k+1}$.

Page 489 line 8 of answer 62 ______ 07 Jul 1995

See papers $\rightsquigarrow$ However, L. J. Rogers had already published a much more general formula in the meantime [*Proc. London Math. Soc.* **26** (1895), 15–32, §8]. See also papers

Page 489 line 10 of answer 62 ______ 26 Sep 1987

(1906) $\rightsquigarrow$ (1907)

Page 489 new answer ______ 26 Mar 1995

63. See *CMath*, exercises 5.83 and 5.106.

Page 489 new answers ______ 09 Jun 1995

65. Notice first that $\binom{z}{n+1}\le\binom{y}{n+1}$. This is obvious if $z\ge n$, since $z\le y$; otherwise $n-1\le z\le n$, hence $\binom{z}{n+1}\le0\le\binom{y}{n+1}$. Therefore $\binom{z+1}{n+1}=\binom{z}{n+1}+\binom{z}{n}\le\binom{y}{n+1}+\binom{z}{n}=\binom{x}{n+1}$, and we have $x\ge z+1$.

Now we can prove that every term of the sum

$$\binom{x}{n+1}-\binom{y}{n+1}=\sum_{k\ge0}\binom{z-k}{n-k}t_k,\qquad t_k=\binom{x-z-1+k}{k+1}-\binom{y-z-1+k}{k+1},$$

is nonnegative. The coefficient $\binom{z-k}{n-k}$ is nonnegative, since $z \ge n-1$; and so is $\binom{x-z-1+k}{k+1}$, since $x \ge z+1$. Therefore $z \le y \le x$ implies $\binom{y-z-1+k}{k+1} \le \binom{x-z-1+k}{k+1}$.

The desired result is obvious when $x = y$ and $z = n-1$. Otherwise $\binom{x}{n} - \binom{y}{n} - \binom{z}{n-1} = \sum_{k\ge 0}\binom{z-k}{n-1-k}(t_k - \delta_{k0}) = \sum_{k\ge 0}\frac{n-k}{z-n+1}\binom{z-k}{n-k}(t_k - \delta_{k0})$ is less than or equal to $\sum_{k\ge 0}\frac{n-1}{z-n+1}\binom{z-k}{n-k}(t_k - \delta_{k0}) = \frac{n-1}{z-n+1}\bigl(\binom{x}{n+1} - \binom{y}{n+1} - \binom{z}{n}\bigr) = 0$, because $t_0 - 1 = x - y - 1 \le 0$. [L. Lovász, *Combinatorial Problems and Exercises*, Problem 13.31(a).]

66. If $k > 0$, exercise 1.2.5–24 gives the slightly sharper (but less memorable) upper bounds $\binom{n}{k} = n^{\underline{k}}/k! \le n^k/k! \le \frac{1}{e}\bigl(\frac{ne}{k}\bigr)^k \le \bigl(\frac{ne}{k+1}\bigr)^k$. The corresponding lower bound is $\binom{n}{k} \ge \bigl(\frac{(n-k)e}{k}\bigr)^k \frac{1}{ek}$.

Page 489 new answer ——— 08 May 1996

67. Let $t_k = k\binom{n}{k}p^k(1-p)^{n+1-k}$; then $t_k - t_{k+1} = \binom{n}{k}p^k(1-p)^{n-k}(k-np)$. So the stated sum is $\sum_{k<\lceil np\rceil}(t_{k+1}-t_k) + \sum_{k\ge\lceil np\rceil}(t_k - t_{k+1}) = 2t_{\lceil np\rceil}$. [De Moivre stated this identity in *Miscellanea Analytica* (1730), 101, in the case that np is an integer; H. Poincaré proved the general case in his *Calcul des Probabilités* (1896), 56–60. See P. Diaconis and S. Zabell, *Statistical Science* **6** (1991), 284–302, for the interesting history of this identity and for a variety of similar formulas.]

Page 490 replacement for answer 11 ——— 25 Oct 1984

11. $2 - H_n/n - 1/n$, for $n > 0$.

Page 490 new answer ——— 26 Sep 1987

13. Differentiate with respect to x; also evaluate at $x = 1$.

▸**Page 490** answer 14 ——— 14 May 1976

$H_{n+4}^2 \rightsquigarrow H_{n+1}^2$

▸**Page 490** line 2 of answer 23 ——— 24 Mar 1986

$H_n - \gamma \rightsquigarrow H_{n-1} - \gamma$

Page 491 line 3 ——— 19 Dec 1996

of r. $\rightsquigarrow$ of r. The infinite product converges for all complex x.

Page 491 answer 19 ——— 13 Apr 1995

i.e. ... $\frac{1}{2}\phi$. $\rightsquigarrow$ i.e., $1 = 2(v-u) = 2v - 4v^2 + 2$. We conclude that $v = \frac{1}{2}\phi$. (Also $u = \frac{1}{2}\phi^{-1}$, $\sin 36° = \frac{1}{2}5^{1/4}\phi^{-1/2}$, $\sin 72° = \frac{1}{2}5^{1/4}\phi^{1/2}$.)

Page 492 last line of answer 29 ——— 22 May 1996

(6). $\rightsquigarrow$ (6). [E. Lucas, *Amer. J. Math.* **1** (1878), 201–204.]

Page 492 line -2 ——— 13 Apr 1995

See also J. Riordan, *Duke Math J.* $\rightsquigarrow$ *References:* Dov Jarden, *Recurring Sequences*, 2nd ed. (Jerusalem, 1966), 30–33; J. Riordan, *Duke Math. J.*

Page 493 answer 35 ——— 17 Aug 1996

110. $\rightsquigarrow$ 110. To represent $x > 0$, we find the largest k with $\phi^k \le x$ and represent x as ϕ^k plus the representation of $x - \phi^k$.

The representation of nonnegative integers can also be obtained from the following all-integer recursive rules, starting with the trivial representations of 0 and 1: Let

$L_n = \phi^n + \widehat{\phi}^{\,n} = F_{n+1} + F_{n-1}$. The representation of $L_{2n} + m$ for $0 \le m \le L_{2n-1}$ and $n \ge 1$ is $\phi^{2n} + \phi^{-2n}$ plus the representation of m. The representation of $L_{2n+1} + m$ for $0 < m < L_{2n}$ and $n \ge 0$ is $\phi^{2n+1} + \phi^{-2n-2}$ plus the representation of $m - \phi^{2n}$, where the latter is obtained by applying the rule $\phi^k - \phi^{k-2j} = \phi^{k-1} + \cdots + \phi^{k-2j+1}$. It turns out that all strings α of 0's and 1's, such that α begins with 1 and has no adjacent 1's, occur to the left of the radix point in the representation of exactly one positive integer, except for the strings that end with $10^{2k}1$; the latter strings never occur in such representations.

Page 493 additional sentences for answer 36 ——— 13 Apr 1995

Also, the kth letter is b iff $k = \lfloor m\phi \rfloor$ for some positive integer m. This sequence was studied by Jean Bernoulli III in the 18th century, by A. A. Markov in the 19th, and by many other mathematicians since then; see K. B. Stolarsky, *Canadian Math. Bull.* **19** (1976), 473–482.

▶**Page 493** line 3 of answer 37 ——— 16 Oct 1985

$F_{k_r-1} \rightsquigarrow F_{k_{r-1}}$

Page 493 new answers ——— 10 Jun 1995

40. We prove, by induction on m, that $f(n) = m$ for $F_m < n \le F_{m+1}$: First, $f(n) \le \max\bigl(1 + f(F_m), 2 + f(n - F_m)\bigr) = m$. Second, if $f(n) < m$ there is some $k < n$ with $1 + f(k) < m$ (hence $k \le F_{m-1}$) and $2 + f(n-k) < m$ (hence $n - k \le F_{m-2}$); but then $n \le F_{m-1} + F_{m-2}$. [Thus the Fibonacci trees defined in Section 6.2.1 minimize the maximum root-to-leaf cost when a right branch costs twice as much as a left branch.]

41. $F_{k_1+1} + \cdots + F_{k_r+1} = \phi n + (\widehat{\phi}^{\,k_1} + \cdots + \widehat{\phi}^{\,k_r})$ is an integer, and the parenthesized quantity lies between $\widehat{\phi}^{\,3} + \widehat{\phi}^{\,5} + \cdots = \phi^{-1} - 1$ and $\widehat{\phi}^{\,2} + \widehat{\phi}^{\,4} + \cdots = \phi^{-1}$. Similarly, $F_{k_1-1} + \cdots + F_{k_r-1} = \phi^{-1}n + (\widehat{\phi}^{\,k_1} + \cdots + \widehat{\phi}^{\,k_r}) = f(\phi^{-1}n)$. [Such Fibonacci shifting is a convenient way to convert mentally between miles and kilometers; see *CMath*, §6.6.]

42. [*Fibonacci Quarterly* **6** (1968), 235–244.] If such a representation exists, we have

$$mF_{N-1} + nF_N = F_{k_1+N} + F_{k_2+N} + \cdots + F_{k_r+N} \qquad (*)$$

for all integers N; hence two different representations would contradict exercise 34.

Conversely, we can prove the existence of such joint representations for all nonnegative m and n by induction. But it is more interesting to use the previous exercise, and to prove that such joint representations exist for possibly negative integers m and n iff $m + \phi n \ge 0$: Let N be large enough so that $|m\widehat{\phi}^{\,N-1} + n\widehat{\phi}^{\,N}| < \phi^{-2}$, and represent $mF_{N-1} + nF_N$ as in $(*)$. Then $mF_N + nF_{N+1} = \phi(mF_{N-1} + nF_N) + (m\widehat{\phi}^{\,N-1} + n\widehat{\phi}^{\,N}) = f\bigl(\phi(mF_{N-1} + nF_N)\bigr) = F_{k_1+N+1} + \cdots + F_{k_r+N+1}$, and it follows that $(*)$ holds for all N. Now set $N = 0$ and $N = 1$.

Page 494 answer 10 ——— 12 Jul 1995

line 2: (25) $\rightsquigarrow$ (17)

lines 6–7: x_j ... *identities.* $\rightsquigarrow$ x_j; the recurrence analogous to (35) is $na_n = S_1 a_{n-1} - S_2 a_{n-2} + \cdots$. [The equations in this recurrence are called *Newton's identities*, since they were first published in Isaac Newton's *Arithmetica Universalis* (1707); see D. J. Struik's *Source Book in Mathematics* (Harvard University Press, 1969), 94–95.]

Page 494 replacement for answer 11 ——— 12 Jul 1995

11. Since $\sum_{m\ge1} S_m z^m/m = \ln G(z) = \sum_{k\ge1} (-1)^{k-1}(h_1 z + h_2 z^2 + \cdots)^k/k$, the desired coefficient is $(-1)^{k_1+k_2+\cdots+k_m-1} m(k_1 + k_2 + \cdots + k_m - 1)!/k_1!\,k_2!\ldots k_m!$. [Multiply by

$(-1)^{m-1}$ to get the coefficient of $a_1^{k_1} a_2^{k_2} \ldots a_m^{k_m}$ when S_m is expressed in terms of the a's of exercise 10. Albert Girard stated the formulas for S_1, S_2, S_3, S_4 in terms of a_1, a_2, a_3, a_4 near the end of his *Invention Nouvelle en Algébre* (Amsterdam, 1629); this was the birth of the theory of symmetric functions.]

Page 494 answer 15 —— 08 Dec 1975

line 1: $zG_{n-2}(z) \rightsquigarrow zG_{n-2}(z) + \delta_{n0}$
line 3 (the displayed formula): delete the '.' and add a new line 4:
when $z \neq -1/4$; $G_n(-1/4) = (n+1)/2^n$ for $n \geq 0$.

Page 496 new answer following line 4 —— 27 Mar 1995

20. $c_{mk} = k!\left\{ {m \atop k} \right\}$, by Eq. 1.2.6–(41).

Page 496 new answer —— 12 Jul 1995

The former answer 11 becomes answer 21.

Page 496 new answer —— 10 Jun 1995

22. When $m = 1$ this is the binomial theorem, with $f_1(z) = z$ and $g_1(z) = 1 + z$. When $m \geq 1$ we can increase m by 1 if we replace z_m by $z_m(1 + z_{m+1}^{-1})$ and let $f_{m+1}(z_1, \ldots, z_{m+1}) = z_{m+1} f_m\bigl(z_1, \ldots, z_{m-1}, z_m(1 + z_{m+1}^{-1})\bigr)$, $g_{m+1}(z_1, \ldots, z_{m+1}) = z_{m+1} g_m\bigl(z_1, \ldots, z_{m-1}, z_m(1 + z_{m+1}^{-1})\bigr)$. Thus $g_2(z_1, z_2) = z_1 + z_2 + z_1 z_2$ and

$$\frac{g_m(z_1, \ldots, z_m)}{f_m(z_1, \ldots, z_m)} = 1 + \cfrac{z_1^{-1}}{1 + \cfrac{z_2^{-1}}{1 + \cfrac{\ddots}{1 + z_m^{-1}}}}.$$

Both polynomials f_m and g_m satisfy the same recurrence $f_m = z_m f_{m-1} + z_{m-1} f_{m-2}$, $g_m = z_m g_{m-1} + z_{m-1} g_{m-2}$, with the initial conditions $f_{-1} = 0$, $f_0 = g_{-1} = g_0 = z_0 = 1$. It follows that g_m is the sum of all terms obtainable by starting with $z_1 \ldots z_m$ and striking out zero or more nonadjacent factors; there are F_{m+2} ways to do this. A similar interpretation applies to f_m, except that z_1 must remain. In part (b) we will encounter the polynomial $h_m = z_m g_{m-1} + z_{m-1} f_{m-2}$; this is the sum of all terms obtained from $z_1 \ldots z_m$ by striking out factors that are not *cyclically* adjacent. For example, $h_3 = z_1 z_2 z_3 + z_1 z_2 + z_1 z_3 + z_2 z_3$.

(b) By part (a), $S_n(z_1, \ldots, z_{m-1}, z) = [z_m^n] \sum_{r=0}^{n} z^r z_m^{n-r} f_m^{n-r} g_m^r$; hence

$$S_n(z_1, \ldots, z_m) = \sum_{0 \leq s \leq r \leq n} \binom{r}{s} \binom{n-r}{s} a^{r-s} b^s c^s d^{n-r-s},$$

where $a = z_m g_{m-1}$, $b = z_{m-1} g_{m-2}$, $c = z_m f_{m-1}$, $d = z_{m-1} f_{m-2}$. Multiplying this equation by z^n and summing first on n, then on r, then on s, yields the closed form

$$S_n(z_1, \ldots, z_m) = [z^n] \frac{1}{(1 - az)(1 - dz) - bcz^2} = \frac{\rho^{n+1} - \sigma^{n+1}}{\rho - \sigma},$$

where $1 - (a + d)z + (ad - bc)z^2 = (1 - \rho z)(1 - \sigma z)$. Here $a + d = h_m$, and $ad - bc$ simplifies to $(-1)^m z_1 \ldots z_m$. [We have, incidentally, established the recurrence $S_n = h_m S_{n-1} - (-1)^m z_1 \ldots z_m S_{n-2}$, a relation that is not easy to derive without the help of generating functions.]

(c) Let $\rho_1 = (z + \sqrt{z^2 + 4z}\,)/2$ and $\sigma_1 = (z - \sqrt{z^2 + 4z}\,)/2$ be the roots when $m = 1$; then $\rho_m = \rho_1^m$ and $\sigma_m = \sigma_1^m$. Egorychev

Carlitz used this result to deduce a surprising fact: The characteristic polynomial $\det(xI - A)$ of the $n \times n$ matrix

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & \binom{0}{0} \\ 0 & 0 & \ldots & \binom{1}{0} & \binom{1}{1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \binom{n-1}{0} & \binom{n-1}{1} & \ldots & \binom{n-1}{n-2} & \binom{n-1}{n-1} \end{pmatrix}$$

of "right justified binomial coefficients" is $\sum_k \binom{n}{k}_{\mathcal{F}} (-1)^{\lceil (n-k)/2 \rceil} x^k$, with Fibonomial coefficients (see exercise 1.2.8–30). He also showed, using similar methods, that

$$\sum_{k_1,\ldots,k_m \ge 0} \binom{k_1+k_2}{k_1}\binom{k_2+k_3}{k_2}\ldots\binom{k_m+k_1}{k_m} z_1^{k_1}\ldots z_m^{k_m} = \frac{1}{\sqrt{z_1^2\ldots z_m^2\, h_m(-z_1^{-1},\ldots,-z_m^{-1})^2 - 4z_1\ldots z_m}}.$$

[*Collectanea Math.* **27** (1965), 281–296.]

Page 496 new answers ———— 09 Feb 1997

23. $G(z) = (1+z)^r(1+z^2)^r(1+z^4)^r(1+z^8)^r\ldots = (1-z)^{-r}$. It follows that the stated sum is $\binom{r+n-1}{n}$.

24. Both sides are equal to $\sum_k \binom{m}{k}[z^n]\bigl(zG(z)\bigr)^k$. When $G(z) = 1/(1-z)$, the identity becomes $\sum_k \binom{m}{k}\binom{n-1}{n-k} = \binom{m+n-1}{n}$, a case of 1.2.6–(21). When $G(z) = (e^z - 1)/z$, it becomes $\sum_k m^{\underline{k}} {n \brace k} = m^n$, Eq. 1.2.6–(41).

25. $\sum_k [w^k](1-2w)^n\,[z^n]\,z^k(1+z)^{2n-2k} = [z^n]\,(1+z)^{2n}\sum_k [w^k]\,(1-2w)^n\bigl(z/(1+z)^2\bigr)^k$, which equals $[z^n]\,(1+z)^{2n}\bigl(1 - 2z/(1+z)^2\bigr)^n = [z^n]\,(1+z^2)^n = \binom{n}{n/2}[n \text{ even}]$. Similarly, we find $\sum_k \binom{n}{k}\binom{2n-2k}{n-k}(-4)^k = (-1)^n\binom{2n}{n}$. Many examples of this summation technique can be found in G. P. Egorychev's book *Integral Representation and the Computation of Combinatorial Sums* (Amer. Math. Soc., 1984), translated from the Russian edition of 1977.

26. See D. E. Knuth, in *A Classical Mind* (Prentice–Hall, 1994), 247–258.

Page 496 new answer ———— 25 Jun 1996

6. For (18), the formula

$$\ln(q + pe^t) = \ln(1 + pt + p\frac{t^2}{2} + p\frac{t^3}{6} + \cdots = pt + p(1-p)\frac{t^2}{2} + p(1-p)(1-2p)\frac{t^3}{6} + \cdots$$

tells us that $\kappa_3 = p(1-p)(1-2p) = pq(q-p)$. (This nice pattern does not continue to the coefficient of t^4.) Setting $p = k^{-1}$ gives us $\kappa_3 = \sum_{k=2}^n k^{-1}(1-k^{-1})(1-2k^{-1}) = H_n - 3H_n^{(2)} + 2H_n^{(3)}$ in the case of distribution (8). And for (20), we have $\ln G(e^t) = t + H(nt) - H(t)$ where $H(t) = \ln\bigl((e^t-1)/t\bigr)$. Since $H'(z) = e^t/(e^t-1) - 1/t$, we have $\kappa_r = (n^r - 1)B_r/r$ for all $r \ge 2$ in this case; in particular, $\kappa_3 = 0$.

Page 497 lines 3–9 ———— 27 Mar 1995

Place lines 6–9 first; also change $k_1 \leadsto \kappa_1$, and **12.** $\leadsto$ **11.**

In the former lines 3–5, change **11** $\leadsto$ **12**, and add the following sentence to that answer: By exercise 11, the same formula applies for central moments except that we leave out all terms with $k_1 > 0$; thus $\kappa_2 = m_2$, $\kappa_3 = m_3$, $\kappa_4 = m_4 - 3m_2^2$.

Page 497 replacement for lines 3–6 of answer 13 ______ 27 Mar 1995

Let $z_n = e^{it/\sigma_n}$. When $n \to \infty$ and t is fixed, we have $z_n \to 1$; hence $\Gamma(z_n + 1) \to 1$, and

$$\begin{aligned}\lim_{n\to\infty} z_n^{-\mu_n} G_n(z_n) &= \lim_{n\to\infty} \exp\left(\frac{-it\mu_n}{\sigma_n} + (e^{it/\sigma_n} - 1)\ln n\right)\\ &= \lim_{n\to\infty} \exp\left(\frac{-t^2 \ln n}{2\sigma_n^2} + O\left(\frac{1}{\sqrt{\log n}}\right)\right) = e^{-t^2/2}.\end{aligned}$$

P. Flajolet and M. Soria [*Disc. Math.* **114** (1993), 159–180] have extended this analysis to show that $G_n(z)$ and a large family of related distributions not only are approximately normal near their mean values, they also have uniformly exponential tails, in the sense that

$$\text{probability}\left(\left|\frac{X_n - \mu_n}{\sigma_n}\right| > x\right) < e^{-ax}$$

for some positive constant a and for all n and x.

[Also change t to it in answer 14 and answer 16(c).]

Page 498 line 2 of answer 5 ______ 07 Jul 1995

$P(x) = 1 \leadsto P(x) = 1 + x$

$|1| > Mx$ when $x < 1/M \leadsto |1 + x| > Mx$ for all $x > 0$ when $M \le 1$, and for $0 < x < 1/(M-1)$ when $M > 1$

Page 498 new answers ______ 12 Aug 1996

19. Let $a_k = l$. Then a_k is a left-to-right maximum of $a_1 \ldots a_n \iff j < k$ implies $a_j < l \iff a_j > l$ implies $j > k \iff j > l$ implies $b_j > k \iff k$ is a right-to-left minimum of $b_1 \ldots b_n$.

20. We have $m_L = \max\{a_1 - b_1, \ldots, a_n - b_n\}$. Proof: If not, let k be the smallest subscript such that $a_k - b_k > m_L$. Then a_k is not a left-to-right maximum, so there is a $j < k$ with $a_j \ge a_k$. But then $a_j - b_j \ge a_k - b_k > m_L$, contradicting the minimality of k. Similarly, $m_R = \max\{b_1 - a_1, \ldots, b_n - a_n\}$.

Page 498 new answer ______ 13 May 1996

21. $g(z) = \sum p_k g_k(z)$; $\operatorname{mean}(g) = \sum p_k \operatorname{mean}(g_k)$; and $\operatorname{var}(g) = \sum p_k \operatorname{var}(g_k) + \sum_{j<k} p_j p_k \bigl(\operatorname{mean}(g_j) - \operatorname{mean}(g_k)\bigr)^2$.

Page 498 new answers ______ 16 May 1996

22. The result is trivial when $\epsilon \ge q$, so we may assume that $\epsilon < q$. Setting $x = \frac{p+\epsilon}{p}\frac{q}{q-\epsilon}$ in (25) gives $\Pr(X \ge n(p+\epsilon) \le \bigl((\frac{p}{p+\epsilon})^{p+\epsilon}(\frac{q}{q-\epsilon})^{q-\epsilon}\bigr)^n$. Now $(\frac{p}{p+\epsilon})^{p+\epsilon} \le e^{-\epsilon}$ since $t \le e^{t-1}$ for all real t. And $(q-\epsilon)\ln\frac{q}{q-\epsilon} = \epsilon - \frac{1}{2\cdot 1}\epsilon^2 q^{-1} - \frac{1}{3\cdot 2}\epsilon^3 q^{-2} - \cdots \le -\frac{1}{2q}\epsilon^2$. (A more detailed analysis yields the slightly stronger estimate $\exp(-\epsilon^2 n/(2pq))$ when $p \ge \frac{1}{2}$; still further work yields the upper bound $\exp(-2\epsilon^2 n)$ for all p.)

By reversing the roles of heads and tails we find $\Pr\bigl(X \le n(p-\epsilon)\bigr) = \Pr\bigl(n - X \ge n(q+\epsilon)\bigr) \le e^{-\epsilon^2 n/(2p)}$. (One should not confuse "tails" with the tail of a probability distribution.)

23. (a) Set $x = r$ in (24) and (25), and note that $q_k + p_k r = 1 + (r-1)p_k \le e^{(r-1)p_k}$. [Cf. H. Chernoff, *Annals of Math. Stat.* **23** (1952), 493–509.]

(b) Let $r = 1+\delta$ where $|\delta| \le 1$. Then $r^{-r}e^{r-1} = \exp(-\frac{1}{2\cdot 1}\delta^2 + \frac{1}{3\cdot 2}\delta^3 - \cdots)$, which is $\le e^{\delta^2/2}$ when $\delta \le 0$ and $\le e^{\delta^2/3}$ when $\delta \ge 0$.

(c) The function $r^{-1}e^{1-r^{-1}}$ decreases from 1 to 0 as r increases from 1 to ∞. If $r \ge 2$ it is $\le \frac{1}{2}e^{1/2} < .825$; if $r \ge 4.32$ it is $< \frac{1}{2}$.

Incidentally, the tail inequalities with $x = r$ give precisely the same estimate $(r^{-r}e^{r-1})^\mu$ when X has the Poisson distribution of exercise 16.

24. Setting $x = \frac{p-\epsilon}{p}\frac{q}{q-\epsilon}$ in (24) gives $\Pr\bigl(X \le n(p-\epsilon)\bigr) \le \bigl((\frac{p}{p-\epsilon})^{p-\epsilon}(\frac{q-\epsilon}{q})^{q-\epsilon}\bigr)^n \le e^{-\epsilon^2 n/(2pq)}$. Similarly, $x = \frac{p+\epsilon}{p}\frac{q}{q+\epsilon}$ yields $\Pr\bigl(X \ge n(p+\epsilon)\bigr) \le \bigl((\frac{p}{p+\epsilon})^{p+\epsilon}(\frac{q+\epsilon}{q})^{q+\epsilon}\bigr)^n$. Let $f(\epsilon) = (q+\epsilon)\ln(1+\frac{\epsilon}{q}) - (p+\epsilon)\ln(1+\frac{\epsilon}{p})$, and note that $f'(\epsilon) = \ln(1+\frac{\epsilon}{q}) - \ln(1+\frac{\epsilon}{p})$. It follows that $f(\epsilon) \le -\epsilon^2/(6pq)$.

Page 498 new answer for Section 1.2.11.1 ———————— 28 Mar 1995

12. See *AMM* **99** (1992), 419.

Page 498 new answer to 1.2.11.2–3 ———————— 15 Oct 1979

3. $|R_{2k}| \le |B_{2k}/(2k)!| \int_1^n |f^{(2k)}(x)|\,dx$. [*Notes:* We have $B_m(x) = (-1)^m B_m(1-x)$, and $B_m(x)$ is $m!$ times the coefficient of z^m in $ze^{xz}/(e^z-1)$. In particular, since $e^{z/2}/(e^z-1) = 1/(e^{z/2}-1) - 1/(e^z-1)$ we have $B_m(\frac{1}{2}) = (2^{1-m}-1)B_m$. It is not difficult to prove that the maximum of $|B_{2m} - B_{2m}(x)|$ for $0 \le x \le 1$ occurs at $x = \frac{1}{2}$. Now when $k \ge 2$ we have $R_{2k-2} = C_{2k} + R_{2k} = \int_1^n \bigl(B_{2k} - B_{2k}(\{x\})\bigr) f^{(2k)}(x)\,dx/(2k)!$, and $B_{2k} - B_{2k}(\{x\})$ is between 0 and $(2-2^{1-2k})B_{2k}$, hence R_{2k-2} lies between 0 and $(2-2^{1-2k})C_{2k}$. It follows that R_{2k} lies between $-C_{2k}$ and $(1-2^{1-2k})C_{2k}$, a slightly stronger result. According to this argument we see that if $f^{(2k+2)}(x)\,f^{(2k+4)}(x) > 0$ for $1 < x < n$, the quantities C_{2k+2} and C_{2k+4} have opposite signs, while R_{2k} has the sign of C_{2k+2} and R_{2k+2} has the sign of C_{2k+4} and $|R_{2k+2}| \le |C_{2k+2}|$; this proves (13). Cf. J. F. Steffensen, *Interpolation* (Baltimore: 1937), §14.]

Page 499 answer 7 ———————— 15 Nov 1978

(It is "Glaisher's constant" 1.2824271. . .) To $\rightsquigarrow$ To

This formula . . . $n = 4$. $\rightsquigarrow$ The number A is "Glaisher's constant" 1.2824271 . . . [*Messenger of Math.* **7** (1877), 43–47], which can be shown to equal $e^{1/12-\zeta'(-1)} = (2\pi e^{\gamma-\zeta'(2)/\zeta(2)})^{1/12}$ [de Bruijn, *Asymptotic Methods in Analysis*, §3.7].

Page 499 new answer for Section 1.2.11.2 ———————— 28 Mar 1995

8. (a) We have $\ln(2n)! = (2n+\frac{1}{2})\ln 2n - 2n + \sigma + \frac{1}{24n} + O(n^{-3})$, and $\ln(n!)^2 = (2n+1)\ln n - 2n + 2\sigma + \frac{1}{6n} + O(n^{-3})$; hence $\binom{2n}{n} = \exp\bigl(2n\ln 2 - \frac{1}{2}\ln \pi n - \frac{1}{8n} + O(n^{-3}\bigr) = 2^{2n}(\pi n)^{-1/2}\bigl(1 - \frac{1}{8}n^{-1} + \frac{1}{128}n^{-2} + O(n^{-3})\bigr)$. (b) Since $\binom{2n}{n} = 2^{2n}\binom{n-1/2}{n}$ and $\binom{n-1/2}{n} = \Gamma(n+1/2)/\bigl(n\Gamma(n)\Gamma(1/2)\bigr) = n^{-1}n^{\overline{1/2}}/\sqrt{\pi}$, we obtain the same result from 1.2.11.1–(14) because

$$\left[{1/2 \atop 1/2}\right] = 1, \quad \left[{1/2 \atop -1/2}\right] = \binom{1/2}{2} = -\frac{1}{8}, \quad \left[{1/2 \atop -3/2}\right] = \binom{1/2}{4} + 2\binom{3/2}{4} = \frac{1}{128}.$$

Method (b) explains why the denominators in

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{\pi n}}\left(1 - \frac{n^{-1}}{8} + \frac{n^{-2}}{128} + \frac{5n^{-3}}{1024} - \frac{21n^{-4}}{32768} - \frac{399n^{-5}}{262144} + \frac{869n^{-6}}{4194304} + O(n^{-7})\right)$$

are all powers of 2 [*AMM* **97** (1990), 629–630].

▸**Page 500** replacement for exercise 5 —— 29 Nov 1977

5. $f''(x) = f(x)\bigl((n+1/2)(n-1/2)/x^2 - (2n+1)/x + 1\bigr)$ changes sign at $r = n + 1/2 - \sqrt{n+1/2}$, so $|R| = O\bigl(\int_0^n |f''(x)|\,dx\bigr) = O\bigl(\int_0^r f''(x)\,dx - \int_r^n f''(x)\,dx\bigr) = O\bigl(f'(n) - 2f'(r) + f'(0)\bigr) = O\bigl(f(n)/\sqrt{n}\,\bigr)$.

▸**Page 500** answer 6 —— 14 May 1976

$O(n^{-3}) \rightsquigarrow O(n^{-3}))$

Page 500 line 3 of answer 7 —— 12 Dec 1985

(cf. next exercise) $\rightsquigarrow$

Page 500 replacement for answer 8 —— 12 Dec 1985

8. (Solution by Miklós Simonovits.) We have $|f(x)| < x$ if x is large enough. Let $R(x) = \int_0^{f(x)} (e^{-g(u,x)} - e^{-h(u,x)})\,du$ be the difference between the two given integrals, where $g(u,x) = u - x\ln(1+u/x)$ and $h(u,x) = u^2/2x - u^3/3x^2 + \cdots + (-1)^m u^m/mx^{m-1}$. Notice that $g(u,x) \ge 0$ and $h(u,x) \ge 0$ when $|u| < x$; also $g(u,x) = h(u,x) + O(u^{m+1}/x^m)$.

According to the mean value theorem, $e^a - e^b = (a-b)\,e^c$ for some c between a and b. Therefore $|e^a - e^b| \le |a-b|$ when $a, b \le 0$. It follows that

$$\begin{aligned} |R(x)| \le \int_{-|f(x)|}^{|f(x)|} |g(u,x) - h(u,x)|\,du &= O\left(\int_{-Mx^r}^{Mx^r} \frac{u^{m+1}du}{x^m}\right) \\ &= O(x^{(m+2)r-m}) = O(x^{-s}). \end{aligned}$$

Page 500 new paragraph at and of answer 9 —— 08 Jul 1995

The answer to exercise 11 gives another way to solve this problem.

▸**Page 501** replacement for answer 11 —— 08 Jul 1995

11. First, $xQ_x(n) + R_{1/x}(n) = n!\,(x/n)^n e^{n/x}$ generalizes (4). Next, we have $R_x(n) = n!\,(e^x/nx)^n \gamma(n, nx)/(n-1)!$, generalizing (9). Since $a\gamma(a,x) = \gamma(a+1,x) + e^{-x}x^a$ we can also write $R_x(n) = 1 + (e^x/nx)^n \gamma(n+1, nx)$, relating this problem to exercise 9. Moreover, we can tackle $Q_x(n)$ and $R_x(n)$ directly by using Eqs. 1.2.9–(27) and (28) to derive series expansions involving Stirling numbers:

$$1 + xQ_x(n) = \sum_{k\ge 0} x^k n^{\underline{k}}/n^k = \sum_{k,m} \frac{(-1)^m}{n^m} \left[{k \atop k-m}\right] x^k\,;$$

$$R_x(n) = \sum_{k\ge 0} x^k n^k/(n+1)^{\overline{k}} = \sum_{k,m} \frac{(-1)^m}{n^m} \left\{{k+m \atop k}\right\} x^k\,.$$

The sums over k are convergent for fixed m when $|x| < 1$, and when $|x| > 1$ we can use the relation between $Q_x(n)$ and $R_{1/x}(n)$; this leads to the formulas

$$Q_x(n) = \frac{1}{1-x} - \frac{x}{(1-x)^3 n} + \cdots + \frac{(-1)^m q_m(x)}{(1-x)^{2m+1} n^m} + O(n^{-1-m}),$$

$$R_x(n) = \frac{1}{1-x} - \frac{x}{(1-x)^3 n} + \cdots + \frac{(-1)^m r_m(x)}{(1-x)^{2m+1} n^m} + O(n^{-1-m}), \quad \text{if } x < 1;$$

$$Q_x(n) = \frac{n!\, x^{n-1} e^{n/x}}{n^n} + \frac{1}{1-x} - \frac{x}{(1-x)^3 n} + \cdots + \frac{(-1)^m q_m(x)}{(1-x)^{2m+1} n^m} + O(n^{-1-m}),$$

$$R_x(n) = \frac{n!\, e^{nx}}{n^n x^n} + \frac{1}{1-x} - \frac{x}{(1-x)^3 n} + \cdots + \frac{(-1)^m r_m(x)}{(1-x)^{2m+1} n^m} + O(n^{-1-m}), \quad \text{if } x > 1.$$

Here $q_m(x) = \left\langle\!\left\langle {m \atop 0} \right\rangle\!\right\rangle x^{2m-1} + \left\langle\!\left\langle {m \atop 1} \right\rangle\!\right\rangle x^{2m-2} + \cdots$ and $r_m(x) = \left\langle\!\left\langle {m \atop 0} \right\rangle\!\right\rangle x + \left\langle\!\left\langle {m \atop 1} \right\rangle\!\right\rangle x^2 + \cdots$ are polynomials whose coefficients are "second-order Eulerian numbers" [*CMath* §6.2; cf. L. Carlitz, *Proc. Amer. Math. Soc.* **16** (1965), 248–252]. The case $x = -1$ is somewhat delicate, but it can be handled by continuity, because the bound implied by $O(n^{-1-m})$ is independent of x when $x < 0$. It is interesting to note that $R_{-1}(n) - Q_{-1}(n) = (-1)^n n!/e^n n^n \sim (-1)^n \sqrt{2\pi n}/e^{2n}$ is extremely small.

Page 501 new answer ——— 30 Jun 1995

13. See Flajolet, Grabner, Kirschenhofer, Prodinger, *J. Computational and Applied Math.* **58** (1995), 103–116.

Page 501 new answer ——— 16 Mar 1981

18. Let $S_n(x,y) = \sum \binom{n}{k}(x+k)^k(y+n-k)^{n-k}$. Then for $n > 0$ we have $S_n(x,y) = x\sum\binom{n}{k}(x+k)^{k-1}(y+n-k)^{n-k} + n\sum\binom{n-1}{k}(x+1+k)^k(y+n-1-k)^{n-1-k} = (x+y+n)^n + nS_{n-1}(x+1,y)$ by Abel's formula 1.2.6–(16); consequently $S_n(x,y) = \sum\binom{n}{k}k!\,(x+y+n)^{n-k}$. [This formula is due to Cauchy, who proved it by quite different means in *Exercices de Mathématiques* (Paris: 1826), 62–73.] The stated sums are therefore equal respectively to $n^n\bigl(1+Q(n)\bigr)$ and $(n+1)^n Q(n+1)$.

Page 501 new answers ——— 19 Aug 1996

19. Suppose C_n exists for all $n \geq N$ and $|f(x)| \leq Mx^\alpha$ for $0 \leq x \leq r$. Let $F(x) = \int_r^x e^{-Nt} f(t)\,dt$. Then when $n > N$ we have

$$\begin{aligned} |C_n| &\leq \int_0^r e^{-nx}|f(x)|\,dx + \left|\int_r^\infty e^{-(n-N)x}e^{-Nx}f(x)\,dx\right| \\ &\leq M\int_0^r e^{-nx}x^\alpha\,dx + (n-N)\left|\int_r^\infty e^{-(n-N)x}F(x)\,dx\right| \\ &\leq M\int_0^\infty e^{-nx}x^\alpha\,dx + (n-N)\sup_{x\geq r}|F(x)|\int_r^\infty e^{-(n-N)x}\,dx \\ &= M\Gamma(\alpha+1)n^{-1-\alpha} + \sup_{x\geq r}|F(x)|e^{-(n-N)r} = O(n^{-1-\alpha}). \end{aligned}$$

[E. W. Barnes, *Phil. Trans.* **A206** (1906), 249–297; G. N. Watson, *Proc. London Math. Soc.* **17** (1918), 116–148.]

20. [C. C. Rousseau, *Applied Math. Letters* **2** (1989), 159–161.] We have $Q(n)+1 = n\int_0^\infty e^{-nx}(1+x)^n\,dx = n\int_0^\infty e^{-n(x-\ln(1+x))}\,dx = n\int_0^\infty e^{-nu}g(u)\,du$, where $u = x -$

$\ln(1+x)$ and $g(u) = dx/du$. Notice that $x = \sum_{k=1}^{\infty} c_k (2u)^{k/2}$ when u is sufficiently small. Hence $g(u) = \sum_{k=1}^{m-1} c_k (2u)^{k/2-1} + O(u^{m/2-1})$, and we can apply Watson's lemma to $Q(n) + 1 - n \int_0^{\infty} e^{-nu} \sum_{k=1}^{m-1} kc_k (2u)^{k/2-1}\,du$.

▸**Page 502** lines 4 and 5 of answer 14 ——— 18 Feb 1997

`ENT`i`0,`i for $1 \le i \le 6$; $\rightsquigarrow$

Page 502 answer 17(b) ——— 15 Mar 1975

`J2NN` $\rightsquigarrow$ `J2P`
(Using assembly ... section.) $\rightsquigarrow$ (A slightly faster, but quite preposterous, program uses 993 `STZ`'s: `JMP 3995`; `STZ 1,2`; `STZ 2,2`; ...; `STZ 993,2`; `J2N 3999`; `DEC2 993`; `J2NN 3001`; `ENN1 0,2`; `JMP 3000,1`.)

Page 502 new sentence at end of answer 18 ——— 15 Mar 1975

(Unless the program itself appears in locations 0000–0015.)

Page 502 answer 19 ——— 19 Apr 1977

$24u \rightsquigarrow 42u$
$1+1)u. \rightsquigarrow 10+10)u.$

Page 504 near the top ——— 14 Jun 1977

line 1: $|X| \le 6 \rightsquigarrow |X| \le 5$
lines 6–10:

```
(3495)  (−5)^13
(3496)  (−4)^13
   ⋮
(3505)  (+5)^13
```

space = 14, time = $4u$.

Page 504 answer 25 ——— 19 Apr 1977

line 11: operations" $\rightsquigarrow$ operations," jumps on register even or odd, and binary shifts
lines 11–12: Chapter 7 $\rightsquigarrow$ Section 7.1

▸**Page 506** answer 6 line 3 ——— 11 May 1976

$\sqrt{n} \rightsquigarrow \sqrt{\mathtt{N}}$

▸**Page 507** lines 3 and 4 ——— 17 Jul 1981

second-last ... is all blank. $\rightsquigarrow$ third-last line ends with `AAAAA` and 35 blanks; the final two lines are entirely blank.

Page 509 new paragraph at end of answer 10 ——— 28 Sep 1996

If an $m \times n$ matrix has distinct elements, we can solve the problem by looking at only $O(m+n)$ elements and doing $O(m \log n)$ auxiliary operations. See Bienstock, Chung, Fredman, Schäffer, Shor, and Suri, *AMM* **98** (1991), 418–419.

Page 511–513, answer 14 ——— 14 Jun 1977

line 1: uses as much $\rightsquigarrow$ due in part to J. Petolino uses a lot of
line 2: as possible, in $\rightsquigarrow$ in
line 9: `INCX 1` $\rightsquigarrow$
line 10: `G` $\rightsquigarrow$ `GMINUS1`

the lines beginning '`INCA 1`' and ending '`JLE EASTER+1`'

```
         INCA 61               below)
         STA  CPLUS60(1:2)
         MUL  =3//4+1=
         STA  XPLUS57(1:2)
CPLUS60  ENTA *
         MUL  =8//25+1=       rA ← Z + 24
GMINUS1  ENT2 *               E5.
         ENT1 1,2             rI1 ← G.
         INC2 1,1
         INC2 0,2
         INC2 0,1
         INC2 0,2
         INC2 773,1           rI2 ← 11G + 773.
XPLUS57  INCA -*,2            rA ← 11G + Z − X + 20 + 24 · 30 (≥ 0).
         SRAX 5
         DIV  =30=            rX ← E.
         DECX 24
         JXN  4F
         DECX 1
         JXP  2F
         JXN  3F
         DEC1 11
         J1NP 2F
3H       INCX 1
2H       DECX 29              E6.
4H       STX  20MINUSN(0:2)
         LDA  Y               E4.
         MUL  =1//4+1=
         ADD  Y
         SUB  XPLUS57(1:2)    rA ← D − 47.
20MINUSN ENN1 *
         INCA 67,1            E7.
         SRAX 5               rX ← D + N
         DIV  =7=
         SLAX 5
         DECA -4,1            rA ← 31 − N
         JAN  1F              E8.
         DECA 31
         CHAR
         LDA  MARCH
         JMP  2F
1H       CHAR
         LDA  APRIL
2H       JBUS *(18)
         ⋮                    (continue as in the old program)
         ORIG *+20
BEGIN    ENTX 1950            "driver"
```

```
ENT6 1950-2000     routine,
JMP  EASTER        uses the
INC6 1             subroutine
ENTX 2000,6        above.
J6NP EASTER+1
```

Page 513 new answer ________________ 14 Jun 1977

15. The first such year is A.D. 10317, although the error *almost* leads to failure in A.D. $10108 + 19k$ for $0 \le k \le 10$.

Incidentally, Robert Hill has pointed out that the date of Easter repeats with a period of 5,700,000 years. The most common date is April 19 (220400 times per period), while the earliest and least common is March 22 (27550 times); the latest, and next-to-least common, is April 25 (42000 times).

▸**Page 514** line 17 ________________ 21 Feb 1997

`0013.05362` $\rightsquigarrow$ `0013.05363`

Page 514 line 18 ________________ 29 Nov 1977

time. $\rightsquigarrow$ time. (It would be faster to calculate $R_n(1/m)$ directly when m is small, and then to apply the suggested procedure.)

Page 515 line 5 ________________ 18 Oct 1979

For ... history, $\rightsquigarrow$ *Historical notes:* C. Haros gave a (more complicated) rule for constructing such sequences, in *J. de l'École Polytechnique* **4**, 11 (1802), 364–368; his method was correct, but his proof was inadequate. Several years later, the geologist John Farey independently conjectured that x_k/y_k is always equal to $(x_{k-1} + x_{k+1})/(y_{k-1} + y_{k+1})$ [*Philos. Magazine and Journal* **47** (1816), 385–386]; a proof was supplied shortly afterwards by A. Cauchy [*Bull. Sociéte Philomathique de Paris* (3) **3** (1816), 133–135], who attached Farey's name to the series. For more of its interesting properties,

▸**Page 516** lines 11–14 ________________ 20 Feb 1997

[The alphabetic constants `CABA`, `CBBB`, `ACAB`, and `BCBB` should be shifted one space to the right.]

▸**Page 516** lines −4, −3 ________________ 19 Apr 1977

$3) + 7 \rightsquigarrow 7.5) + 16$

Page 518 lines 4 and 5 ________________ 27 May 1978

memory. Reference: $\rightsquigarrow$
memory, and it performs any desired permutation *in situ.* References:
19–27. $\rightsquigarrow$
19–27; E. G. Cate and D. W. Twigg, *ACM Trans. Math. Software* **3** (1977), 104–110; F. E. Fich, J. I. Munro, and P. V. Poblete, *SICOMP* **24** (1995), 266–278.

Page 519 better answer to exercise 22(d) ________________ 14 May 1976

Since the α's are independently chosen, the $\rightsquigarrow$ The

Page 520 line 6 ________________ 14 May 1976

dev $\sqrt{1/m}\,$). $\rightsquigarrow$ dev $\sqrt{1/m}\,$), when $n \ge 2m$.

Page 520 better answer to exercise 23 ——— 14 May 1976

line 1: $\int_0^1 \ldots (\ln t)$. $\leadsto$ $\int_0^\infty \exp(-t - E_1(t))\,dt$, where $E_1(x) = \int_x^\infty e^{-t}dt/t$.
line 4: $\ln n/e^\gamma$ $\leadsto$ $e^{-\gamma}\ln n$
line 6: 8310. . . ; $\leadsto$ 83100 83724 41796+ [*Math. Comp.* **22** (1968), 411–415];
line 7: known. $\leadsto$ known. The same constant had, however, been computed in another context by Karl Dickman in *Arkiv för Mat., Astron. och Fys.* **22A**, 10 (1930), 1–14; the coincidence wasn't noticed until many years later [*Theor. Comp. Sci.* **3** (1976), 373].

Page 520 line 4 of answer 23 ——— 27 Apr 1996

are not yet known. $\leadsto$ have been found by Xavier Gourdon [to appear]; the series begins

$$\lambda n + \tfrac{1}{2}\lambda - \tfrac{1}{24}e^\gamma n^{-1} + \bigl(\tfrac{1}{48}e^\gamma - \tfrac{1}{8}(-1)^n\bigr)n^{-2} + \bigl(\tfrac{17}{3840}e^\gamma + \tfrac{1}{8}(-1)^n + \tfrac{1}{6}\omega^{1-n} + \tfrac{1}{6}\omega^{n-1}\bigr)n^{-3},$$

where $\omega = e^{2\pi i/3}$.

Page 521 new answers ——— 30 Jun 1995

31. See *CMath*, Section 3.3.

32. (a) In fact, $k-1 \le \pi_k \le k+2$ when k is even; $k-2 \le \pi_k \le k+1$ when k is odd. (b) Choose the exponents from left to right, setting $e_k = 1$ iff k and $k+1$ are in different cycles of the permutation so far. [Steven Alpern, *J. Combinatorial Theory* **B25** (1978), 62–73.]

Page 521 new answers ——— 22 Feb 1997

34. Let $N = m + n$. If $m \perp n$ there is only one cycle, because every element can be written in the form $am \bmod N$ for some integer a. And in general if $d = \gcd(m, n)$, there are exactly d cycles C_0, C_1, $\ldots$, C_{d-1}, where C_j contains the elements $\{j, j+d, \ldots, j+N-d\}$ in some order. To carry out the permutation, we can therefore proceed as follows for $0 \le j < d$ (in parallel, if convenient): Set $t \leftarrow x_j$ and $k \leftarrow j$; then while $(k+m) \bmod N \ne j$, set $x_k \leftarrow x_{(k+m) \bmod N}$ and $k \leftarrow (k+m) \bmod N$; finally set $x_k \leftarrow t$. In this algorithm the relation $(k+m) \bmod N \ne j$ will hold if and only if $(k+m) \bmod N \ge d$, so we can use whichever test is more efficient.

35. Let $M = l + m + n$ and $N = l + 2m + n$. The cycles for the desired rearrangement are obtained from the cycles of the permutation on $\{0, 1, \ldots, N-1\}$ that takes k to $(k+l+m) \bmod N$, by simply striking out all elements of each cycle that are $\ge M$. (Compare this behavior with the similar situation in exercise 29.) *Proof:* When the hinted interchange sets $x_k \leftarrow x_{k'}$ and $x_{k'} \leftarrow x_{k''}$ for some k with $k' = (k+l+m) \bmod N$ and $k'' = (k'+l+m) \bmod N$ and $k' \ge M$, we know that $x_{k'} = x_{k''}$; hence the rearrangement $\alpha\beta\gamma \to \gamma\beta\alpha$ replaces x_k by $x_{k''}$.

It follows that there are exactly $d = \gcd(l+m, m+n)$ cycles, and we can use an algorithm similar to the one in the previous exercise.

A slightly simpler way to reduce this problem to the special case in exercise 34 is also noteworthy, although it makes a few more references to memory: Suppose $\gamma = \gamma'\gamma''$ where $|\gamma''| = |\alpha|$. Then we can change $\alpha\beta\gamma'\gamma''$ to $\gamma''\beta\gamma'\alpha$, and interchange γ'' with $\beta\gamma'$. A similar approach works if $|\alpha| > |\gamma|$. [See J. L. Mohammed and C. S. Subi, *J. Algorithms* **8** (1987), 113–121.]

▸**Page 522** replacement for lines 7 and 8 of answer 6 ——— 09 Aug 1981

```
STA  *+2(0:3)
LD2  5F(0:2)
ENTA *
LD2  3F(0:2)
```

Page 521 new answer ——— 26 Sep 1996

33. For $l = 0$, let $(\alpha_{01}, \alpha_{02}; \beta_{01}, \beta_{02}) = (\pi, \rho; \epsilon, \epsilon)$ and $(\alpha_{11}, \alpha_{12}; \beta_{11}, \beta_{12}) = (\epsilon, \epsilon; \pi, \rho)$, where $\pi = (1\,4)(2\,3)$, $\rho = (1\,5)(2\,4)$, and $\epsilon = (1)$.

Suppose we have made such a construction for some $l \geq 0$, where $\alpha_{jk}^2 = \beta_{jk}^2 = (1)$ for $0 \leq j < m$ and $1 \leq k \leq n$. Then the permutations

$$\begin{aligned}(A_{(jm+j')1}, \ldots, A_{(jm+j')(4n)}; B_{(jm+j')1}, \ldots, B_{(jm+j')(4n)}) &= \\ (\sigma^- \alpha_{j1}\sigma, \ldots, \sigma^- \alpha_{jn}\sigma,\ \tau^- \alpha_{j'1}\tau, \ldots, \tau^- \alpha_{j'n}\tau,& \\ \sigma^- \beta_{jn}\sigma, \ldots, \sigma^- \beta_{j1}\sigma,\ \tau^- \beta_{j'n}\tau, \ldots, \tau^- \beta_{j'1}\tau;& \\ \sigma^- \beta_{j1}\sigma, \ldots, \sigma^- \beta_{jn}\sigma,\ \tau^- \beta_{j'1}\tau, \ldots, \tau^- \beta_{j'n}\tau,& \\ \sigma^- \alpha_{jn}\sigma, \ldots, \sigma^- \alpha_{j1}\sigma,\ \tau^- \alpha_{j'n}\tau, \ldots, \tau^- \alpha_{j'1}\tau)&\end{aligned}$$

have the property that

$$\begin{aligned}A_{(im+i')1}B_{(jm+j')1} \ldots A_{(im+i')(4n)}B_{(jm+j')(4n)} &= \\ \sigma^-(1\,2\,3\,4\,5)\sigma\,\tau^-(1\,2\,3\,4\,5)\tau\,\sigma^-(5\,4\,3\,2\,1)\sigma\,\tau^-(5\,4\,3\,2\,1)\tau&\end{aligned}$$

if $i = j$ and $i' = j'$, otherwise the product is (1). Choosing $\sigma = (2\,3)(4\,5)$ and $\tau = (3\,4\,5)$ will make the product $(1\,2\,3\,4\,5)$ as desired, when $im + i' = jm + j'$.

The construction that leads from l to $l+1$ is due to David A. Barrington [*J. Comp. Syst. Sci.* **38** (1989), 150–164], who proved a general theorem by which any Boolean function can be represented as a product of permutations of $\{1, 2, 3, 4, 5\}$. With a similar construction we can, for example, find sequences of permutations $(\alpha_{j1}, \ldots, \alpha_{jn}; \beta_{j1}, \ldots, \beta_{jn})$ such that

$$\alpha_{i1}\beta_{j1}\alpha_{i2}\beta_{j2} \ldots \alpha_{in}\beta_{jn} = \begin{cases}(1\,2\,3\,4\,5), & \text{if } i < j; \\ (1), & \text{if } i \geq j;\end{cases}$$

for $0 \leq i, j < m = 2^{2^l}$ when $n = 6^{l+1} - 4^{l+1}$.

▸**Page 525** changes to answer 1 ——— 26 Feb 1997

`      ST1   2F(0:2)` $\rightsquigarrow$ `ST1   4F`
`2H    ENT1  *` $\rightsquigarrow$ `LD1   4F`
and add a new line “`4H    CON   0`”.

▸**Page 526** line 3 ——— 26 Feb 1997

`I1REG ENT1  *` $\rightsquigarrow$ `LD1   I1REG(0:2)`

▸**Page 529** line 5 ——— 14 Mar 1976

process would loop indefinitely; $\rightsquigarrow$ algorithm breaks down (possibly referring to the buffer while I/O is in progress);

▸**Page 529** line 2 of answer 18 ——— 26 Feb 1997

active $\rightsquigarrow$ idle
idle $\rightsquigarrow$ active

Page 529 new answer ——— 20 Sep 1996

19. If $C \le L$ we can have $t_k = (k-1)L$, $u_k = t_k + T$, and $v_k = u_k + C$ if and only if $NL \ge T + C$. If $C > L$ the situation is more complex; we can have $u_k = (k-1)C + T$ and $v_k = kC + T$ if and only if there are integers $a_1 \le a_2 \le \cdots \le a_n$ such that $t_k = (k-1)L + a_k P$ satisfies $u_k - T \ge t_k \ge v_{k-N}$ for $N < k \le n$. An equivalent condition is that $NC \ge b_k$ for $N < k \le n$, where $b_k = C + T + \bigl((k-1)(C-L)\bigr) \bmod P$. Let $c_l = \max\{b_{l+1}, \ldots, b_n, 0\}$; then c_l decreases as l increases, and the smallest value of N that keeps the process going steadily is the minimum l such that $c_l/l \le C$. Since $c_l < C + T + P$ and $c_l \le L + T + n(C-L)$, this value of N never exceeds $\lceil \min\{C+T+P, L+T+n(C-L)\}/C \rceil$. [Cf. A. Itai and Y. Raz, *CACM* **31** (1988), 1339–1342.]

In the stated example we have therefore (a) $N = 1$; (b) $N = 2$; (c) $N = 3$, $c_N = 2.5$; (d) $N = 35$, $c_N = 51.5$; (e) $N = 51$, $c_N = 101.5$; (f) $N = 41$, $c_N = 102$; (g) $N = 11$, $c_N = 109.5$; (h) $N = 3$, $c_N = 149.5$; (i) $N = 2$, $c_N = 298.5$.

Page 531 line -2 ——— 18 Oct 1979

X's. For the history of the ballot problem $\rightsquigarrow$ X's. This problem was actually resolved as early as 1708 by Abraham De Moivre, who showed that the number of sequences containing l A's and m B's, and containing at least one initial substring with n more A's than B's, is $f(l,m,n) = \binom{l+m}{\min(m,l-n)}$. In particular, $a_n = \binom{2n}{n} - f(n,n,1)$ as above. (De Moivre stated this result without proof [*Philos. Trans.* **27** (1711), 262–263]; but it is clear from other passages in his paper that he knew how to prove it, since the formula is obviously true when $l \ge m+n$, and since his generating-function approach to similar problems yields the symmetry condition $f(l,m,n) = f(m+n, l-n, n)$ by simple algebra.) For the later history of the ballot problem

Page 533 additional paragraph for answer 5 ——— 29 Sep 1996

P. V. Ramanan [*SICOMP* **13** (1984), 167–169] has shown how to characterize the permutations obtainable when m auxiliary storage locations can be used freely in addition to a stack. (This generalization of the problem is surprisingly difficult.)

▸**Page 533** answer 9 ——— 08 Dec 1975

in reverse, we can get the inverse $\rightsquigarrow$ backwards, we can get the reverse of the inverse of the reverse

▸**Page 534** line 10 ——— 15 Mar 1975

$(1 - 2r_2(z)) \rightsquigarrow (1 - 2xr_2(z))$

Page 534 replacement for answer 12 ——— 28 Mar 1995

12. We have $a_n = 4^n/\sqrt{\pi n^3} + O(4^n n^{-5/2})$ by Stirling's formula. To analyze b_n, let us first consider the general problem of estimating the coefficient of w^n in the power series for $\sqrt{1-w}\,\sqrt{1-\alpha w}$ when $|\alpha| < 1$. We have

$$\sqrt{1-w}\,\sqrt{1-\alpha w} = \sqrt{1-w}\,\sqrt{1-\alpha+\alpha(1-w)} = \sqrt{1-\alpha}\,\sum_k \binom{1/2}{k} \beta^k (1-w)^{k+1/2},$$

where $\beta = \alpha/(1-\alpha)$; hence the desired coefficient is $(-1)^n \sqrt{1-\alpha}\,\sum_k \binom{1/2}{k}\beta^k\binom{k+1/2}{n}$. Now

$$(-1)^n \binom{k+1/2}{n} = \binom{n-k-3/2}{n} = \frac{\Gamma(n-k-1/2)}{\Gamma(n+1)\Gamma(-k-1/2)} = -\frac{(k+1/2)^{\underline{k+1}}}{\sqrt{\pi}\,n}\, n^{\overline{-k-1/2}},$$

and $n^{\overline{-k-1/2}} = \sum_{j=0}^{m} \left[{-k-1/2 \atop -k-1/2-j}\right] n^{-k-1/2-j} + O(n^{-k-3/2-m})$ by Eq. 1.2.11.1–(14). Thus we obtain the asymptotic series $[w^n]\sqrt{1-w}\sqrt{1-\alpha w} = c_0 n^{-3/2} + c_1 n^{-5/2} + \cdots + c_m n^{-m-3/2} + O(n^{-m-5/2})$ where

$$c_j \;=\; -\sqrt{\frac{1-\alpha}{\pi}} \sum_{k=0}^{j} \binom{1/2}{k} \bigl(k+\tfrac12\bigr)^{\underline{k+1}} \left\{ {j+1/2 \atop k+1/2} \right\} \frac{\alpha^k}{(1-\alpha)^k}\,.$$

For b_n, we write $1-6z+z^2 = \bigl(1-(3+\sqrt{8}\,)z\bigr)\bigl(1-(3-\sqrt{8}\,)z\bigr)$ and let $w = (3+\sqrt{8}\,)z$, $\alpha = (3-\sqrt{8}\,)/(3+\sqrt{8}\,)$, obtaining the asymptotic formula

$$b_n = \frac{(\sqrt{2}-1)(3+\sqrt{8}\,)^n}{2^{3/4}\pi^{1/2}n^{3/2}}\bigl(1+O(n^{-1})\bigr) = \frac{(\sqrt{2}+1)^{2n-1}}{2^{3/4}\pi^{1/2}n^{3/2}}\bigl(1+O(n^{-1})\bigr)\,.$$

Page 534 new answer ——— 30 Jun 1995

14. (Solution by R. Melville, 1980.) Let R and S be stacks such that the queue runs from top to bottom of R followed by bottom to top of S. When R is empty, pop the elements of S onto R until S becomes empty. To delete from the front, pop the top of R, which will not be empty unless the entire queue is empty. To insert at the rear, push onto S (unless R is empty). Each element is pushed at most twice and popped at most twice before leaving the queue.

▸**Page 535** answer 4(ii) ——— 15 Mar 1975

`X,7:7` $\rightsquigarrow$ `X,7:7(0:2)`

▸**Page 537** replacement for last line of answer 11 ——— 07 Mar 1997

$$\sum_m E_{mnt} z^m = \frac{n-1}{2n}\,\frac{z^2}{(1-z)^3}\left(\frac{z}{n-(n-1)z}\right)^{t}.$$

Page 538 new answer ——— 01 Mar 1979

13. A. C. Yao has shown that $\max(k_1,k_2)$ will be $\frac12 m + \bigl(2\pi(1-2p)\bigr)^{-1/2}\sqrt{m} + O\bigl(m^{-1/2}(\log m)^2\bigr)$ for large m, when $p < \frac12$. [*SICOMP* **10** (1981), 398–403.]

Page 541 new answer ——— 11 Feb 1995

13. Sha and Kleitman, *Discrete Math.* **63** (1987), 271–278, have proved that the number is at most $\prod_{k=0}^{n}\binom{n}{k}^{\binom{n}{k}}$. This exceeds the obvious lower bound $\prod_{k=0}^{n}\binom{n}{k}! = 2^{2^n(n+O(1))}$ by a factor of $2^{2^n+O(n)}$; they conjecture that the lower bound is closer to the truth.

Page 543 better answer to exercises 23 and 24 ——— 14 May 1976

T10. If `P` $\neq \Lambda$, set `QLINK[SUC(P)]` $\leftarrow k$, `P` $\leftarrow$ `NEXT(P)`, and repeat this step.

85 `ST6 X,1(QLINK)` `QLINK[`rI1`]` $\leftarrow k$.

Delete lines *86–87*; renumber lines *88–118* as *86–116*; and change the next two lines to *Note:* If the relations $9 \prec 1$ and $6 \prec 9$ are added to the data (18), this program will print "9, 6, 8, 5, 9" as the loop.

Page 546 new answer ——— 21 Sep 1976

30. To insert, set `P` $\Leftarrow$ `AVAIL`, `INFO(P)` $\leftarrow$ `Y`, `LINK(P)` $\leftarrow \Lambda$, if `F` $= \Lambda$ then `F` $\leftarrow$ `P` else `LINK(R)` $\leftarrow$ `P`, and `R` $\leftarrow$ `P`. To delete, do (9) with `F` replacing `T`.

Page 547 answer 5 ______ 03 Mar 1981

(Solution by B. Young.) $\leadsto$ (Cf. exercise 2.2.3–7.)

Page 548 first line of answer 9 ______ 17 Apr 1979

should. $\leadsto$ should — except in the anomalous case that `COEF` $= 0$ for some term with `ABC` ≥ 0, when it fails badly.

▸**Page 549** answer 16 ______ 14 May 1976

line 2: $29\sum \leadsto 27\sum$ (twice)
line 8: $6\alpha pr \leadsto 4\alpha pr$
line 10: often. $\leadsto$ often. (See exercise 5.2.3–29 for a faster algorithm.)

Page 550 top ______ 02 Mar 1977

denotes, ... are included $\leadsto$ denotes "exclusive or." Other invertible operations, such as addition or subtraction modulo the pointer field size, could also be used. It is convenient to include two adjacent list heads

▸**Page 550** at end of answer 1 ______ 03 Mar 1975

`AVAIL` $\Leftarrow$ `P` $\leadsto$ `Y` $\leftarrow$ `INFO(P)`; `AVAIL` $\Leftarrow$ `P`

Page 550 answer 2 ______ 02 Mar 1977

line 2: next ... list point $\leadsto$ next, so the links in the list must point
line 3: So ... implies $\leadsto$ Deletion at both ends therefore implies that
line 4: ways. $\leadsto$ ways. On the other hand, exercise 2.2.4–18 explains how to represent two links in a single link field; in this way general deque operations *are* possible.

▸**Page 550** line 6 of answer 3 ______ 10 Mar 1997

1398 $\leadsto$ 1378

▸**Page 554** line 2 ______ 01 Apr 1975

`COL(P)` $\leadsto$ `COL(P0)`

▸**Page 554** lines *16*, *17*, and *20* of the program comments ______ 09 Nov 1981

`([` $\leadsto$ `[`

▸**Page 554** line *26* of the program comments ______ 29 Mar 1997

`J` $< 0 \leadsto$ `J` ≥ 0

▸**Page 555** line *55* of the program ______ 29 Mar 1997

`1,5(COL)` $\leadsto$ `0,5(ROW)`

▸**Page 555** line *56* of the program comments ______ 29 Mar 1997

`COL`(rI5) $\leftarrow$ `I`. $\leadsto$ `ROW`(rI5) $>$ `I`.

▸**Page 556** line *92* of the program ______ 19 Mar 1997

`M4A` $\leadsto$ `S4A`

Page 556 new answer ______ 01 Jul 1995

17. For each row i and each element `A`$[i,k] \neq 0$, add `A`$[i,k]$ times row k of `B` to row i of `C`. Maintain only the `COL` links of `C` while doing this; the `ROW` links are easily filled in afterwards. [A. Schoor, *Inf. Proc. Letters* **15** (1982), 87–89.]

▸**Page 556** answer 20 ——— 14 May 1976

`A(1,1)` $\rightsquigarrow$ `A[1,1]`

Page 556 new answer ——— 14 May 1976

21. For example, $\mathtt{M} \leftarrow \max(\mathtt{I},\mathtt{J})$, $\mathtt{LOC(A[I,J])} = \mathtt{LOC(A[1,1])} + \mathtt{M}(\mathtt{M}-1) + \mathtt{I} - \mathtt{J}$. (Such formulas have been proposed independently by many people. A. L. Rosenberg and H. R. Strong have suggested the following k-dimensional generalization: $\mathtt{LOC(A[I_1,\ \dots\ ,I_{\mathit{k}}])} = \mathtt{L}_k$ where $\mathtt{L}_1 = \mathtt{LOC(A[1,\ \dots\ ,1])} + \mathtt{I}_1 - 1$, $\mathtt{L}_r = \mathtt{L}_{r-1} + (\mathtt{M}_r - \mathtt{I}_r) \times (\mathtt{M}_r^{r-1} - (\mathtt{M}_r - 1)^{r-1})$, and $\mathtt{M}_r = \max(\mathtt{I}_1, \dots, \mathtt{I}_r)$ [*IBM Tech. Disclosure Bull.* **14** (1972), 3026–3028]. See *Current Trends in Programming Methodology* **4** (Prentice-Hall, 1978), 263–311, for further results of this kind.)

Page 556 new answer ——— 30 Mar 1995

22. According to the binomial number system (exercise 1.2.6–56), we can let

$$p(i_1,\dots,i_k) = \binom{i_1}{1} + \binom{i_1+i_2+1}{2} + \cdots + \binom{i_1+i_2+\cdots+i_k+k-1}{k}.$$

[*Det Kongelige Norske Videnskabers Selskabs Forhandlinger* **34** (1961), 8–9.]

Page 556 new answer ——— 01 Jul 1995

23. Let $c[\mathtt{J}] = \mathtt{LOC(A[0,J])} = \mathtt{LOC(A[0,0])} + m\mathtt{J}$, if there were m rows when the matrix grew from $\mathtt{J}$ to $\mathtt{J}+1$ columns; similarly, let $r[\mathtt{I}] = \mathtt{LOC(A[I,0])} = \mathtt{LOC(A[0,0])} + n\mathtt{I}$, if there were n columns when we created row $\mathtt{I}$. Then we can use the allocation function

$$\mathtt{LOC(A[I,J])} = \begin{cases} \mathtt{I} + c[\mathtt{J}], & \text{if } c[\mathtt{J}] \ge r[\mathtt{I}]; \\ \mathtt{J} + r[\mathtt{I}], & \text{otherwise.} \end{cases}$$

It is not hard to prove that $c[\mathtt{J}] \ge r[\mathtt{I}]$ implies $c[\mathtt{J}] \ge r[\mathtt{I}] + \mathtt{J}$, and $c[\mathtt{J}] \le r[\mathtt{I}]$ implies $c[\mathtt{J}] + \mathtt{I} \le r[\mathtt{I}]$; therefore the relation

$$\mathtt{LOC(A[I,J])} = \max\bigl(\mathtt{I} + \mathtt{LOC(A[0,J])}, \mathtt{J} + \mathtt{LOC(A[I,0])}\bigr)$$

also holds. We need not restrict allocation to mn consecutive locations; the only constraint is that, when the matrix grows, we allocate m or n consecutive new cells in locations greater than those previously used. This construction is due to E. J. Otoo and T. H. Merrett [*Computing* **31** (1983), 1–9], who also generalized it to k dimensions.

Page 556 new answer ——— 01 Sep 1996

24. [Aho, Hopcroft, and Ullman, *The Design and Analysis of Computer Algorithms* (Addison-Wesley, 1974), exercise 2.12.] Besides the array `A`, maintain also a verification array `V` of the same size, and a list `L` of the locations used. Let n be the number of items in `L`; initially $n = 0$ and the contents of `L`, `A`, and `V` are arbitrary. Whenever you want to access $\mathtt{A}[k]$ for a value of k that you might not have used before, first check whether $0 \le \mathtt{V}[k] < n$ and $\mathtt{L}[\mathtt{V}[k]] = k$. If not, set $\mathtt{V}[k] \leftarrow n$, $\mathtt{L}[n] \leftarrow k$, $\mathtt{A}[k] \leftarrow 0$, and $n \leftarrow n+1$. Otherwise you can be sure that $\mathtt{A}[k]$ already contains legitimate data. (By a slight extension of this method, it is possible to save and eventually restore the contents of all entries of `A` and `V` that change during the computation.)

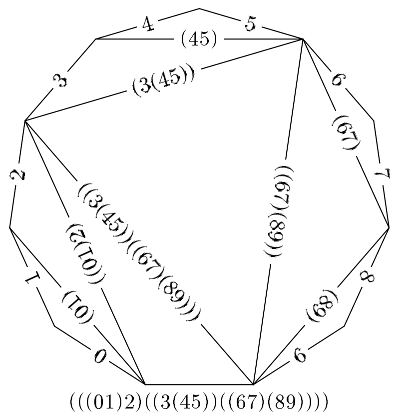
Page 559 answer 20 ——— 30 Mar 1995

b-tree $\rightsquigarrow$ 0-2 tree (five times)

Page 559 new answer ——— 31 Mar 1995

22. Choose one edge of the polygon and call it the base. Given a triangulation, let the triangle on the base correspond to the root of a binary tree, and let the other two sides of that triangle define bases of left and right subpolygons, which correspond to left and right subtrees in the same way. We proceed recursively until reaching "2-sided" polygons, which correspond to empty binary trees.

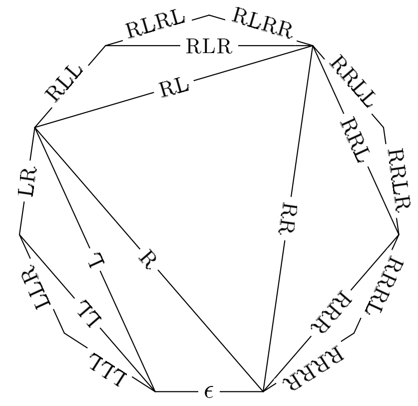
Stating this correspondence another way, we can label the non-base edges of a triangulated polygon with the integers 0, ..., n; and when two adjacent sides of a triangle are labeled α and β in clockwise order, we can label the third side $(\alpha\beta)$. The label of the base then characterizes the binary tree and the triangulation. For example,



corresponds to the binary tree shown in 2.3.1–(1).

Page 559 new answer ——— 01 Apr 1995

23. Let $m = n + 2$; by the previous exercise, we want a correspondence between triangulated m-gons and $(m - 1)$-rowed friezes. First let's look more closely at the previous correspondence, by giving a "top-down" labeling to the edges of a triangulation instead of the "bottom-up" one considered earlier: Assign the empty label ϵ to the base, then recursively give the labels αL and αR to the opposite edges of a triangle whose base is labeled α. For example, the previous diagram becomes



under these the new conventions. If the base edge in this example is called 10, while the other edges are 0 to 9 as before, we can write $0 = 10LLL$, $1 = 10LLR$, $2 = 10LR$, $3 = 10RLL$, etc. Any of the other edges can also be chosen as the base; thus, if 0 is chosen we have $1 = 0L$, $2 = 0RL$, $3 = 0RRLLL$, etc. It is not difficult to verify that if $u = v\alpha$ we have $v = u\alpha^T$, where α^T is obtained by reading α from right to left and interchanging L with R. For example, $10 = 0RRR = 1LRR = 2LR = 3RRL$, etc. If u, v, and w are edges of the polygon with $w = u\alpha L\gamma$ and $w = v\beta R\gamma$, then $u = v\beta L\alpha^T$ and $v = u\alpha R\beta^T$.

Given a triangulation of a polygon whose edges are numbered $0, 1, \ldots, m-1$, we define (u, v) for any pair of distinct edges u and v as follows: Let $u = v\alpha$, and interpret α as a 2×2 matrix by letting

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then (u, v) is defined to be the element in the upper left corner of α. Notice that α^T is the transpose of the matrix α, since $R = L^T$; hence we have $(v, u) = (u, v)$. Notice also that $(u, v) = 1$ iff u_- and v_- are joined by an edge of the triangulation, where u_- denotes the vertex between u and $u - 1$.

Let $(u, u) = 0$ for all polygon edges u. We can now prove that $v = u\alpha$ implies

$$\alpha \;=\; \begin{pmatrix} (u,v) & (u,v+1) \\ (u+1,v) & (u+1,v+1) \end{pmatrix} \qquad \text{for all } u \neq v, \tag{$*$}$$

where $u+1$ and $v+1$ are the clockwise successors of u and v. The proof is by induction on m: Eq. $(*)$ is trivial when $m = 2$, since the two parallel edges u and v are then related by $u = v\epsilon$, and $\alpha = \epsilon$ is the identity matrix. If any triangulation is augmented by extending some edge v with a triangle $v\,v'\,v''$, then $v = u\alpha$ implies $v' = u\alpha L$ and $v'' = u\alpha R$; hence (u, v') and (u, v'') in the extended polygon are respectively equal to (u, v) and $(u, v) + (u, v + 1)$ in the original one. It follows that

$$\alpha L = \begin{pmatrix} (u,v') & (u,v'') \\ (u+1,v') & (u+1,v'') \end{pmatrix} \qquad \text{and} \qquad \alpha R = \begin{pmatrix} (u,v'') & (u,v''+1) \\ (u+1,v'') & (u+1,v''+1) \end{pmatrix},$$

and $(*)$ remains true in the extended polygon.

The frieze pattern corresponding to the given triangulation is now defined to be the periodic sequence

$$\begin{array}{cccccccccccc}
(0,1) & & (1,2) & & (2,3) & & \ldots & & (m{-}1,0) & & (0,1) & & (1,2) & & \ldots & \\
& (0,2) & & (1,3) & & (2,4) & & \ldots & & (m{-}1,1) & & (0,2) & & (1,3) & & \ldots \\
(m{-}1,2) & & (0,3) & & (1,4) & & \ldots & & (m{-}2,1) & & (m{-}1,2) & & (0,3) & & \ldots & \\
& (m{-}1,3) & & (0,4) & & (1,5) & & \ldots & & (m{-}2,2) & & (m{-}1,3) & & (0,4) & & \ldots
\end{array}$$

and so on until $m - 1$ rows have been defined; the final row begins with $(\lceil m/2\rceil + 1, \lceil m/2\rceil)$ when $m > 3$. Condition $(*)$ proves that this pattern is a frieze, i.e., that

$$(u,v)(u+1,v+1) - (u,v+1)(u+1,v) = 1, \tag{$**$}$$

because $\det L = \det R = 1$ implies $\det\alpha = 1$. Our example triangulation yields

$$\begin{array}{l}
1\quad \ldots\\
\;\;1\quad 2\quad 4\quad 2\quad 1\quad 5\quad 1\quad 3\quad 1\quad 4\quad 3\quad 1\quad 2\quad 4\quad 2\quad 1\quad 5\quad 1\quad 3\quad 1\quad 4\quad \ldots\\
2\quad 1\quad 7\quad 7\quad 1\quad 4\quad 4\quad 2\quad 2\quad 3\quad 11\quad 2\quad 1\quad 7\quad 7\quad 1\quad 4\quad 4\quad 2\quad 2\quad 3\quad \ldots\\
\;\;1\quad 3\quad 12\quad 3\quad 3\quad 3\quad 7\quad 1\quad 5\quad 8\quad 7\quad 1\quad 3\quad 12\quad 3\quad 3\quad 3\quad 7\quad 1\quad 5\quad 8\quad \ldots\\
3\quad 2\quad 5\quad 5\quad 8\quad 2\quad 5\quad 3\quad 2\quad 13\quad 5\quad 3\quad 2\quad 5\quad 5\quad 8\quad 2\quad 5\quad 3\quad 2\quad 13\quad \ldots\\
\;\;5\quad 3\quad 2\quad 13\quad 5\quad 3\quad 2\quad 5\quad 5\quad 8\quad 2\quad 5\quad 3\quad 2\quad 13\quad 5\quad 3\quad 2\quad 5\quad 5\quad 8\quad \ldots\\
3\quad 7\quad 1\quad 5\quad 8\quad 7\quad 1\quad 3\quad 12\quad 3\quad 3\quad 3\quad 7\quad 1\quad 5\quad 8\quad 7\quad 1\quad 3\quad 12\quad 3\quad \ldots\\
\;\;4\quad 2\quad 2\quad 3\quad 11\quad 2\quad 1\quad 7\quad 7\quad 1\quad 4\quad 4\quad 2\quad 2\quad 3\quad 11\quad 2\quad 1\quad 7\quad 7\quad 1\quad \ldots\\
5\quad 1\quad 3\quad 1\quad 4\quad 3\quad 1\quad 2\quad 4\quad 2\quad 1\quad 5\quad 1\quad 3\quad 1\quad 4\quad 3\quad 1\quad 2\quad 4\quad 2\quad \ldots\\
\;\;1\quad \ldots
\end{array}$$

The relation $(u,v) = 1$ defines the edges of the triangulation, hence different triangulations yield different friezes. To complete the proof of one-to-one correspondence, we must show that every $(m-1)$-rowed frieze pattern of positive integers is obtained in this way from some triangulation.

Given any frieze of $m-1$ rows, extend it by putting a new row 0 at the top and a new row m at the bottom, both consisting entirely of zeros. Now let the elements of row 0 be called $(0,0)$, $(1,1)$, $(2,2)$, etc., and for all nonnegative integers $u < v \le u+m$ let (u,v) be the element in the diagonal southeast of (u,u) and in the diagonal southwest of (v,v). By assumption, condition $(**)$ holds for all $u < v < u+m$. We can in fact extend $(**)$ to the considerably more general relation

$$(t,u)(v,w) + (t,w)(u,v) = (t,v)(u,w) \qquad \text{for } t \le u \le v \le w \le t+m. \qquad (***)$$

For if $(***)$ is false, let (t,u,v,w) be a counterexample with the smallest value of $(w-t)m + u - t + w - v$. *Case 1:* $t+1 < u$. Then $(***)$ holds for $(t,t+1,v,w)$, $(t,t+1,u,v)$, and $(t+1,u,v,w)$, so we find $\bigl((t,u)(v,w) + (t,w)(v,u)\bigr)(t+1,v) = (t,v)(u,w)(t+1,v)$; this implies $(t+1,v) = 0$, a contradiction. *Case 2:* $v+1 < w$. Then $(***)$ holds for $(t,u,w-1,w)$, $(u,v,w-1,w)$, and $(t,u,v,w-1)$; we obtain a similar contradiction $(u,w-1) = 0$. *Case 3:* $u = t+1$ and $w = v+1$. In this case $(***)$ reduces to $(**)$.

Now we set $u = t+1$ and $w = t+m$ in $(***)$, obtaining $(t,v) = (v,t+m)$ for $t \le v \le t+m$, because $(t+1,t+m) = 1$ and $(t,t+m) = 0$. We conclude that the entries of any $(m-1)$-rowed frieze are periodic: $(u,v) = (v,u+m) = (u+m,v+m) = (v+m,u+2m) = \cdots$.

Every frieze pattern of positive integers contains a 1 in row 2. For if we set $t = 0$, $v = u+1$, and $w = u+2$ in $(***)$ we get $(0,u+1)(u,u+2) = (0,u) + (0,u+2)$, hence $(0,u+2) - (0,u+1) \ge (0,u+1) - (0,u)$ iff $(u,u+2) \ge 2$. This cannot hold for all u in the range $0 \le u \le m-2$, because $(0,1) - (0,0) = 1$ and $(0,m) - (0,m-1) = -1$.

Finally, if $m > 3$ we cannot have two consecutive 1s in row 2, because $(u,u+2) = (u+1,u+3) = 1$ implies $(u,u+3) = 0$. Therefore we can reduce the frieze to another one with m reduced by 1, as illustrated here for 7 rows reduced to 6:

$$\begin{array}{l}
1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad \ldots\\
\;\;a\quad b\quad c\quad d{+}1\quad 1\quad e{+}1\quad y\quad z\quad \ldots\\
\;\;\;\;p\quad q\quad c{+}r\quad d\quad e\quad u{+}y\quad v\quad w\quad \ldots\\
\;\;\;\;\;\;u\quad q{+}v\quad r\quad s\quad u\quad q{+}v\quad r\quad s\quad \ldots\\
\;\;\;\;\;\;\;\;u{+}y\quad v\quad w\quad p\quad q\quad c{+}r\quad d\quad e\quad \ldots\\
\;\;\;\;\;\;\;\;\;\;y\quad z\quad a\quad b\quad c\quad d{+}1\quad 1\quad e{+}1\;\ldots\\
\;\;\;\;\;\;\;\;\;\;\;\;1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad \ldots
\end{array}
\qquad
\begin{array}{l}
1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad \ldots\\
\;\;a\quad b\quad c\quad d\quad e\quad y\quad z\quad \ldots\\
\;\;\;\;p\quad q\quad r\quad s\quad u\quad v\quad w\quad \ldots\\
\;\;\;\;\;\;u\quad v\quad w\quad p\quad q\quad r\quad s\quad \ldots\\
\;\;\;\;\;\;\;\;y\quad z\quad a\quad b\quad c\quad d\quad e\quad \ldots\\
\;\;\;\;\;\;\;\;\;\;1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad \ldots
\end{array}$$

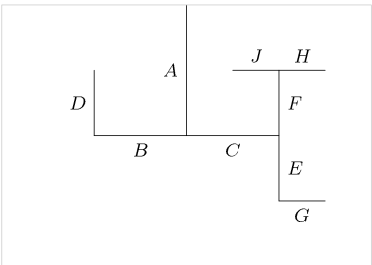
The reduced frieze corresponds to a triangulation, by induction, and the unreduced frieze corresponds to attaching one more triangle. [Cf. *Math. Gazette* **57** (1974), 87–94, 175–183.]

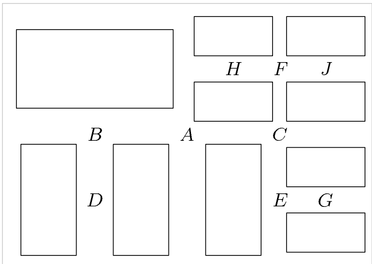
Notes: This proof demonstrates that the function (u, v) we defined on any triangulation via 2×2 matrices satisfies $(***)$ whenever (t, u, v, w) are consecutive edges of the polygon in clockwise order. We can express each (u, v) as a polynomial in the quantities $a_j = (j-1, j+1)$; these polynomials are essentially identical to the "continuants" discussed in Section 4.5.3, except for the signs of individual terms. In fact, $(j, k) = i^{1-k+j} Q_{k-j-1}(ia_{j+1}, ia_{j+2}, \ldots, ia_{k-1})$. Thus $(***)$ is equivalent to Euler's identity for continuants in the answer to exercise 4.5.3–32. The matrices L and R have the interesting property that any 2×2 matrix of nonnegative integers with determinant 1 can be expressed uniquely as a product of L's and R's. Many other interesting relationships are present; for example, the numbers in row 2 of an integer frieze count the number of triangles touching each vertex of the corresponding triangulated polygon. The total number of occurrences of $(u, v) = 1$ in the basic region $0 \le u < v - 1 < m - 1$ and $(u, v) \ne (0, m-1)$ is the number of diagonals (chords) of the triangulation, namely $m - 3 = n - 1$. The total number of 2s is also $n - 1$, because $(u, v) = 2$ iff u_- and v_- are opposing vertices of the two triangles adjacent to a chord. Another interpretation of (u, v) was found by D. Broline, D. W. Crowe, and I. M. Isaacs [*Geometriae Dedicata* **3** (1974), 171–176]: It is the number of ways to match the $v - u - 1$ vertices between edges u and $v - 1$ with distinct triangles adjacent to those vertices.

Page 566 new answer ———————————— 24 Aug 1996

24. The basic idea is to proceed recursively, with the representation of a nonempty binary tree defined to be the representation of its root plus half-size-and-rotated representations of its left and right subtrees. Thus an arbitrarily large binary tree can be represented on a single sheet of paper, if one has a sufficiently powerful magnifying glass.

Many variations on this theme are possible. For example, one idea is to represent the root by a line from the center of a given landscape-oriented page to the top edge, and to rotate the left-subtree representation by 90° clockwise in the left halfpage, the right-subtree representation by 90° counterclockwise in the right halfpage. Each node is then represented by a line. (When this method is applied to a complete binary tree having $2^k - 1$ nodes on k levels, it yields so-called "H-trees," which are the most efficient layouts of such binary trees on a VLSI chip; see R. P. Brent and H. T. Kung, *Inf. Proc. Letters* **11** (1980), 46–48.)





Another idea is to represent an empty binary tree by some sort of box, and to rotate the subtree representations of nonempty binary trees so that left subsubtrees are alternately to the left of or below the corresponding right subsubtrees, depending on whether the depth of recursion is even or odd. Then the boxes correspond to external nodes in an extended binary tree (see Section 2.3.4.5). This representation, which is strongly related to the 2-D trees and quadtrees discussed in Section 6.5, is especially appropriate when the external nodes carry information but the internal nodes do not.

Page 559 new paragraph for end of answer 5 ——— 23 Aug 1996

(Moreover, if we imagine the blanks at the left and treat the Dewey labels as ordinary binary numbers we get "level order," 2.3.3–(8).)

Page 560 new paragraph to follow answer 11 ——— 14 Apr 1995

When the binary tree represents a forest as described in Section 2.3.2, the quantity analyzed here is the *height* of that forest (the furthest distance between a node and a root, plus one). Generalizations to many other varieties of trees have been obtained by Flajolet and Odlyzko [*J. Computer and System Sci.* **25** (1982), 171–213]; the asymptotic distribution of heights, both near the mean and far away, was subsequently analyzed by Flajolet, Gao, Odlyzko, and Richmond [*Combinatorics, Probability, and Computing* **2** (1993), 145–156].

Page 560 better answer to exercise 13 ——— 14 May 1976

13. (Solution by S. Araújo.) Let steps T1 through T4 be unchanged, except that a new variable `Q` is initialized to Λ in step T1; `Q` will point to the last node visited, if any. Step T5 becomes two steps: "**T5.** [Right branch done?] If `RLINK(P)` $= \Lambda$ or `RLINK(P)` $=$ `Q`, go on to T6; otherwise set `A` $\Leftarrow$ `P`, `P` $\leftarrow$ `RLINK(P)` and return to T2. **T6.** [Visit `P`.] "Visit" `NODE(P)`, set `Q` $\leftarrow$ `P`, and return to T4." A similar proof applies. (Steps T4 and T5 can be streamlined so that nodes are not taken off the stack and immediately reinserted.)

Page 562 replacement for answer 21 ——— 23 Feb 1995

21. The following solution by Joseph M. Morris [*Inf. Proc. Letters* **9** (1979), 197–200] traverses also in preorder (cf. exercise 18).

U1. [Initialize.] Set `P` $\leftarrow$ `T` and `R` $\leftarrow \Lambda$.

U2. [Done?] If `P` $= \Lambda$, the algorithm terminates.

U3. [Look left.] Set `Q` $\leftarrow$ `LLINK(P)`. If `Q` $= \Lambda$, visit `NODE(P)` in preorder and go to U6.

U4. [Search for thread.] Set `Q` $\leftarrow$ `RLINK(Q)` zero or more times until either `Q` $=$ `R` or `RLINK(Q)` $= \Lambda$.

U5. [Insert or remove thread.] If `Q` $\neq$ `R`, set `RLINK(Q)` $\leftarrow$ `P` and go to U8. Otherwise set `RLINK(Q)` $\leftarrow \Lambda$ (it had been changed temporarily to `P`, but we've now traversed `P`'s left subtree).

U6. [Inorder visit.] Visit `NODE(P)` in inorder.

U7. [Go to right or up.] Set `R` $\leftarrow$ `P`, `P` $\leftarrow$ `RLINK(P)`, and return to U2.

U8. [Preorder visit.] Visit `NODE(P)` in preorder.

U9. [Go to left.] Set `P` $\leftarrow$ `LLINK(P)` and return to step U3. ▌

Morris also suggested a slightly more complicated way to traverse in postorder.

A completely different solution was found by J. M. Robson [*Inf. Proc. Letters* **2** (1973), 12–14]. Let's say that a node is "full" if its `LLINK` and `RLINK` are nonnull, "empty" if its `LLINK` and `RLINK` are both empty. Robson found a way to maintain a stack of pointers to the full nodes whose right subtrees are being visited, using the link fields in empty nodes!

Yet another way to avoid an auxiliary stack was discovered independently by G. Lindstrom and B. Dwyer, *Inf. Proc. Letters* **2** (1973), 47–51, 143–145. Their algorithm traverses in *triple order* — it visits every node exactly three times, once in each of preorder, inorder, and postorder — but it does not know which of the three is currently being done.

W1. [Initialize.] Set `P` ← `T` and `Q` ← `S`, where `S` is a sentinel value — any link that is known to be different from any link in the tree (e.g., -1).

W2. [Bypass null.] If `P` $= \Lambda$, set `P` ← `Q` and `Q` ← Λ.

W3. [Done?] If `P` = `S`, terminate the algorithm. (We will have `Q` = `T` at termination.)

W4. [Visit.] Visit `NODE(P)`.

W5. [Rotate.] Set `R` ← `LLINK(P)`, `LLINK(P)` ← `RLINK(P)`, `RLINK(P)` ← `Q`, `Q` ← `P`, `P` ← `R`, and return to W2. ▌

Correctness follows from the fact that if we start at W2 with `P` pointing to the root of a binary tree `T` and `Q` pointing to `X`, where `X` is not a link in that tree, the algorithm will triple traverse the tree and reach step W3 with `P` = `X` and `Q` = `T`.

If $\alpha(\mathtt{T}) = x_1x_2\ldots x_{3n}$ is the resulting sequence of nodes in triple order, we have $\alpha(\mathtt{T}) = \mathtt{T}\,\alpha(\mathtt{LLINK(T)})\,\mathtt{T}\,\alpha(\mathtt{RLINK(T)})\,\mathtt{T}$. Therefore, as Lindstrom observed, the three subsequences $x_1x_4\ldots x_{3n-2}$, $x_2x_5\ldots x_{3n-1}$, $x_3x_6\ldots x_{3n}$ each include every tree node just once. (Since x_{j+1} is either the parent or child of x_j, these subsequences visit the nodes in such a way that each is at most three links away from its predecessor. Section 7.2.1 describes a general traversal scheme called *prepostorder* that has this property not only for binary trees but for trees in general.)

Page 562 replacement for answer 22 (extends to page 563) ———— 23 Feb 1995

22. This program uses the conventions of Programs T and S, with `Q` in rI6 and/or rI4. The old-fashioned `MIX` computer is not good at comparing index registers for equality, so variable `R` is omitted and the test '`Q` = `R`' is changed to '`RLINK(Q)` = `P`'.

```
U1    LD5  HEAD(LLINK)        1          U1. Initialize. P ← T.
U2A   J5Z  DONE               1          Stop if P = Λ.
U3    LD6  0,5(LLINK)       n+a−1        U3. Look left. Q ← LLINK(P).
      J6Z  U6               n+a−1        To U6 if Q = Λ.
U4    CMP5 1,6(RLINK)       2n−2b        U4. Search for thread.
      JE   5F               2n−2b        Jump if RLINK(Q) = P.
      ENT4 0,6            2n−2b−a+1      rI4 ← Q.
      LD6  1,6(RLINK)     2n−2b−a+1
      J6NZ U4             2n−2b−a+1      To U4 with Q ← RLINK(Q) if it's ≠ 0.
U5    ST5  1,4(RLINK)        a−1         U5a. Insert thread. RLINK(Q) ← P.
U9    LD5  0,5(LLINK)        a−1         U9. Go to left. P ← LLINK(P).
      JMP  U3                a−1         To U3.
5H    STZ  1,6(RLINK)        a−1         U5b. Remove thread. RLINK(Q) ← Λ.
U6    JMP  VISIT              n          U6. Inorder visit.
U7    LD5  1,5(RLINK)         n          U7. Go to right or up. P ← RLINK(P).
U2    J5NZ U3                 n          U2. Done? To U3 if P ≠ Λ.
DONE  ...
```

The total running time is $21n + 6a - 3 - 14b$, where n is the number of nodes, a is the number of null `RLINK`s (hence $a - 1$ is the number of nonnull `LLINK`s), and b is the number of nodes on the tree's "right spine" `T`, `RLINK(T)`, `RLINK(RLINK(T))`, etc.

▸**Page 567** answer 12 line 2 ——— 02 May 1976

`INFO(P2)` $- 1$ $\rightsquigarrow$ `TREE(INFO(P2)` $- 1$`)`

▸**Page 567** the missing `MIX` program on bottom four lines ——— 08 Jun 1980

```
ST3  6F(0:2)
ST2  7F(0:2)
ENT2 8F
JMP  1F
```

▸**Page 568** program line *86* ——— 08 Jun 1980

`0,2` $\rightsquigarrow$ `0,2(RLINKT)`

Page 568 improvements to program lines *93* to *100* ——— 08 Jun 1980

```
 93  C4 LDA  0,1(LLINK)    C4. Anything to left?
 94     JANZ 4B            Jump if LLINK(P) ≠ Λ.
 95     STZ  0,2(LLINK)    LLINK(Q) ← Λ.
 96  C5 LD2N 0,2(RLINKT)   C5. Advance. Q ← −RLINKT(Q).
 97     LD1  0,1(RLINK)    P ← RLINK(P).
 98     J2P  C5            Jump if RTAG(Q) was "−".
 99     ENN2 0,2           Q ← −Q.
100  C6 J2NZ C2            C6. Test if complete.
```

Page 568 lines 3 and 4 of answer 14 ——— 08 Jun 1980

89–95, ... $18u$); $\rightsquigarrow$ 89–94, n; 95, $n - a$; 96–98, $n + 1$; 99–100, $n - a$; 101–103, 1. The total time is $(36n + 22)u$;

Page 570 new answer ——— 01 Jul 1995

18. First set `LLINK`$[j]$ $\leftarrow$ `RLINK`$[j]$ $\leftarrow j$ for all j, so that each node is in a circular list of length 1. Then for $j = n, n - 1, \ldots, 1$ (in this order), if `FATHER`$[j] = 0$ set $r \leftarrow j$, otherwise insert the circular list starting with j into the circular list starting with `FATHER`$[j]$ as follows: $k \leftarrow$ `FATHER`$[j]$, $l \leftarrow$ `RLINK`$[k]$, $i \leftarrow$ `LLINK`$[j]$, `LLINK`$[j] \leftarrow k$, `RLINK`$[k] \leftarrow j$, `LLINK`$[l] \leftarrow i$, `RLINK`$[i] \leftarrow l$. This works because (a) each nonroot node is always preceded by its parent or by a descendant of its parent; (b) nodes of each family appear in their parent's list, in order of location; (c) preorder is the unique order satisfying (a) and (b).

▸**Page 573** exercise 18 line 5 ——— 21 Mar 1976

preorder $\rightsquigarrow$ postorder

▸**Page 574** answer 7 ——— 02 May 1976

The diagrams for Case 1 have two arrowheads in the wrong direction; the arrows should lead away from α' and towards β', both Before and After.
On the final line of the answer, 313–332 $\rightsquigarrow$ 313–322

▸**Page 575** answer 12 line 5 ——— 03 Mar 1975

$a(i)$ set $a(i) \leftarrow c(i,j)$ and $b(i) \leftarrow j$; $\rightsquigarrow$ $a(j)$ set $a(j) \leftarrow c(i,j)$ and $b(j) \leftarrow i$
$a(i) \leftarrow \infty$. $\rightsquigarrow$ $a(i) \leftarrow \infty$. Here $c(i,j)$ means $c(j,i)$ when $j < i$.

Page 575 at end of answer 12 . . . . 23 Mar 1995

Eventually I'll refer here to Section 7.5.4, where minimum spanning trees are discussed in depth.

▸**Page 577** last line of answer 14 ——— 21 Mar 1997

e_{21} or e_{22} $\rightsquigarrow$ e_{20} or e_{22}

▸**Page 578** answer 19 ——— 15 Mar 1987

line 1: $a_{ij} = 1 \rightsquigarrow a_{jj} = 1$
lines 14–16: deleted. ... Then $\det A_0 = \det B_0 + \det C_0$, $\rightsquigarrow$ deleted. *Example:* If $A_0 = \begin{pmatrix} 3 & -2 \\ -1 & 2 \end{pmatrix}$, $j = 1$, and e is an arc from V_1 to V_0, then $B_0 = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$, $C_0 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$. In general we have $\det A_0 = \det B_0 + \det C_0$,

Page 578 last three lines of answer 19 ——— 03 Jul 1995

This important ... 56]. $\rightsquigarrow$ *Notes:* If we delete any set S of rows from the matrix A, and the same set of columns, the determinant of the resulting matrix is the number of oriented forests whose roots are the vertices $\{V_k \mid k \in S\}$ and whose arcs belong to the given digraph. The matrix tree theorem for oriented trees is due to W. T. Tutte [*Proc. Cambridge Phil. Soc.* **44** (1948), 463–482, §3], who generalized a classic theorem of Borchardt [*Crelle* **57** (1860), 111–121]. Borchardt's theorem covered the case of free trees in *undirected* graphs, when the matrix A is symmetric. Several authors have ascribed Borchardt's theorem to Kirchhoff, but Kirchhoff proved a quite different (though related) result. Borchardt's work followed that of Sylvester, who came close to proving a theorem even more general than Tutte's; see exercise 28.

Page 578 bottom line ——— 14 May 1976

D. E. Knuth, $\rightsquigarrow$ R. Dawson and I. J. Good, *Ann. Math. Stat.* **28** (1957), 946–956; D. E. Knuth,

▸**Page 579** in the biggest matrix ——— 01 May 1979

change the label on row 3 and the label on column 3 from [10] to [20]

▸**Page 579** in the second-biggest matrix ——— 01 May 1979

$a_{0m} \rightsquigarrow a_{0n}$

Page 579 new paragraph for end of answer 21 ——— 03 Jul 1995

An alternative, purely combinatorial proof of this result has been given by J. B. Orlin, *J. Combinatorial Theory* **B25** (1978), 187–198.

Page 580 addition to answer 23 ——— 14 May 1976

[For $m = 2$ this result is due to C. Flye Sainte-Marie, L'*Intermédiaire des Mathématiciens* **1** (1894), 107–110.]

▸**Page 580** line −9 ——— 22 Mar 1997

$f \quad j \quad + \quad h \quad - \rightsquigarrow f \quad j \quad + \quad d \quad -$

Page 581 new answer before Section 2.3.4.3 ——— 19 May 1981

27. Let a_{ij} be the sum of $p(e)$ over all arcs e from V_i to V_j. We are to prove that $t_j = \sum_i a_{ij} t_i$ for all j. Since $\sum_i a_{ji} = 1$, we must prove that $\sum_i a_{ji} t_j = \sum_i a_{ij} t_i$. But this is not difficult, because both sides of the equation represent the sum of all products $p(e_1) \ldots p(e_n)$ taken over subgraphs $\{e_1, \ldots, e_n\}$ of G such that $\operatorname{init}(e_i) = V_i$

and such that there is a unique oriented cycle contained in $\{e_1, \ldots, e_n\}$, where this cycle includes V_j. Removing any arc of the cycle yields an oriented tree; the left-hand side of the equation is obtained by factoring out the arcs that leave V_j, while the right-hand side corresponds to those that enter V_j.

In a sense, this exercise is combination of exercises 19 and 26.

Page 581 another new answer before Section 2.3.4.3 ——————— 09 Dec 1991

28. Every term in the expansion is $a_{1p_1} \ldots a_{mp_m} b_{1q_1} \ldots b_{nq_n}$, where $0 \le p_i \le n$ for $1 \le i \le m$ and $0 \le q_j \le m$ for $1 \le j \le n$, times some integer coefficient. Represent this product as a directed graph on the vertices $\{0, u_1, \ldots, u_m, v_1, \ldots, v_n\}$, with arcs from u_i to v_{p_i} and from v_j to u_{q_j}, where $u_0 = v_0 = 0$.

If the digraph contains a cycle, the integer coefficient is zero. For each cycle corresponds to a factor of the form

$$a_{i_0 j_0}\, b_{j_0 i_1}\, a_{i_1 j_1}\ \ldots\ a_{i_{k-1} j_{k-1}}\, b_{j_{k-1} i_0} \qquad (*)$$

where the indices $(i_0, i_1, \ldots, i_{k-1})$ are distinct and so are the indices $(j_0, j_1, \ldots, j_{k-1})$. The sum of all terms containing $(*)$ as a factor is the determinant obtained by setting $a_{i_l j} \leftarrow [j = j_l]$ for $0 \le j \le n$ and $b_{j_l i} \leftarrow [i = i_{(l+1) \bmod k}]$ for $0 \le i \le m$, for $0 \le l < k$, leaving the variables in the other $m + n - 2k$ rows unchanged. This determinant is identically zero, because the sum of rows $i_0, i_1, \ldots, i_{k-1}$ in the top section equals the sum of rows $j_0, j_1, \ldots, j_{k-1}$ in the bottom section.

On the other hand, if the directed graph contains no cycles, the integer coefficient is $+1$. This follows because each factor a_{ip_i} and b_{jq_j} must have come from the diagonal of the determinant: If any off-diagonal element $a_{i_0 j_0}$ is chosen in row i_0 of the top section, we must choose some off-diagonal $b_{j_0 i_1}$ from column j_0 of the left section, hence we must choose some off-diagonal $a_{i_1 j_1}$ from row i_1 of the top section, etc., forcing a cycle.

Thus the coefficient is $+1$ iff the corresponding digraph is an oriented tree with root 0. The number of such terms (hence the number of such oriented trees) is obtained by setting each a_{ij} and b_{ji} to 1; for example,

$$\det\begin{pmatrix} 4 & 0 & 1 & 1 & 1 \\ 0 & 4 & 1 & 1 & 1 \\ 1 & 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 0 & 3 \end{pmatrix} = \det\begin{pmatrix} 4 & 0 & 1 & 1 & 1 \\ -4 & 4 & 0 & 0 & 0 \\ 1 & 1 & 3 & 0 & 0 \\ 0 & 0 & -3 & 3 & 0 \\ 0 & 0 & -3 & 0 & 3 \end{pmatrix} = \det\begin{pmatrix} 4 & 0 & 3 & 1 & 1 \\ 0 & 4 & 0 & 0 & 0 \\ 2 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

$$= \det\begin{pmatrix} 4 & 3 \\ 2 & 3 \end{pmatrix} \cdot 4 \cdot 3 \cdot 3\,.$$

In general we obtain $\det\left(\begin{smallmatrix} n+1 & n \\ m & m+1 \end{smallmatrix}\right) \cdot (n+1)^{m-1} \cdot (m+1)^{n-1}$.

Note: J. J. Sylvester considered the special case $m = n$ and $a_{10} = a_{20} = \cdots = a_{m0} = 0$ in *Quarterly J. of Pure and Applied Math.* **1** (1857), 42–56, where he conjectured (correctly) that the total number of terms is then $n^n (n+1)^{n-1}$. He also stated without proof that the $(n+1)^{n-1}$ nonzero terms present when $a_{ij} = \delta_{ij}$ correspond to all connected cycle-free graphs on $\{0, 1, \ldots, n\}$. In that special case, he reduced the determinant to the form in the matrix tree theorem of exercise 19, e.g.,

$$\det\begin{pmatrix} b_{10} + b_{12} + b_{13} & -b_{12} & -b_{13} \\ -b_{21} & b_{20} + b_{21} + b_{23} & -b_{23} \\ -b_{31} & -b_{32} & b_{30} + b_{31} + b_{32} \end{pmatrix}.$$

Cayley quoted this result in *Crelle* **52** (1856), 279, ascribing it to Sylvester; thus it is ironic that the theorem about the number of such graphs is often attributed to Cayley.

By negating the first m rows of the given determinant, then negating the first m columns, we can reduce this exercise to the matrix tree theorem.

Page 581 addendum to the (new) answer 28 ———— 14 Dec 1996

[Matrices having the general form considered in this exercise are important in iterative methods for the solution of partial differential equations, and they are said to have "Property $\mathcal{A}$." See, for example, Louis A. Hageman and David M. Young, *Applied Iterative Methods* (Academic Press, 1981), Chapter 9.]

▶**Page 581** answer 3 line 3 ———— 02 May 1976

upper half plane $\rightsquigarrow$ right half plane

Page 582 near the top ———— 14 Apr 1995

line 3: exhibits a set of *six* polygonal shapes which $\rightsquigarrow$ exhibited a set of *six* squarish shapes that

line 4: allowed. $\rightsquigarrow$ allowed. In 1974, Roger Penrose discovered a set of only *two* polygons, based on the golden ratio instead of a square grid, that tile the plane only aperiodically [see Grünbaum and Shephard, *Tilings and Patterns* (Freeman, 1987), Chapter 10; Martin Gardner, *Penrose Tiles to Trapdoor Ciphers* (Freeman, 1989), Chapters 1–2].

Page 582 line -9 ———— 01 Mar 1979

Note: Kruskal's $\rightsquigarrow$ *Note:* Kruskal actually proved a stronger result, using a weaker form of embedding. His

Page 582 line -6 ———— 25 Mar 1981

305. $\rightsquigarrow$ 305. See also N. Dershowitz, *Inf. Proc. Letters* **9** (1979), 212–215, for applications to termination of algorithms.

Page 583 answer 5 ———— 05 Jan 1979

There is ... exist. $\rightsquigarrow$ When $n > 1$, the number of series-parallel networks with n edges is $2c_n$ [see P. A. MacMahon, *Proc. London Math. Soc.* **22** (1891), 330–339].

Page 583 replacement for second line of answer 6 ———— 14 Apr 1995

$23z^7 + 46z^8 + 98z^9 + \cdots$. The function $F(z) = 1 - zG(z)$ satisfies the simpler relation $F(z^2) = 2z + F(z)^2$. [J. H. M. Wedderburn, *Annals of Math.* **24** (1922), 121–140.]

▶**Page 584** answer 10 ———— 02 May 1976

height $\rightsquigarrow$ weight (three times)

Page 584 new paragraph for end of answer 10 ———— 04 Jul 1995

[It is a nice programming exercise to compute $s(X, Y)$ for all adjacent X and Y in $O(n)$ steps; from this information we can quickly find the centroid(s). An efficient algorithm for centroid location was first given by A. J. Goldman, *Transportation Sci.* **5** (1971), 212–221.]

Page 585 the line before answer 18 ———— 26 Apr 1995

3.1–14.) $\rightsquigarrow$ 3.1–14. The formula $m^{m-1}Q(m)$ was first obtained by L. Katz, *Annals of Math. Statistics* **26** (1955), 512–517.)

Page 587 new sentence to follow answer 29 ——— 04 May 1995

[Exercise 4.7–22 derives considerably more general results.]

Page 587 line 2 of answer 31 ——— 04 Oct 1996

$\frac{1}{4}(1+z-\sqrt{1-6z+z^2})$. $\rightsquigarrow$ $\frac{1}{4}(1+z-\sqrt{1-6z+z^2}\,) = z+z^2+3z^3+11z^4+45z^5+\cdots$.

Page 587 line 3 of answer 31 ——— 14 Apr 1995

Zeitschrift für Mathematik $\rightsquigarrow$ *Zeitschrift für Mathematik und Physik*

Page 587 line 7 of answer 31 ——— 04 Oct 1996

12.] $\rightsquigarrow$ 12. Curiously, the value $[z^{10}]\,G(z) = 103049$ seems to have been calculated already by Hipparchus in the second century B.C., as the number of "affirmative compound propositions that can be made from only ten simple propositions"; see R. P. Stanley, *AMM* **104** (1997), to appear.]

Page 588 fourth line before answer 33 ——— 27 May 1978

minimal. $\rightsquigarrow$ minimal. [This argument in the case of binary trees was apparently first discovered by C. S. Peirce in an unpublished manuscript; see his *New Elements of Mathematics* **4** (The Hague: Mouton, 1976), 303–304. It was discovered in the case of t-ary trees by Dvoretzky and Motzkin, *Duke Math. J.* **14** (1947), 305–313.]

Page 588 third line before answer 33 ——— 16 Mar 1981

Either of the methods $\rightsquigarrow$ Still another proof, by G. Bergman, inductively replaces d_kd_{k+1} by $(d_k+d_{k+1}-1)$ if $d_k>0$ [*Algebra Universalis* **8** (1978), 129–130].
The methods

▶**Page 589** line 1 of answer 4 ——— 18 Oct 1979

$l_j > l_{j+1}$ $\rightsquigarrow$ $l_j \ge l_{j+1}$

Page 590 addendum to answer 10 ——— 18 Oct 1979

(place the figure at the right margin and set the copy narrower, to its left)
The desired ternary is $\rightsquigarrow$
The desired ternary tree is shown at the right. F. K. Hwang has observed [*SIAM J. Appl. Math.* **37** (1979), 124–127] that a similar procedure is valid for minimum weighted path length trees having any prescribed multiset of degrees: At each step the smallest t weights are combined, where t is as small as possible.

▶**Page 590** replacement for answer 12 ——— 14 Dec 1983

12. By exercise 9, it is the internal path length divided by n, plus 1. (This result holds for general trees as well.)

Page 590 new answers ——— 18 Oct 1979

13. [Cf. J. van Leeuwen, *Proc. 3rd International Colloq. Automata, Languages and Programming* (Edinburgh University Press, 1976), 382–410.]

H1. [Initialize.] Set $A[m-1+i] \leftarrow w_i$ for $1 \le i \le m$. Then set $A[2m] \leftarrow \infty$, $x \leftarrow m$, $i \leftarrow m+1$, $j \leftarrow m-1$, $k \leftarrow m$. (During this algorithm $A[i] \le \cdots \le$

$A[2m-1]$ is the queue of unused external weights; $A[k] \ge \cdots \ge A[j]$ is the queue of unused internal weights, empty if $j < k$; the current left and right pointers are x and y.)

H2. [Find right pointer.] If $j < k$ or $A[i] \le A[j]$, set $y \leftarrow i$ and $i \leftarrow i+1$; otherwise set $y \leftarrow j$ and $j \leftarrow j-1$.

H3. [Create internal node.] Set $k \leftarrow k-1$, $L[k] \leftarrow x$, $R[k] \leftarrow y$, $A[k] \leftarrow A[x]+A[y]$.

H4. [Done?] Terminate the algorithm if $k = 1$.

H5. [Find left pointer.] (At this point $j \ge k$ and the queues contain a total of k unused weights. If $A[y] < 0$ we have $j = k$, $i = y+1$, and $A[i] > A[j]$.) If $A[i] \le A[j]$, set $x \leftarrow i$ and $i \leftarrow i+1$; otherwise set $x \leftarrow j$ and $j \leftarrow j-1$. Return to step H2. ▮

14. The proof for $k = m-1$ applies with little change. [Cf. *SIAM J. Appl. Math.* **21** (1971), 518.]

15. Use the combined-weight functions (a) $1+\max(w_1, w_2)$ and (b) $xw_1 + xw_2$, respectively, instead of $w_1 + w_2$ in (9). [Part (a) is due to M. C. Golumbic, *IEEE Trans.* **C-25** (1976), 1164–1167; part (b) to T. C. Hu, D. Kleitman, and J. K. Tamaki, *SIAM J. Appl. Math.* **37** (1979), 246–256. Huffman's problem is the limiting case of (b) as $x \to 1$, since $\sum(1+\epsilon)^{l_j} w_j = \sum w_j + \epsilon \sum w_j l_j + O(\epsilon^2)$.]

D. Stott Parker, Jr., has pointed out that a Huffman-like algorithm will also find the minimum of $w_1 x^{l_1} + \cdots + w_m x^{l_m}$ when $0 < x < 1$, if the two *maximum* weights are combined at each step. In particular, the minimum of $w_1 2^{-l_1} + \cdots + w_m 2^{-l_m}$, when $w_1 \le \cdots \le w_m$, is $w_1/2 + \cdots + w_{m-1}/2^{m-1} + w_m/2^{m-1}$. See D. E. Knuth, *J. Comb. Theory* **A32** (1982), 216–224, for further generalizations.

16. Let $l_{m+1} = l'_{m+1} = 0$. Then

$$\sum_{j=1}^{m} w_j l_j \le \sum_{j=1}^{m} w_j l'_j = \sum_{k=1}^{m} (l'_j - l'_{j+1}) \sum_{j=1}^{k} w_j$$
$$\le \sum_{k=1}^{m} (l'_j - l'_{j+1}) \sum_{j=1}^{k} w'_j = \sum_{j=1}^{m} w'_j l'_j\,,$$

since $l'_j \ge l'_{j+1}$ as in exercise 4. The same proof holds for many other kinds of optimum trees, including those of exercise 10.

17. (a) This is exercise 14. (b) We can extend $f(n)$ to a concave function $f(x)$, so the stated inequality holds. Now $F(m)$ is the minimum of $\sum_{j=1}^{m-1} f(s_j)$, where the s_j are internal node weights of an extended binary tree on the weights 1, 1, ..., 1. Huffman's algorithm, which constructs the complete binary tree with $m-1$ internal nodes in this case, yields the optimum tree. The choice $k = 2^{\lceil (n/3) \rceil}$ defines a binary tree with the same internal weights, so it yields the minimum in the recurrence, for each n. [*SIAM J. Appl. Math.* **31** (1976), 368–378.] We can evaluate $F(n)$ in $O(\log n)$ steps; cf. exercises 5.2.3–20 and 21. If $f(n)$ is convex instead of concave, so that $\Delta^2 f(n) \ge 0$, the solution to the recurrence is obtained when $k = \lfloor n/2 \rfloor$.

Page 592 line 1 of answer 5 ______ 04 Jul 1995

algorithm ⤳ algorithm in a one-level memory

Page 593 additional line for answer 9 ______ 04 Jul 1995

See also exercise 2.5–33.

Page 594 replacement for lines 19–21 ———— 14 May 1976

Several elegant List-moving and List-copying algorithms that make substantially weaker assumptions about List representation have been devised. See D. W. Clark, *CACM* **19** (1976), 352–354; J. M. Robson, *CACM* **20** (1977), 431–433.

Page 594 bottom line ———— 14 May 1978

exhausted. $\rightsquigarrow$ exhausted. [For further information, see Guy L. Steele Jr., *CACM* **18** (1975), 495–508; P. Wadler, *CACM* **19** (1976), 491–500; E. W. Dijkstra, L. Lamport, A. J. Martin, C. S. Scholten, and E. F. M. Steffens, *CACM* **21** (1978), 966–975; H. G. Baker, Jr., *CACM* **21** (1978), 280–294.

▶**Page 596** line 5 ———— 08 Apr 1975

miniscule $\rightsquigarrow$ minuscule

Page 597 last line of answer 6 ———— 14 Apr 1995

list. $\rightsquigarrow$ list. [Sometimes, however, it is useful to have small blocks at the beginning, as in the strict first-fit method; for example, we might want to keep a sequential stack at the high end of memory. In such cases we can reduce the search time by using trees as suggested in exercise 6.2.3–30.]

Page 597 answer 8 ———— 25 Jun 1976

line 1: also set $\texttt{R} \rightsquigarrow$ also set $\texttt{M} \leftarrow \infty$, $\texttt{R}$
line 3: If $\texttt{R} = \Lambda$ or $\texttt{M} \rightsquigarrow$ If $\texttt{M}$

▶**Page 600** replacement for lines 3–5 of answer 19 ———— 29 Jun 1982

new step A2.5 between steps A2 and A3: "Set $\texttt{P1} \leftarrow \texttt{P} + \texttt{SIZE(P)}$. If $\texttt{TAG(P1)} =$ "+", proceed to step A3. Otherwise set $\texttt{P2} \leftarrow \texttt{LINK(P1)}$, $\texttt{LINK(P2+1)} \leftarrow \texttt{LINK(P1 + 1)}$, $\texttt{LINK(LINK(P1 + 1))} \leftarrow \texttt{P2}$, $\texttt{SIZE(P)} \leftarrow \texttt{SIZE(P)} + \texttt{SIZE(P1)}$. If $\texttt{ROVER} = \texttt{P1}$, set $\texttt{ROVER} \leftarrow \texttt{P2}$. Repeat step A2.5."

Page 600 new paragraph to follow answer 19 ———— 10 Jul 1995

See *TEX: The Program* (Addison-Wesley, 1986), §125, for several refinements.

▶**Page 601** answer 26 ———— 05 Dec 1977

power of two. $\rightsquigarrow$ power of two, with blocks in decreasing order of size.
$\texttt{P} \geq \texttt{M} \rightsquigarrow \texttt{P} \geq \texttt{M} - 2^k$.

Page 602 new answer ———— 28 Feb 1978

31. See David L. Russell, *SICOMP* **6** (1977), 607–621.

Page 603 line 5 of step G8 ———— 04 Jul 1995

$\texttt{P} < \texttt{Q} \leq \texttt{P} + \texttt{T(P)} \rightsquigarrow \texttt{P} < \texttt{Q} \leq \texttt{P} + \texttt{T(P)}, \texttt{LINK(Q)} \neq \Lambda$

Page 603 replacement for lines 22–24 ———— 04 Jul 1995

[This method is called the "LISP 2 garbage collector." An interesting alternative, which does not require the LINK field at the beginning of a node, can be based on the idea of linking together all pointers that point to each node — see Lars-Erik Thorelli, *BIT* **16** (1976), 426–441; R. B. K. Dewar and A. P. McCann, *Software Practice & Exp.* **7** (1977), 95–113; F. Lockwood Morris, *CACM* **21** (1978), 662–665, **22** (1979), 571; H. B. M. Jonkers, *Inf. Proc. Letters* **9** (1979), 26–30; J. J. Martin, *CACM* **25** (1982), 571–581; F. Lockwood Morris, *Inf. Proc. Letters* **15** (1982), 139–142, **16** (1983), 215.

Other methods have been published by B. K. Haddon and W. M. Waite, *Comp. J.* **10** (1967), 162–165; B. Wegbreit, *Comp. J.* **15** (1972), 204–208; D. A. Zave, *Inf. Proc. Letters* **3** (1975), 167–169. Cohen and Nicolau have analyzed four of these approaches in *ACM Trans. Prog. Languages and Systems* **5** (1983), 532–553.]

Page 606 line 4 ———— 12 Mar 1984

$1.22 \ln n \rightsquigarrow 1.1825 \ln n$

▶**Page 606** line 16 ———— 14 Jun 1977

$4\frac{19559}{32768} \rightsquigarrow 4\frac{18535}{32768}$

Page 606 new answers ———— 01 Apr 1981

42. We can assume that $m \geq 6$. The main idea is to establish the occupancy pattern $R_{m-2}(F_{m-3}R_1)^k$ at the beginning of the memory, for $k = 0, 1, \ldots$, where R_j and F_j denote reserved and free blocks of size j. The transition from k to $k+1$ begins with

$$\begin{aligned} R_{m-2}(F_{m-3}R_1)^k &\to R_{m-2}(F_{m-3}R_1)^k R_{m-2}R_{m-2} \\ &\to R_{m-2}(F_{m-3}R_1)^{k-1} F_{2m-4}R_{m-2} \\ &\to R_{m-2}(F_{m-3}R_1)^{k-1} R_m R_{m-5} R_1 R_{m-2} \\ &\to R_{m-2}(F_{m-3}R_1)^{k-1} F_m R_{m-5} R_1\,; \end{aligned}$$

then the commutation sequence $F_{m-3}R_1F_mR_{m-5}R_1 \to F_{m-3}R_1R_{m-2}R_2R_{m-5}R_1 \to F_{2m-4}R_2R_{m-5}R_1 \to R_mR_{m-5}R_1R_2R_{m-5}R_1 \to F_mR_{m-5}R_1F_{m-3}R_1$ is used k times until we get $F_mR_{m-5}R_1(F_{m-3}R_1)^k \to F_{2m-5}R_1(F_{m-3}R_1)^k \to R_{m-2}(F_{m-3}R_1)^{k+1}$. Finally when k gets large enough there is an endgame that forces overflow unless the memory size is at least $(n-4m+11)(m-2)$; details appear in *Comp. J.* **20** (1977), 242–244. [Notice that the worst conceivable worst case, which begins with the pattern $F_{m-1}R_1F_{m-1}R_1F_{m-1}R_1\ldots$, is only slightly worse than this; the next-fit strategy of exercise 6 can produce this pattern.]

43. We will show that if $D_1, D_2, \ldots$ is any sequence of numbers such that $D_1/m + D_2/(m+1) + \cdots + D_m/(2m-1) \geq 1$ for all $m \geq 1$, and if $C_m = D_1/1 + D_2/2 + \cdots + D_m/m$, then $N_{\mathrm{FF}}(n,m) \leq nC_m$. In particular, since

$$\frac{1}{m} + \frac{1}{m+1} + \cdots + \frac{1}{2m+1} = 1 - \frac{1}{2} + \cdots + \frac{1}{2m-3} - \frac{1}{2m-2} + \frac{1}{2m-1} > \ln 2,$$

the constant sequence $D_m = 1/\ln 2$ satisfies the necessary conditions. The proof is by induction on m. Let $N_j = nC_j$ for $j \geq 1$, and suppose that some request for a block of size m cannot be allocated in the leftmost N_m cells of memory. Then $m > 1$. For $0 \leq j < m$, we let N'_j denote the rightmost position allocated to blocks of sizes $\leq j$, or 0 if all reserved blocks are larger than j; by induction we have $N'_j \leq N_j$. Furthermore we let N'_m be the rightmost occupied position $\leq N_m$, so that $N'_m \geq N_m - m + 1$. Then the interval $(N'_{j-1} \mathinner{.\,.} N'_j\,]$ contains at least $\lceil j(N'_j - N'_{j-1})/(m+j-1)\rceil$ occupied cells, since its free blocks are of size $< m$ and its reserved blocks are of size $\geq j$. It follows that $n - m \geq$ number of occupied cells $\geq \sum_{j=1}^{m} j(N'_j - N'_{j-1})/(m+j-1) = mN'_m/(2m-1) - (m-1)\sum_{j=1}^{m-1} N'_j/(m+j)(m+j-1) > mN_m/(2m-1) - m - (m-1)\sum_{j=1}^{m-1} N_j\bigl(1/(m+j-1) - 1/(m+j)\bigr) = \sum_{j=1}^{m} nD_j/(m+j-1) - m \geq n - m$, a contradiction.

[This proof establishes slightly more than was asked. If we define the D's by $D_1/m + \cdots + D_m/(2m-1) = 1$, then the sequence $C_1, C_2, \ldots$ is $1, \frac{7}{4}, \frac{161}{72}, \frac{7483}{2880}, \ldots$; and the result can be improved further, even in the case $m = 2$, cf. exercise 38.]

Page 606 new answer —— 06 Jul 1995

44. $\lceil F^{-1}(1/N)\rceil$, $\lceil F^{-1}(2/N)\rceil$, $\ldots$, $\lceil F^{-1}(N/N)\rceil$.

Page 608 replacing the line for Kronecker delta —— 11 Feb 1995

$[B]$	characteristic function of condition B: $(B \Longrightarrow 1;\ 0)$	1.2.3
δ_{jk}	Kronecker delta: $[j=k]$	1.2.6
$[z^n]\,g(z)$	coefficient of z^n in power series $g(z)$	1.2.9
$j \perp k$	j is relatively prime to k: $\gcd(j,k)=1$	1.2.4

Page 609 in several places —— 14 Apr 1995

x upper k: $\leadsto$ x to the k rising: $\Gamma(x+k)/\Gamma(x) =$
x lower k: $\leadsto$ x to the k falling: $x!/(x-k)! =$
n factorial: $\leadsto$ n factorial: $\Gamma(n+1) =$
$\|S\| \leadsto |S|$

Page 610 omit notations rarely used —— 11 Feb 1995

delete the definition of the Beta function $\mathrm{B}(x,y)$ and the partition function $p(n)$

Page 610 near the middle —— 14 Apr 1995

gamma function ... positive integer $\leadsto$ gamma function: $(x-1)! = \gamma(x,\infty)$
incomplete gamma function $\leadsto$ incomplete gamma function: $\int_0^y e^{-t}t^{x-1}\,dt$

Page 610 the missing Section reference for $x \approx y$ —— 14 Apr 1995

1.2.5, 4.2.2

Page 611 after the entry for $\mathrm{var}(g)$ —— 03 Aug 1995

$\Pr\bigl(S(X)\bigr)$	probability that statement $S(X)$ is true, for random values of X	1.2.10
$\mathrm{E}\,X$	expected value of X: $\sum_x x\Pr(X=x)$	1.2.10

▸**Page 614** line -20 (the entry for 1 degree of arc) —— 19 Jun 1975

1154 $\leadsto$ *1155*

Page 615 replacement for lines 1–7 —— 01 Nov 1995

Several of the 40-digit values in Tables 1 and 2 were computed on a desk calculator by John W. Wrench, Jr., for the first edition of this book. See the answer to exercise 1.3.3–23 for the 40-digit value of another fundamental constant.

Page 615 more lines for Table 3 —— 14 Apr 1995

26	34395742267/8923714800	8553103/6	121393	26
27	312536252003/80313433200	0	196418	27
28	315404588903/80313433200	−23749461029/870	317811	28
29	9227046511387/2329089562800	0	514229	29
30	9304682830147/2329089562800	8615841276005/14322	832040	30

Page 616 lines 7–10 ______ 01 Nov 1995

$$H_{1/5} = 5 - \tfrac{1}{2}\pi\phi^{3/2}5^{-1/4} - \tfrac{5}{4}\ln 5 - \tfrac{1}{2}\sqrt{5}\ln\phi,$$
$$H_{2/5} = \tfrac{5}{2} - \tfrac{1}{2}\pi\phi^{-3/2}5^{-1/4} - \tfrac{5}{4}\ln 5 + \tfrac{1}{2}\sqrt{5}\ln\phi,$$
$$H_{3/5} = \tfrac{5}{3} + \tfrac{1}{2}\pi\phi^{-3/2}5^{-1/4} - \tfrac{5}{4}\ln 5 + \tfrac{1}{2}\sqrt{5}\ln\phi,$$
$$H_{4/5} = \tfrac{5}{4} + \tfrac{1}{2}\pi\phi^{3/2}5^{-1/4} - \tfrac{5}{4}\ln 5 - \tfrac{1}{2}\sqrt{5}\ln\phi,$$

Page 617 and following ______ 03 Jun 1995

Miscellaneous changes to the existing index of Volume 1 are collected here, including corrections and amendments to the old entries as well as new entries that are occasioned by the new material. Thus, the lines of the full index that have changed serve also as an index to the present document. However, when a correction or amendment has caused an old index entry to be deleted, the deletion is usually not indicated.

Page 634 bottom two lines ———— 01 Jun 1995

For additional definitions . . . 1966).

Page 636 and 637 (the endpapers) ———— 01 Jan 1980

replace by new table from endpapers of Volume 2