# GENERATING AND TESTING PSEUDORANDOM NUMBERS

BY CHARLES A. WHITNEY

## Analyze haphazard occurrences with linear congruential generators

F A DRUNKARD starts from a lamppost and randomly staggers away, how far will he have progressed in one thousand steps? Expressed in varving forms, the "drunkard's walk" has become a staple of mathematical physics. How far will an impurity atom migrate in a crystal lattice? How many steps will be required for a photon to emerge from a foggy atmosphere? They are all the same question, and they can be treated with "Monte Carlo" calculations. These calculations are finding increasing applications in business, as well. For example, they provide an analysis of how best to serve customers arriving haphazardly at a counter. Carrying out such calculations depends on being able to imitate randomly generated numbers. and this is not easy. It has been said that more time has been spent generating and testing random numbers than using them.

You can find numerical examples of random series all around you. The final integers in a list of telephone numbers gives a good random series in the range 0 to 9. The face values of cards drawn from a well-shuffled deck and the final digits in licenseplate numbers on passing cars are usually quite reliable. (But the first digits of such numbers are often far from random.)

The property that defines a randomly generated series is that each number is independent of all earlier numbers. In other words, the process that generates the random series has no memory. Therefore, even if you know all the previous numbers, you cannot predict with certainty the next one. (And if you have been losing at a truly random game, you have no reason to think you will start winning.) That's why flipping a coin is such a good method of generating a random series of zeros and ones. Each flip is entirely independent of the others.

Before going further, I should try to clarify a point of language that can lead to confusion. Strictly speaking, no such thing as a random number exists-only a random process. The number 12345 is neither less nor more random than the number 32719. In a series of 100.000 randomly generated numbers, both of these would have the same chance of occurring. The idea that 12345 is less random comes from comparing it with a simple pattern of ascending digits. Thus, instead of saying you are trying to produce random numbers. you should say you are trying to construct a method that will produce a series of numbers that imitates the results of a random process. But for the purpose of this article, I will use the word "random" more loosely to refer to a random sequence or a random number, without worrying about the niceties. If a sequence looks like it was generated by a random process, I will call it random and will put off the question of how to judge appearances until I discuss methods of testing random-number generators.

To carry out a Monte Carlo calculation, you need to work with random numbers inside a computer. But you are faced with the fact that a purely digital computer is a deterministic machine-except on its "off" daysand such a device cannot truly generate a random process. You have to be satisfied with deterministic algorithms that imitate random processes. You can thus generate "pseudorandom" numbers with some of the earmarks of randomness. Naturally, some random-number generators work better than others, and you must be wary.

#### LINEAR CONGRUENTIAL GENERATORS

Mathematicians have suggested many methods for generating pseudorandom numbers with a digital computer. Happily, the most common and powerful one involves simple arithmetic. This method is the linear congruential generator, or LCG for short. An LCG produces a series of numbers,  $I_i$ , where the subscript "i" indicates the location of the number; i = 1 indicates the first number, i = 3 is the third, and so on. Since each successive term,  $I_{i+1}$ , is computed from its predecessor, you can see right away that this series is not truly random because it has memory. That is why the LCG is called a "pseudorandom"-number generator.

To understand the LCG, look at the following linear expression:

$$l_{i+1} = al_i + c$$
,  $(a > 0, c \ge 0)$ 

This expression multiplies each number by the factor a and then adds c to find the next member. It produces an ascending series of numbers whose differences are given by the expression

$$I_{i+1} - I_i = (a-1)I_i + c$$

We call the process "mapping" because it carries the integer,  $I_i$  to  $I_{i+1}$ , from one point to the next in a one-dimensional space, as shown in figure 1.

Suppose a = 2 and c = 1: then you have 1, 3, 7, 15. As it stands, this series doesn't look random. To create a series that looks random, you need a method of scrambling the output, perhaps by cutting off the upward run. One way to achieve this appearance is to imagine you've subdivided the number line representing the *I*-space into segments of length *m*. If you make the multiplier, *a*, in the LCG suf-(continued on page 442)

Charles A. Whitney is a physicist at the Harvard-Smithsonian Center for Astrophysics (60 Garden St., Cambridge, MA 02138) and is a professor of astronomy at Harvard.

#### (continued from page 129)

ficiently large (compared with the segment length, m), the mapping will take the point out of the current segment at each jump. The successive positions of the point within each segment will then be an erratic function. You can consider these successive positions pseudorandom numbers if you suppose each segment is renumbered from 0 to m-1, as in figure 2. You can accomplish such a subdivision of the I-space by introducing the modulus function

#### mod(I,m) = I - [I/m]m

where |I/m| is the integer part of the quotient, derived by truncation.

When, for example, mod(11,3) = 2, the result always lies between 0 and m-1. In some forms of BASIC, this is written 11 MOD 3.

Now rewrite the mapping as

#### $I_{i+1} = \mod (aI_i + c,m)$

This equation is a general form for the LCG, and it produces a series of integers in the range  $0 < I_i < m-1$ . Three parameters describe this mapping: the multiplier (a), the difference (c), and the modulus (m).

Often it is convenient to "normalize" the values, dividing each by the modulus. The result is a series in the range 0 to (m-1)/m < 1, which can be written as

 $x_{i+1} = FLOAT(I_{i+1}) / FLOAT(m)$ 

Note that the smallest nonzero difference between terms is 1/m, which means the numbers the LCG produces comprise a set of m equally spaced rational fractions in the range  $0 \leq x \leq (m-1)/m.$ 

To see some of the properties of series that are generated this way, look at the following series:

#### $I_{i+1} = \text{mod} (5I_i + 3, 8)$

By starting with an arbitrary value, the "seed," and taking  $I_1 = 1$ , you find 1, 0, 3, 2, 5, 4, 7, 6, 1, 0, 3, 2, 5, ....

This series starts off with a haphazard appearance, but since it repeats itself every 8 terms, it is said to have a "period" of 8. It is not hard to see why this is the period. First, the modulus of the series is 8, so the series cannot have more than 8 integers. (You remove larger integers by subtracting the modulus.) Second, the series is deterministic. Each appearance of a particular integer must be followed by a uniquely determined integer. That is, each appearance of "2" must be followed by "5." As a result, the series must repeat itself with a period no longer than its modulus.

This reasoning suggests adopting a large modulus if you want a long period. But it isn't the only possibility. Some generators will skip many of the possible numbers and give an in-

(continued)

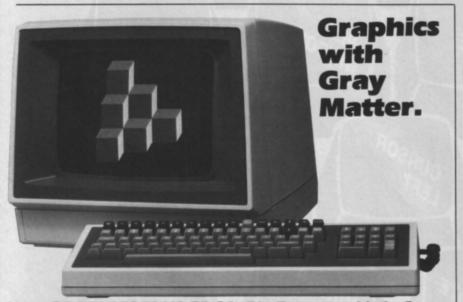
			1 de sil
		shall	n.
AVAL. H	1.11	VOV	-
		PRODUCTS	_
	UTER	PRO	
COM	0.	IBM Personal Com	nutor
	all provide the second		
MONITORS	MARKER -	256KB Memory, DS/D	
Amdali 200	\$139.00	FDC, Color Card, Am	
Amdek 300 Amdek 300A	\$139.00	Monitor with 10MB Ha	ard Disk
Amdek 310A	\$179.00	Sub System.	
Amdek Color I	Best Price	<b>\$2899.</b>	
Amdek Color II	Best Price	(We configure and test th	
PGS-HX-12	\$480.00	for you at no extra cost)	ie system
PGS-MAX-12	\$195.00		
PGS-SR-12 IBM Monochrome Display	Best Price \$265.00	MULTIFUNCTION B	OARDS
IBM Color Display	\$590.00	AST I/O + Ser & Par	\$179.00
		AST Six Pack 64K	\$269.00
PRINTERS		AST Mega Plus 64K	\$269.00
Epson FX 80	\$499.00	AST Combo 64K	\$269.00
Epson FX 100	\$699.00	Quadboard 64K	\$269.00 \$225.00
Okidata 92 A	\$465.00	IBM Color Graphic Adapter IBM Mono/Printer Adapter	\$240.00
Okidata 93 A	\$699.00	Hercules Graphic	\$359.00
Okidata ML 84	\$899.00	Plantronics	<b>Best Price</b>
Okidata 2410 Okidata 2350	\$2085.00 \$1925.00	Paradise Multifunction Card	<b>Best Price</b>
Toshiba P1351 & 1340	Best Price	Orchid - Blossom	Best Price
NEC Spinwriters 3550	\$1699.00	64K Ram Upgrade Kit	\$50.00
NEC Spinwriters 7730	\$1799.00	MODEMS	Crist 1811
C-Itoh Starwriter 40CPS	Best Price		
C-Itoh Printmaster 55CPS	Best Price	Hayes Smart Modem 1200	\$489.00
Gemini 10X Gemini 15X	\$290.00 \$399.00	Hayes Smart Modem 300 Hayes 1200B Plug in Modem	\$209.00
Also available DX-15, HR-2		Card	\$429.00
Silver-Reed.	, m, u	Gard	Q120.00
	A Real Property and	HARD DISK FOR I	BM PC
DRIVES		10MB Hard Disk Sub System	m includes:
Tandon TM 100-2	\$205.00	Software, Controller Cables,	
Slimline Drives - Your Choic		Internal	\$875.00
Hitachi, Panasonic	\$175.00	External	\$1025.00
Many other products	available	, Please call for low - low	w Prices!
Which in the Same of the	micr	coshop	
2640 Waln		it K, Tustin, CA 92680	
Loto Wall			
Holanta attent	The second second second	38-7530	
Prices & availaility subject to cha	ange without no	otice - IBM is a registered trademark	of IBM Corp.
the second			

442 BYTE . OCTOBER 1984

complete set of "random" numbers. A series that generates all of the *m* distinct integers (0 < n < m-1) during each period is called a "full period." Whether a series will have a full period or not depends in large measure on the values you choose for the parameters (*a.c.m*). Table 1 illustrates some of the series with *a* =

5 and m = 8 for various values of the additive constant, *c*. Take the first series as an example. If you use a seed, 6, you obtain the pseudorandom series 6, 7, 4, 5, 2, .... If you use the seed 5, you get 5, 2, 3, 0, 1, 6, ....

Because several of the series in table 1 do not have a full period, they



#### New GRAPHICS-PLUS GP-29. Big Features. Little Cost.

The GP-29 delivers the graphics and text versatility of powerful, expensive terminals at significant cost savings. And you can buy the GP-29 as a ready to use terminal or as a retrofit board for the Zenith Z29.

Expand your applications with dual plane graphics. Create images with "shades of gray." Overlay two separate images. Animate the images. Store multiple images in local memory.

The GP-29 gives you low cost graphics you never imagined possible. 1024 x 500 hi-res and 512 x 250 lo-res selectable resolution. 128k of display memory. And our dedicated graphics processor provides selective vector erase, area erase, area fill, area move, arc drawing, pan and zoom.

You also get Tektronix 4014 compatibility as well as DEC VT100 and VT220 compatibility.

#### **Enhanced Text Too!**

The GP-29 offers four selectable display formats: 80 and 132 columns with 24 or 49 lines. Offscreen scrolling memory. Storage for pages and pages of text. Offline editing capability. Double high and double wide characters and much, much more.

Plus, you get operator convenience features like "plain English" set-up menu, easy to use programmable keys, local function keys, transparent mode and non-volatile memory. And a serial printer port is standard. Call or write today to place your order.

GP-29 retrofit for Zenith Z29 .... \$ 995. GP-29 complete terminal...... 1695. Optional long-persistence CRT .. 95.

Northwest Digital Systems P.O. Box 15288, Seattle, WA 98115 (206) 524-0014

GP-19 board for the Zenith Z19 terminal is also available for \$695.

generate subsets of integers with many useful properties. In the first place, the sum of the periods of the subsets equals the modulus of the series. This property helps you decide whether you have found all the subsets. Second, if the series does not have a full period, different seeds start each subset. For example, the series with c = 4 has six subseries of periods 1 or 2, and the seed you use will determine which subseries you generate.

When you set up a random-number generator, look for a long and full period because it will produce the richest set of numbers. Several established rules for selecting the parameters will achieve a full period. These rules are discussed in Donald Knuth's The Art of Computer Programming. referenced at the end of this article. One rule is that the modulus, m, and the constant, c. must have no factors in common. Another rule is that a must be greater than the square root of  $m (a > \sqrt{m})$  to avoid the serial correlation that upward runs produce. (This rule ensures that the mapping quickly takes the number out of the current segment, the same condition mentioned earlier.) Finally, you will get the longest possible periods if the modulus is a prime number equal to or less than the largest integer your computer can handle. (For a 16-bit processor, this condition implies that m < 32.768.)

Two additional examples of LCGs with short, full periods follow:

 $l_{i+1} = mod(7l_i + 5.9)$ 

 $I_{i+1} = mod(7I_i + 7.9)$ 

In practice, you develop a pseudorandom-number generator in a cutand-try process, testing and modifying various possibilities.

In describing the LCG, I assumed that the coefficients were all integers. You can increase the series' apparent period by taking real (decimal) values. For example, if you substitute 5.1 and 3.111 for the coefficients in the series above, the terms won't repeat precisely after a period of eight terms. But the terms in successive cycles will (continued)

444 BYTE • OCTOBER 1984

Circle 252 on inquiry card.

show only a slight shift, and the overall pattern of the cycle will be the same—except for occasional jumps when the fractional parts accumulate sufficiently. Thus, the increase of period is only illusory, and, since I don't want to fool you with this illusion. I will restrict myself to integer coefficients.

#### STATISTICAL TESTS FOR RANDOMNESS

Often, you can look at a series and see that it was probably not generated by a random process. For example, who would claim that a real coin would lead to a series of heads (H) and tails (T) such as HTHHTTHHHTT-THHHHTTTTHHHHHTTTT? It could happen (with a probability of about  $(1/2)^{29} = 0.00000001$ ), but you wouldn't expect it in your lifetime.

But some series are not so obvious, and you need a more reliable test than the eyeball and a hunch. For this, you must compare some statistical properties of the series with some theoretical predictions you make after assuming that the series was generated by a truly random process. When I refer to a statistical property, I'm talking about one that is independent of the seed you used to start the series, including those tests that are obliv-

## THE CHI-SQUARE TEST

This test is commonly used for agreement between a series of observed values, Oi, and expected values, Ei. In this case, the "observations" are the actual bin populations listed in table 2, and the expected value is the average 1000/100 = 10. There are n = 100 observations, and because n >> 1, the test can be expressed in a simplified form, as follows. For each bin compute the difference.  $O_i - E_i$ ; square these values, divide by the corresponding value of Ei, and sum the quotients over all the observations. Divide the result by the number of observations, obtaining

#### $X^2 = (1/n)\Sigma (O_i - E_i)^2/E_i$

If this quantity lies between 0.5 and 2.0, the scatter of the observations is consistent with random numbers. If  $X^2 < 0.5$ , the observations are suspiciously close to average values; if  $X^2 > 2.0$ , they are too far from the average. This test is applicable only to quantities that are expected to obey the classic law of errors—and this is often a debatable point. For the numbers in table 2,  $X^2 = 0.79$ ; therefore, the populations have the properties of random numbers.

ious to the order of terms in the series. (The mean value is one such property.) These statistical tests reveal how likely it is that a random process generated a certain series.

No statistical test is a sure bet, and few tests are reliable in themselves. Some pseudorandom series will pass one test with flying colors, only to fail miserably in another. Therefore, you have to apply several different tests. I will apply some tests to the LCG and the generator that the IBM PC's Advanced BASIC supplies. Then I'll discuss how to develop a more powerful random-number generator that anyone can use.

Let's start with the simplest test: determining the period of the series.

#### PERIOD

You can determine a period by noting a series' first number and then stepping through it while computing one number after another until the first number recurs, that is, until  $x_n = x_1$ . Then n-1 is the period of the series. (In order to ensure that  $x_n$  is, indeed, the start of a repeat cycle, the first few values,  $x_1 - x_{1+i}$ . can be saved for comparison with  $x_n - x_{n+i}$ .)

Figure 3 shows the unfortunate effect of a short period on the random (continued)

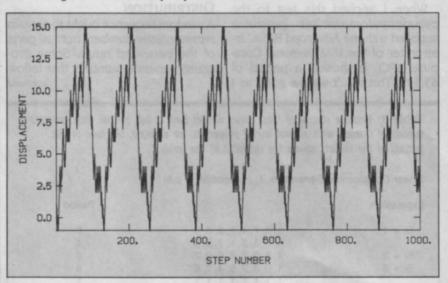


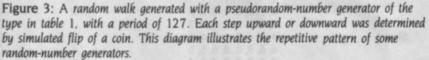


special distribution laws (the normal distribution, for example), but I will consider only the ones that are intended to produce uniformly distributed numbers. If I normalize an LCG, my program should produce numbers in the range 0 to 1.0 with equal prob-

ability. However, the numbers won't arrive in a perfectly uniform way. They will exhibit a tendency to clump, just as the flips of a real coin will show runs of more than one head or tail instead of HTHTHT.

(continued)





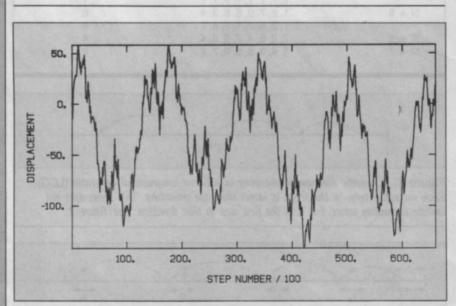


Figure 4: Similar to figures 1 and 2, this walk was generated with the RND function of the Microsoft Advanced BASIC supplied with the IBM PC operating under MS-DOS 2.0. It shows an approximate periodicity of about 18,000 steps, although the rigorous period is about 64,000. Using such a function for Monte Carlo simulations requiring more than 8000 steps could produce misleading results.

walk generated by one of the LCGs in table 1. You can easily spot the periodicity, and you wouldn't want it as the imitation of a very long random walk. (This walk was generated from a normalized LCG by stepping upward if x > 0.5 and downward if x < 0.5.)

When I applied this test to the pseudorandom-number generator supplied with the Advanced BASIC Interpreter of the IBM Personal Computer (PC), it showed a period of 65,535. This result was as good as I could have hoped, but a detailed plot of a walk shows that it also has a much shorter wave-like cycle superposed. Figure 4 shows such a plot and reveals a subcycle that is about 18,000 steps long.

#### DISTRIBUTION

A random sequence ought to contain representative numbers from all parts of the permitted range. Some programs generate numbers that follow (continued)

Period

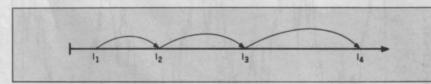
**Table 1**: Because several of these series do not have a full period, they generate subsets of integers with several useful properties, for example, the sum of the periods of the subsets equals the modulus of the series.

 Linear Congruential Generators,  $I_{i+1} = mod(5I_i + c, 8)$  

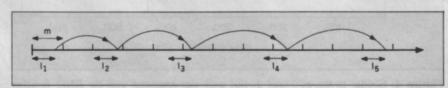
 Expression
  $I_i$  

 5I + 1
 1, 6, 7, 4, 5, 2, 3, 0

51 + 1	1, 6, 7, 4, 5, 2, 3, 0	8
	1, 7, 5, 3, 1, 7, 5, 3	4
51 + 2	4, 6, 0, 2, 4, 6, 0, 2	4
51 + 3	1, 0, 3, 2, 5, 4, 7, 6	8
	1, 1, 1, 1, 1, 1, 1, 1	1
	2, 6, 2, 6, 2, 6, 2, 6	2
	3, 3, 3, 3, 3, 3, 3, 3	the lable 1 with a period of 13
51 + 4	4, 0, 4, 0, 4, 0, 4, 0	2
	5, 5, 5, 5, 5, 5, 5, 5	The second second
	7, 7, 7, 7, 7, 7, 7, 7	1
51 + 5	1, 2, 7, 0, 5, 6, 3, 4	8
	1, 3, 5, 7, 1, 3, 5, 7	4
51 + 6	2, 0, 6, 4, 2, 0, 6, 4	4
51 + 7	1, 4, 3, 6, 5, 0, 7, 2	8



**Figure 1**: Schematic diagram of mapping in a linear congruential generator (LCG). Each successive term in the series is larger than the preceding. This series does not imitate a random series, but it is the first step in that direction. See figure 2.



**Figure 2**: Schematic of an LCG, showing how the division of the number line into equal intervals, m, can produce pseudorandom numbers. The location of each number inside the corresponding interval is haphazard. It results from using the modulus function and leads to a pseudorandom series.

449

ur disc problems, buy 100% surface s. All orders shi d Dysan die stock, within 24 hours. Call toll PREE 0) 235-4137 for prices and is and Master Card accepted. ACIFIC EXCHANGES 100 Foothill Blvd. San Luis Obispo. CA 93401. (In Cal. call (805) 543-1037.) Circle 268 on inquiry card. **PROMPT DELIVERY!!! IC**s SAME DAY SHIPPING (USUALL DYNAMIC RAM 150 ns \$35.77 256K SHOWN 64K 200 ns 4.62 64K 150 ns 4 87 64K 120 ns 5.59 200 ns 1.21 16K EPROM 300 ns \$49.97 27256 32Kx8 27128 16Kx8 300 ns 23.99 22.50 7.97 27C64 8Kx8 200 ns 2764 8Kx8 250 ns 5.40 2732 4Kx8 450 ns 450 ns 3.60 2716 2Kx8 STATIC RAM 150 ns \$35.97 6264LP-15 6116P-3 150 ns 6.36 Factory New, Prime Parts JPoo MICROPROCESSORS UNLIMITED 24.000 South Peoria Ave. (918) 267-4961 BEGGS, OK. 74421 Circle 231 on inquiry card.

LOWE	The second s		
"GUAF			
MONTHS BYTE AM			BY 5%
BABY BLUE II (25 PROCESSING, C	P/M-80 EN	GROUND	\$649
SMART TERMIN AST SIXPACK PL AST SIXPACK PL	US (64K) US (384K)		256 531
EVEREX 10MB HI SYSGEN 10MB HI STREAMING TA	WITH 10	MB	749 2199
			229
	RINTERS		~~~
BROTHER HR25 OKIDATA 92	23 CF 160 CF 160 CF	PS (DAISY S (DOT)	) 689 419 685
OKIDATA 93 SILVER REED 55 ComRiter CR-II		S (DAISY	
TETON D			-
BOX 20320 J/ (307	ACKSO 7) 733-9		3001
	-	-	-

**Table 2**: Sample bin populations; the results for N = 1000 numbers in 100 bins computed with the BASICA RND. The observed values cluster about the expected mean,  $\langle NB \rangle = N/Q = 10$ .

Sample Bin Populations with BASICA RND

7	12	17	9	10	12	7	10	11	12		
8	13	9	12	12	9	9	11	11	15		
9	10	9	11	10	8	9	10	9	8		
11	8		8	10	11	11	10	17	11		
6	9	12	9	8	14	9	10	13	17		
9	5	18	10	9	12	9	9	13	10		
12	9	6	7	8	5	6	12	6	11		
6	7	9	11	9	7	13	7	9	11		
9	6	7	5	15	6	12	8	12	19		
12	11	12	8	10	8	9	. 10	12	8		

You can test the distribution of numbers by setting up *Q* bins and putting each member of the series into one of the bins. For example, if the numbers are restricted to the interval  $0 < x_i < 1$ , each can be put

into bin J, where J is computed from

 $J = int(Q \times x_i), \ 0 < J < Q - 1$ 

On each occurrence of *J*, the bin count, NB(J), is incremented, so that NB(J) = NB(J) + 1.



Table 2 shows the result for N = 1000 numbers in 100 bins computed with the BASICA RND. The observed values cluster about the expected mean,  $\langle NB \rangle = N/Q = 10$ . When you run the test several times, the excesses and deficiencies appear in different bins. As a result, no evidence appears that any particular bins consistently receive more than 1/Q of the counts.

A quantitative measure of performance is the conventional chi-square test, which evaluates a measure of the spread (see "The Chi-Square Test" on page 446). This test estimates how likely it is that the actual value will be different from the expected value in a randomly generated series. If you look at table 2, you find no bins with less than 5, two bins with NB = 5. seven with NB = 6, and so on. The chi-square test examines all the bin populations and tells how often you can expect this particular distribution of populations from a randomly generated series, where you expect NB = 10 on the average.

Applying the chi-square test to the bin populations of table 2 and then for much longer runs using the BASICA generator, you will find that if the random-number generator is pushed to 30,000 terms, it still performs well. The story changes as soon as you get close to the full period of about 65,000 terms. There, all bins are more or less equally filled and the histogram of bin populations, NB(J), becomes tightly peaked about the mean value, < NB>, because all possible values have been achieved. The generator has displayed its entire full period. Near this extreme limit, the generator fails the chi-square test because the chances are small that actual values will be any different from the expected value.

What happens when a randomnumber generator comes to the end of its period is similar to what happens in a game of blackjack when the cards are not collected into the deck after each hand. When 51 cards have been laid out, there is no doubt what the next card will be. You pro-(continued)



NOW! WINDOWS FOR IBM!

#### It also does 54 other things to keep your disk in line.

EVERYTHING YOU ALWAYS WANTED TO DO, BUT WERE AFRAID TO TRY

Unlike some utility programs that are a headache to use, POWER! is engineered to spoil you with 55 features, simple and uniform commands, and utter simplicity of use. POWER! automatically alphabetizes and numbers your files. You select by the number and never type file names again. Need to **[COPY]. [RENAME]. [ERASE].** or **[RUM]** programs? Just type in their menu number! POWER! also locks out your disk's bad sectors **[TEST]** without destroying files-a critical difference from other utilities that search and destroy, without informing you what they've done, leaving you to wonder why your programs won't run. (And POWER! still has 50 commands to go!)

#### **POWER! ONE PROGRAM DOES IT ALL!**

You may own a few utility programs for your computer housekeeping, each with its own commands to memorize. POWER! has all the programs rolled into one 16K integrated package, so you do things you've never tried before –every day. Save sensitive data from prying eyes with [PASS] word protect, move a block of memory [MOVE], look for data [SEARCH] or compare files [CHECK]. POWER! also makes easy work of patching, [DISPLAY/SUBSTITUTE], customizing software [IOAD/SAVE]. Among the other commands are [SIZE], [STAT] [LOG], [DUMP], [TYPE], [JUMP], [FILL], [SET], and the CP/M version lets you restore erased files –even when you don't remember the filename –at a flick of the POWER! [RECLAIM] command. (Still 31 commands to go!)

#### POWER! NOW FOR IBM's PC-DOS AS WELL AS CP/M

We first developed POWER! for CP/M two years ago, and a stack of testimonials from FORD to XEROX testify to its excellence. For IBM-PC<sup>™</sup> users, special features like managing sub-directories, **[CHANGE]**, and a separate creation of up to 8 simultaneous, on-screen **[WINDOWS]** have been added.

#### MONEY-BACK GUARANTEE AND A 10 DAY TRIAL

POWER! has the Seal of Approval from the Professional Software Programmers Association, and you, too, must be happy with POWER!-or your money back! For only \$169 you can now really be in control of your computer. Call Computing! at (415) 567-1634, or your local dealer. For IBM-PC or any CP/M machine. Please specify disk format.

#### **TO ORDER CALL 800 TOLLFREE**

IBM and IBM-PC are registered trademarks of International Business Machines Corporation.



#### CREATE NEW TEXT WITHOUT RETYPING.

DOCU-POWER! turns your existing text files into a database. Now you can create new documents from parts of old files by simply picking sections from the DOCU-POWER! master index. You never have to retype the same words again.

#### DOCU-POWER! WORKS WITH ANY WORD PROCESSOR.

At your leisure, you set up your library files, and then give a DOCU-POWER! mark to any section, paragraph, or even groups of pages you think you may want to use again. DOCU-POWER! automatically indexes them for you, and, at the same time, extracts a comment description from your text — up to 40 characters long.

## **NOW YOU CAN WRITE BY NUMBER.** To create your new text, simply scroll through your DOCU-POWER! index — you have instant window preview into any text — and pick the appropriate numbers. Now you can walk away, free to work on something else. DOCU-POWER! pulls together all the pieces of text, and gets it ready for printing or further editing with your own word processor.

#### MONEY-BACK GUARANTEE AND A 10 DAY TRIAL

DOCU-POWER! is available by mail or through your software dealer — for only \$149. To order, call our 800 Toll Free number. For more information, call Computing! at 415-567-1634. For IBM-PC or any CP/M machine. Please specify disk format.



#### **RANDOM NUMBERS**

You can probably use a generator out to one-half of its period. For small simulations this limit should be good enough.

gressively lose randomness as the dealer approaches the end of the deck—if you keep track.

The upshot is that you can probably use a generator out to one-half of its period. For small simulations this limit should be good enough.

#### AMPLITUDE SPECTRUM FROM FOURIER TRANSFORM

I've already remarked that the random-walk pattern in figure 4, generated from BASICA RND, shows clear signs of waves. The Fourier amplitude spectrum lets you quantitatively measure the waves' size (see "Fourier Spectrum" on page 464). You can derive this spectrum from the fundamental definition of the Fourier coefficients that you'll find in introductory books on applied mathematics. Or you can derive it from the fast Fourier transform subroutines in some software packages. As an example, figure 5 shows the frequency spectrum of the random walk in figure 3, which was generated with an LCG with a period of 127. This spectrum shows the relative amplitudes of waves of various frequencies. You plot the frequencies in terms of the walk's full length, namely 1000 steps, so that the primary period of 127 shows as a peak in the spectrum at about 8 cycles (1000/127) on this spectrum. Because it quickly becomes repetitious, you can't use such an LCG for simulations involving more than a few dozen steps, and the Fourier spectrum puts you on your guard. Figure 6 shows the amplitude spectrum for the BASICA RND pseudorandomnumber generator that comes with

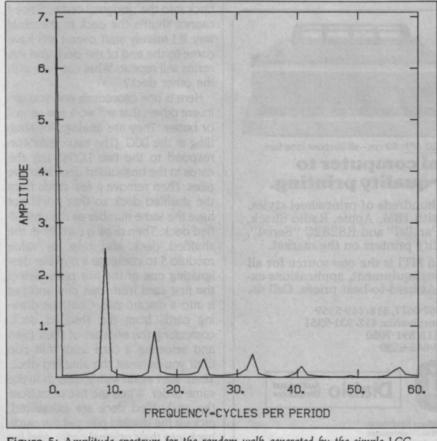
the IBM PC's MS-DOS 2.0 operating system. In this diagram, the frequency is the number of peaks in every 64,000-step run. For example, significant waves show 4, 12, 21, 28, and 37 peaks per 64,000 steps. The most prominent is the wave with frequency 4, which accounts for the main randomwalk plot pattern. The fact that you can divide two of the higher frequencies, 12 and 28, by 4 accounts for the repeated pattern of details on the waves. This test is a clear call for caution in using BASICA RND, and it implies that you need an improved generator.

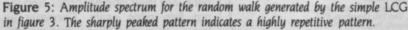
#### SHUFFLING A GENERATOR

How can you extend the period? As mentioned earlier, an LCG's period has two limitations: only integers less than the modulus, *m*, are generated; and the series is deterministic, meaning a particular number always has a

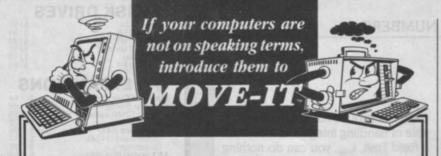
particular sequel. For a computer capable of handling integers smaller than a fixed limit,  $I_{max}$ , you can do nothing about the first restriction. You can, however, alleviate the second restriction and alter the simple determinism of the series using a technique called "shuffling."

Consider an LCG with a period of 8. Each member of the series is an integer from 0 to 7, and if the selection is purely sequential, the period inevitably will be 8. But suppose you set up a secondary list of five numbers and use a second LCG to select the next member of the series from one of them. Then, after each five selections, you replace the five numbers in the secondary list with a randomly selected set. By using two LCGs, you can shuffle the series and extend the effective period. You do not increase the possible integers but prevent an (continued)









Now you can move files, programs and data between incompatible computers. When adding a new computer, data and programs from one system can be transferred to another with MOVE-IT, even if the two have different disk formats and use different operating systems.

A mini-network can be set up by equipping each computer with MOVE-IT, increasing utilization by sharing information freely between machines.

Error Free. A 16 bit error checking technique assures accurate and error-free data transmission even at high speeds and over "dirty" telephone lines or long cable runs. MOVE-IT is the ideal program for telephone line use with modems as well as direct wire applications. If you have ever faced the problem of intercomputer communication, MOVE-IT is what you need.

The simplest dependable solution, the lowest in cost. No expensive circuit boards to install. Easiest to set up. No programming knowledge necessary. Over 17,000 programs sold and operating.

Also MOVE-IT puts you in touch with information networks such as "The Source," "CompuServe," "Dow Jones" and "News Net."

With MOVE-IT you can communicate with over 120 different makes of computers using PC DOS, CPM/86 and CP/M systems. Retail price is \$125 for all CP/M systems and \$150 for all CPM/86 and MS-DOS systems including the IBM PC. Available from your local dealer or from:

SEE US AT COMDEX

#### WOOLF SOFTWARE SYSTEMS, INC. 6754 ETON AVE., CANOGA PARK, CA 91303 • (818) 703-8112 Dealer inquiries invited. © corrigint 1964 woolf Bortware Systems inc



Series 36: 35 cps - all-purpose interface

630 API: 40 cps - all-purpose interface

### Treat your personal computer to famous Diablo letter-quality printing.

**Don't settle for less:** your choice of hundreds of printwheel styles, fully-formed characters, interfaces with IBM, Apple, Radio Shack, Commodore IEEE 488, Centronics "Parallel" and RS232C "Serial". In MTI's opinion, the best letter-quality printers on the market.

Whether you lease or buy, you'll find MTI is the one source for all the computer and data communications equipment, applications expertise and service you'll ever need. At hard-to-beat prices. Call us.

> NewYork: 516/621-6200, 212/767-0677, 518/449-5959 New Jersey: 201/227-5552 Pennsylvania: 412/931-9351 Ohio: 216/464-6688, 513/891-7050 Outside N.Y.S.: 800/645-6530



Distribution/Systems Integration/Maintenance DEC, Intel, Texas Instruments, Hewlett-Packard, Dataproducts, Diablo, Lear Siegler, Esprit, C.Itoh, Racal-Vadic, MICOM, Ven-Tel, Develcon, PCI, U.S.Design, Digital Eng., Cipher, MicroPro, Microsoft, Polygon & Select.

458 BYTE • OCTOBER 1984

Circle 244 on inquiry card. 5 3

#### RANDOM NUMBERS

An LCG's period has two limitations: only integers less than the modulus are generated; and the series is deterministic.

early repetition of the pattern.

Another way to think of this manipulation is in terms of a fictitious game of solitaire. Suppose I have two decks of cards. One of them has been shuffled, but I am asked to produce the longest possible nonrepeating series by laying down the cards from the unshuffled deck. I must follow some well-defined rule of my own choosing, and when I have gone through the deck I can start again. Before continuing, however, I must put the cards back into their original order. Since I cannot shuffle the deck in the usual way, if I merely start over, I will have come to the end of the deck and the series will repeat. What can I do with the other deck?

Here is one procedure, and you can invent others that will work just as well or better. They are analogs for shuffling in the LCG. (The two decks correspond to the two LCGs.) Lay the cards in the unshuffled deck into five piles. Then remove a few cards from the shuffled deck so that it will not have the same number as the unshuffled deck. Then draw a card from the shuffled deck and take its value modulo 5 to compute a number designating one of the five piles. Select the first card from that pile and put it into a discard pile. Continue drawing cards from the shuffled deck, computing the number of their piles and selecting a card from that pile until you exhaust the shuffled deck. Then, start again with the cards in the same order. When the five piles from the unshuffled deck are exhausted, pick up the discard pile, put the cards back into their original order, and lay out five piles again. Continue as before. You will probably find that the

order is quite different the second time and in succeeding sequences. The success of this method depends on having an appropriate number of piles and taking an appropriate number of cards out of the shuffled deck before starting. In a similar way, the success of the shuffled LCG depends on the two series having appropriate relationships between them.

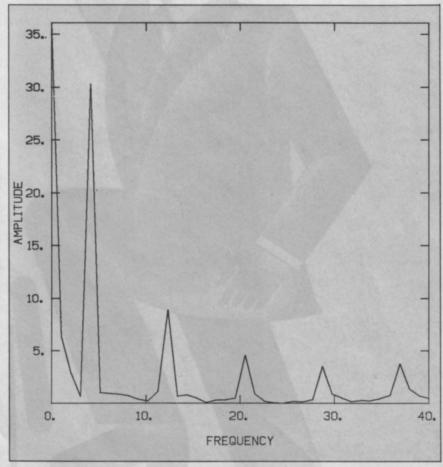
The efficiency of the shuffling technique is quite spectacular. For instance, from a pair of LCGs with periods of only 8 and 9, you can generate a pseudorandom series with a period greater than 200. By tailoring the pair of LCGs to the word length of the computer, you can create shuffled LCGs with much longer periods.

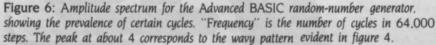
Listing 1 is a BASIC shuffling pro-

gram that generates a random walk consisting of NG groups of NS steps and prints the displacement after each group. By analyzing this listing, you can develop a random-number subroutine suitable for virtually any computer. Notice that subroutine 1010 initializes the program by filling the "piles" with numbers, as in laying out the five piles of cards described above. The program uses two distinct LCGs, and their parameters are listed in statement 32. Before demonstrating the power of this program, I will describe how I selected these parameters.

#### TAILORING A SHUFFLED LCG

I looked for the largest modulus and the largest multiplier consistent with (continued)









#### Buy/Sell Used Hardware Without Risk!!!

In addition to receiving a monthly newsletter containing latest computer information and sales, membership allows you 1 free ad plus additional ads at reduced prices. Buyer selects ad, sends purchase price to Computer Swap Shop who holds same in escrow and notifies seller who ships to buyer. Buyer has 7 days to examine the equipment and if satisfied, seller receives sales price less small commission; otherwise, money is refunded. You must be a member to buy or sell with Computer Swap Shop Inc.

NO RISK! BONDED. Send \$20 subscription fee to: Computer Swap Shop, Inc. Box 2988

Delray Beach, FL 33444

Circle 73 on inquiry card.



Det Matrix Printer Interfaces with Apple II Featuring an Apple II®-compatible parallel interface, Addmaster Corporation has produced a new dot matrix printer, Model 170. The Interface Includes a Centronics-type handshake and DB-25 interface connector, Baudot, and day and time clock. The Model 170 provides 18 or 21 characters per line, 6 lines per inch print density, on standard 2½" adding machine tape. Designed to use with personal computers, Model 170 will produce hard and carbonless copies of programs, data or results. Write Addmaster Corporation, 416 Junipere Serra Dr., San Gabriel, CA \$1776 or call 213/285-1121. **RANDOM NUMBERS** 

 $I_{max} = 32,768$ , the largest integer my computer can handle.

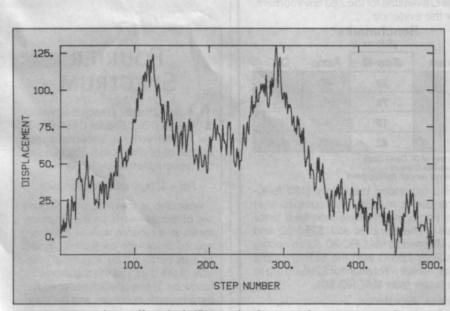
I first needed an expression for the largest integer that will be computed by a particular LCG. It is approximately  $a \times m + c$  before it is reduced modulo, m. This integer must satisfy  $a \times m + c < I_{max}$ .

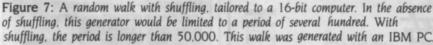
The condition  $a > \sqrt{m}$  also constrains the multiplier to avoid serial correlations. These two constraints suggest taking a small value for the additive constant, *c*, and setting a = m, which leads to the approximate relationship  $a = m = \sqrt{I_{max}}$ . For a 16-bit computer,  $I_{max} = 32.768$ , which

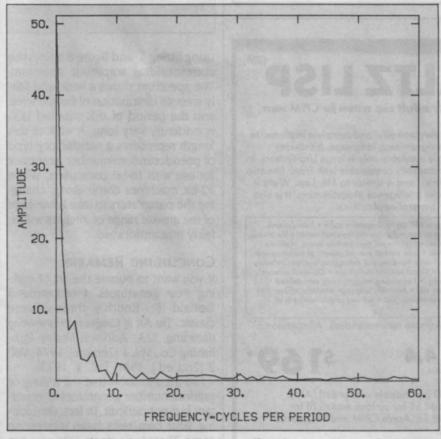
implies a = m < 181. Thus, you can construct a shuffled LCG from a pair of LCGs with periods less than 181. I then wrote a program to run through the output of an LCG and test for full period. After some experimenting, I found that the following pair of LCGs have full periods:  $I_{i+1} = mod(111I_i + 11, 151)$ ; and  $I_{i+1} = mod(113I_i + 13, 137)$ . Finally, the choice of the number of piles did not seem to me to be critical, and I settled on NPILES = 121.

Figure 7 shows a 50,000-step random walk generated on an IBM PC by (continued)

Listing 1: A program in BASIC for generating pseudorandom sequence from two LCGs with shuffling. PROGRAM RANWALK5 with shuffling; MS-BASICA 10 OPEN "RANWALK5.OUT" FOR OUTPUT AS #2 15 **DIM IXS(200)** 16 INPUT "NPILES"; NPILES ' \*\*\* Use NPILES = 121 for secondary list 18 30 INPUT "SEED":IS IM2 = 137:IA2 = 113:IC2 = 13:IM1 = 151:IA1 = 111:IC1 = 11 ' \*\*\* LCGs 32 GOSUB 1010 ' TO INITIALIZE RAND 34 INPUT "STEPS PER GROUP":NS 40 50 INPUT "NUMBER OF GROUPS"; NG 60 PRINT -2, "RANWALK5, LCG#(111,11,151),LCGC(113,13,137):";NS;NG;NPILES 68 ICOUNT = 0 FOR I = 1 TO NG 70 90 FOR J = 1 TO NS GOSUB 2010 ' TO COMPUTE RAND 92 93 ICOUNT = ICOUNT + 1 94 IF ICOUNT = NPILES + 1 THEN GOSUB 1010: ICOUNT = 0 ' reset piles IF RAND > .50 THEN MD = 1 ELSE MD = - 1 100 M = M + MD110 111 NEXT J PRINT #2,I;",";M;"," ' \*\*\* output of random displacement 114 NEXT I 128 GOTO 5000 130 ' \*\*\* end of main program 140 1000 ' \*\*\* Initialize piles of secondary list for shuffling 1010 FOR IK = 1 TO NPILES IX1 = IX1\*IA1 + IC1 1020 1022 JJ = INT(IX1/IM1):IX1 = IX1 - IM1\*JJ 1029 IXS(IK) = IX1 1030 NEXT IK 1040 RETURN 2000 ' \*\*\* RAND drawing from piles without replacement 2010 IX2 = IX2\*IA2 + IC2 2011  $JJ = INT(IX2/IM2):IX2 = IX2 = IX2 - JJ^{*}IM2$ 2012 ICELL = INT(NPILES\*IX2/IM2) + 1 2020 RAND = IXS(ICELL)/IM1 2040 RETURN 5000 END



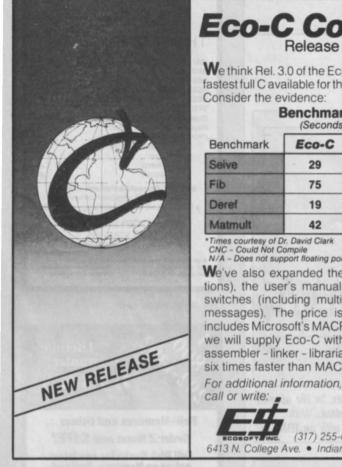


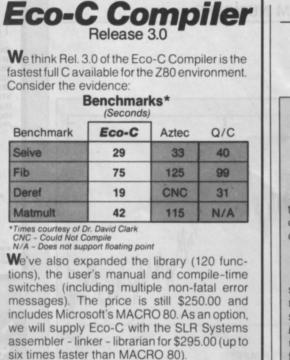


**Figure 8:** Amplitude spectrum of the random walk in figure 7, showing absence of significant periodicities and indicating that the shuffled pair of LCGs passes this test successfully.

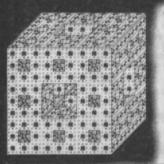
The best, best value, best reliability, or best feature on sale. OFFERING THE BEST FOR LESS \$ IBM PC Call BEST graphic card BEST VALUE word processing \$349 128 **BEST VALUE spreadsheet** 139 BEST VALUE diskettes DS/DD 51/4", 10 35 STB Super Rio 64 K 315 Quadram Expanded Quadboard Hayes Smartmodem 1200B Call 419 PRINTERS Oki 92 par 425 Brothers HR-25 695 Call Epson Diablo Call Ricoh Call **BEST PC Compatible Package: including** a BEST VALUE computer system and a BEST VALUE word processing \$2189 COMPUTER SALES Co (619) 576-9185 CA 92111 Circle 72 on inquiry card. Lifetime Guarantee!! DISKS Fuji-Memorex and Others... **Order 2 Boxes and SAVE! Call Disk Works for our latest** prices on Memorex, Fuji and 3M diskettes. 1-800-292-1492 Nationwide or (312) 368-0359 in Illinois 5 1/4 SSDD \$1.79 each 5 1/4 DSDD \$2.29 each 3 1/5 DSDD \$3.80 each Prices are per disk in quantity 2 boxes of 10. Add \$2.50 shipping and handling. Call for quantity pricing and shipping. All orders shipped same day via UPS ground. **DISK WORKS** S. LaSalle St., Suite 2601 Chicago, IL 60603 Circle 105 on inquiry card. NEW! **KEYBOARD PROTECTOR** Remains in place during keyboard use. Prevents damage from liquid spills, dust, ashes, etc. Fits like a second skin, excellent feel. Homerow and numeric locators. Available for: IBM-PC, Apple IIe, Radio Shack Model 100, Commodore 64. Send \$29.95, check or M.O., Visa & MC include exp. date. Specify computer type. Dealer inquires invited. Free brochure available MERRITT Computer Products, Inc. 2925 LBJ, #180 / Dallas, Texas 75234 (214) 942-1142 Circle 220 on inquiry card.

Circle 123 on inquiry card.







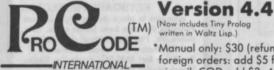


(TM) VALTZL The one and only adult Lisp system for CP/M users.

Waltz Lisp is a very powerful and complete implementation of the Lisp programming language. It includes features previously available only in large Lisp systems. In fact, Waltz is substantially compatible with Franz (the Lisp running under Unix), and is similar to MacLisp. Waltz is perfect for Artificial Intelligence programming. It is also most suitable for general applications.

Much faster than other microcomputer Lisps. • Long integers (up to 611 digits). Selectable radix • True dynamic character strings. Full string operations including fast matching/extraction. • Flexibly implemented random file access. • Binary files. • Standard CP/M devices. • Access to disk directories. • Functions of type lambda (expr), nlambda Bindry lites, \* Jainbard Cr/M devices, \* Access to ank alrectories, \* Puncins of type lambda (expr), inambda (fexpr), lexpr, macro. \* Splicing and non-splicing character macros. \* User control over all aspects of the interpreter.
 Built-in prettyprinting and formatting facilities, \* Complete set of error handling and debugging functions including user programmable processing of undefined function references. \* Virtual function definitions. \* Optional automatic loading of initialization file. \* Powerful CP/M command line parsing. \* Fast sorting/merging using user defined comparison predicates. \* Full suite of mapping functions, iterators, etc. \* Assembly language interface. \* Over 250 functions in total. \* The best documentation ever produced for a micro Lisp (300 + full size pages, hundreds of illustrative examples) illustrative examples).

Waltz Lisp requires CP/M 2.2, Z80 and 48K RAM (more recommended). All common 5" and 8" disk formats available.



15930 SW Colony Pl. Portland, OR 97224 Unix' Bell Laboratories. CP/M\* Digital Research Corp. \*Manual only: \$30 (refundable with order). All foreign orders: add \$5 for surface mail, \$20 for airmail. COD add \$3. Apple CP/M and hard sector formats add \$15.

Call free 1-800-LIP-4000 Dept. #12 In Oregon and outside USA call 1-503-684-3000

#### RANDOM NUMBERS

## FOURIER SPECTRUM

amed for the French mathematician J. B. J. Fourier (1768-1830), the Fourier series is a representation of a series of data points, F(t), as a sum of harmonic terms in the form

 $F(t) = \sum |a_n \sin(nwt) + b_n \cos(nwt)|$ 

When F(t) is specified at a discrete set of points-such as the displacements in a random walk at uniformly spaced times-the coefficients. a, and  $b_{\mu}$  can be found by summing products of the data and the trigonometric functions. These coefficients comprise the amplitude spectrum, and the quantity  $\operatorname{sqrt}(a_n^2 + \hat{b}_n^2)$  is a measure of the importance of harmonic frequency nwin the data. Large values of a, or b, indicate that the data has a significant component of variation with a period  $2\pi/(nw)$ .

using listing 1, and figure 8 shows the corresponding amplitude spectrum. The spectrum shows a wide and fairly smooth distribution of frequencies, and the period of this shuffled LCG is evidently very long. A walk of this length represents a satisfactory type of pseudorandom-number generator for use with 16-bit computers. When 32-bit machines come along, changing the parameters to take advantage of the greater range of integers will be fairly straightforward.

#### CONCLUDING REMARKS

If you want to pursue the art of making new generators, I recommend Donald E. Knuth's three-volume classic, The Art of Computer Programming (Reading, MA: Addison-Wesley Publishing Co., Vol. 1 [2nd ed.], 1974; Vol. 2 [2nd ed.], 1981; Vol. 3, 1973).

I've tried to show that the testing of random-number generators is important but not difficult. In fact, developing your own tests is an interesting game. There is no single right way, but the listing for the program provided here works quite well on my 16-bit machine.

Circle 288 on inquiry card.

5**9**\*