# Random-Number Generation

**DILBERT** *By* Scott Adams

# Random-Number

- As the term suggests, a random number is a number chosen by chance -- i.e., randomly, from a set of numbers.
- All the numbers in a specified distribution have equal probability of being chosen randomly.
- For a number in a sequence or distribution to be truly random, it must be independent.
- The independence of numbers means there is no correlation between successive numbers.

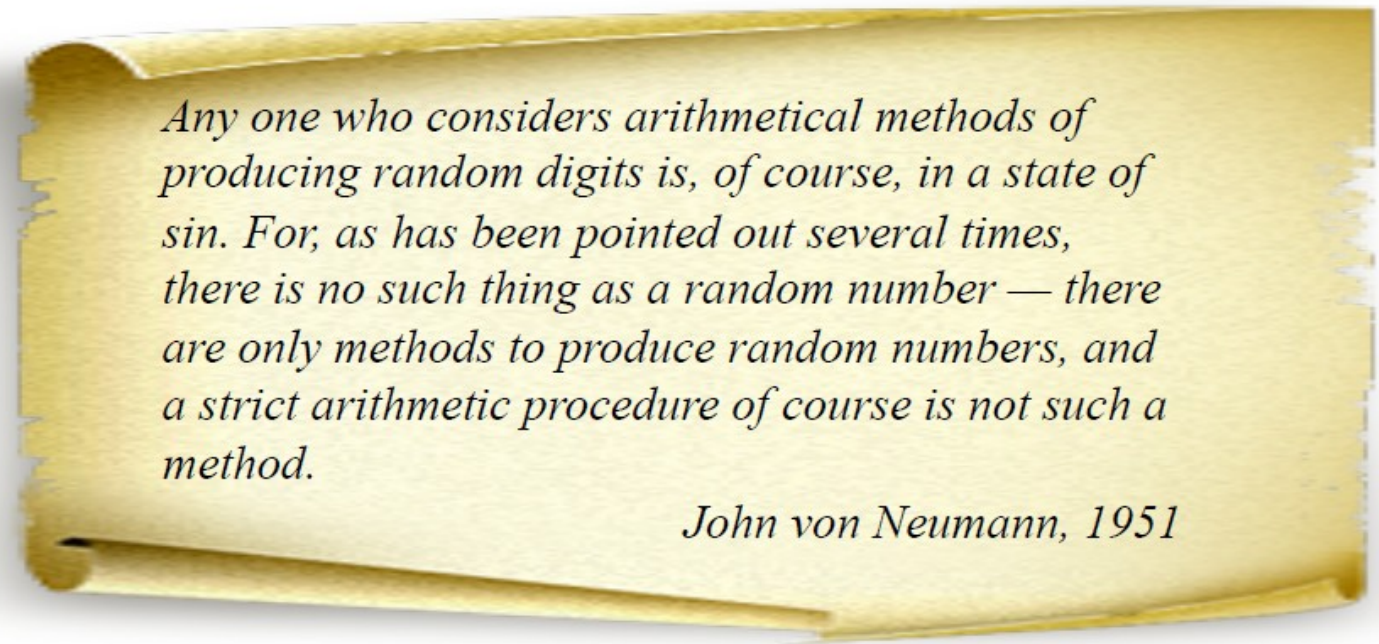
Historically

- Throw dices
- Deal out cards
- Draw numbered balls
- Use digits of π
- Mechanical devices (spinning disc, etc.)
- Electric circuits
- Electronic Random Number Indicator (ERNIE)
- Counting gamma rays
- In combination with a computer
  - Hook up an electronic device to the computer
  - Read-in a table of random numbers

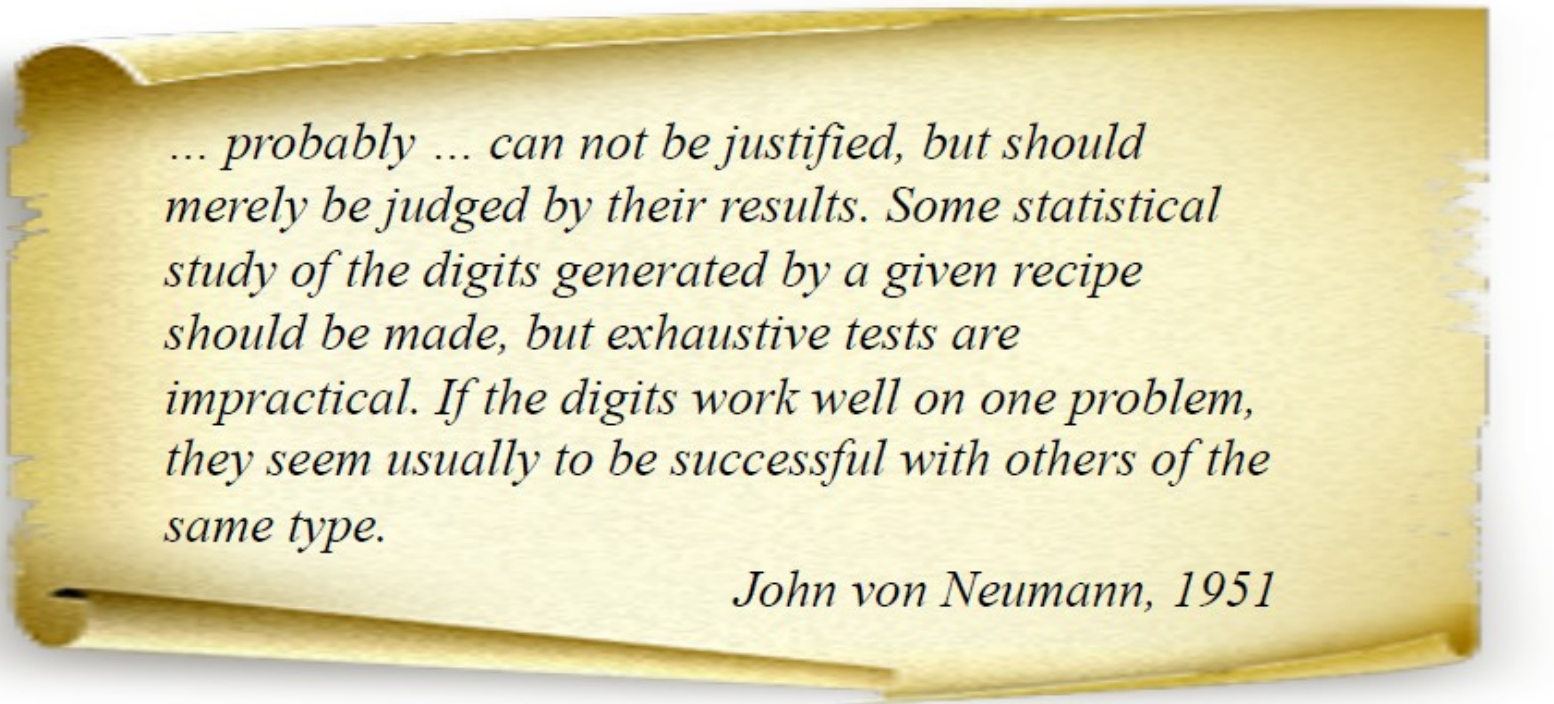
**The gamma counter is an instrument used to measure gamma radiation emitted by a radionuclide in samples collected from patients for detecting certain disease conditions. The gamma counter uses crystals, which emit light when photons from the gamma rays interact with them**

# Pseudo-Random Numbers

- Approach: Arithmetically generation (calculation) of random numbers
- "Pseudo", because generating numbers using a known method removes the potential for true randomness.



# Pseudo-Random Numbers




- Goal: To produce a sequence of numbers in [0,1] that simulates, or imitates, the ideal properties of random numbers (RN).

# Pseudo-Random Numbers

Important properties of good random number routines:

- Fast
- Portable to different computers
- Have sufficiently long cycle
- Replicable
- Verification and debugging
- Use identical stream of random numbers for different systems
- Closely approximate the ideal statistical properties of uniformity and independence

# Pseudo-Random Numbers

Problems when generating pseudo-random numbers

- The generated numbers might not be uniformly distributed
- The generated numbers might be discrete-valued instead of continuous-valued
- The mean of the generated numbers might be too high or too low
- The variance of the generated numbers might be too high or too low

There might be dependence:

- Autocorrelation between numbers
- Numbers successively higher or lower than adjacent numbers
- Several numbers above the mean followed by several numbers below the mean

# Generating Random Numbers

- Midsquare method
- Linear Congruential Method (LCM)
- Combined Linear Congruential Generators (CLCG)
- Random-Number Streams

# Midsquare method

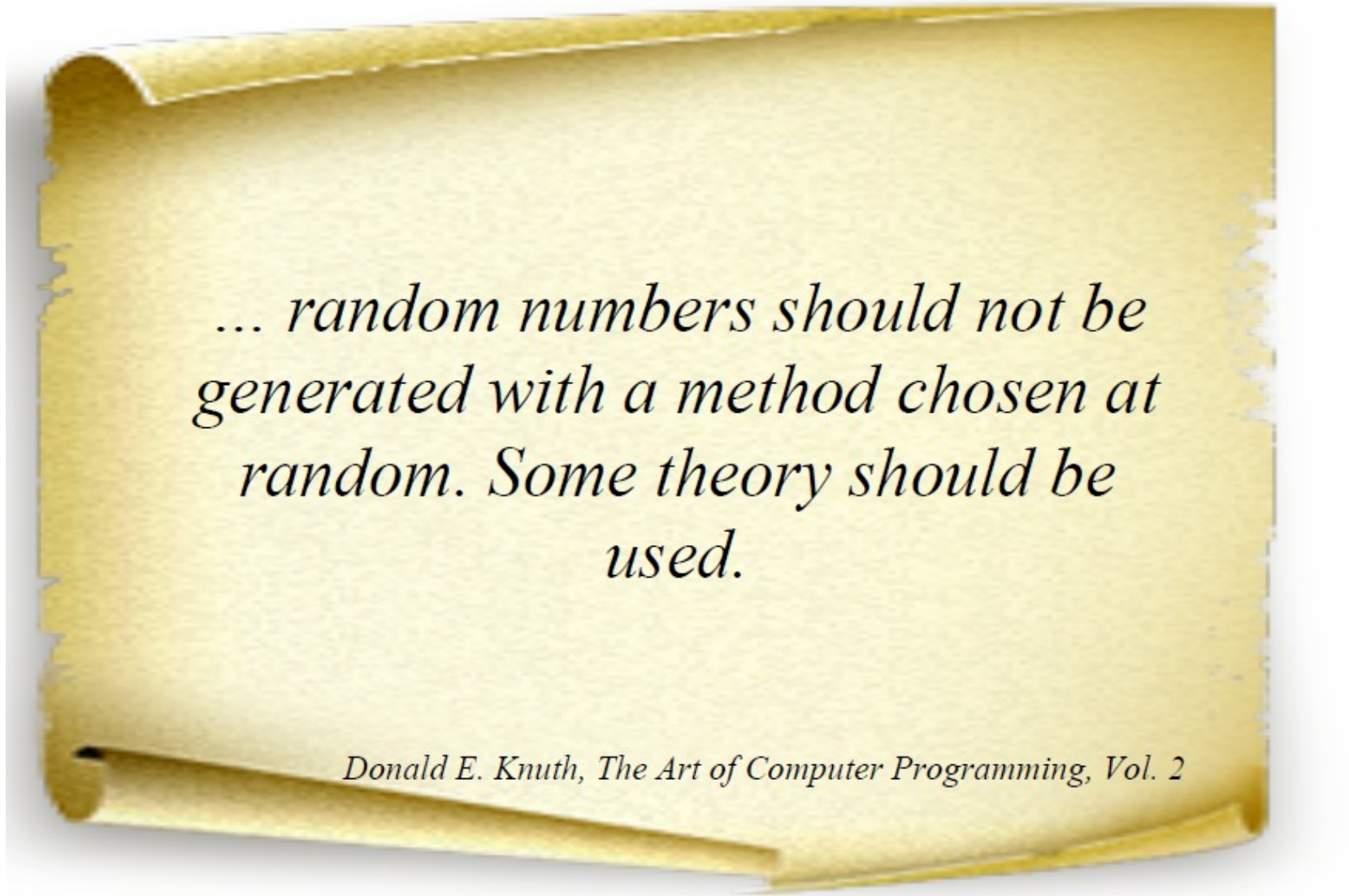
- First arithmetic generator: Midsquare method
  - von Neumann and Metropolis in 1940s
- The Midsquare method:
  - Start with a four-digit positive integer $Z_0$
  - Compute: $Z_0^2 = Z_0 \times Z_0$ to obtain an integer with up to eight digits
  - Take the middle four digits for the next four-digit number

| $i$ | $Z_i$ | $U_i$ | $Z_i \times Z_i$ |
|---|---|---|---|
| 0 | 7182 | - | 51581124 |
| 1 | 5811 | 0.5811 | 33767721 |
| 2 | 7677 | 0.7677 | 58936329 |
| 3 | 9363 | 0.9363 | 87665769 |
| … | | | |

# Midsquare method

- Problem: Generated numbers tend to 0

| $i$ | $Z_i$ | $U_i$ | $Z_i \times Z_i$ |
|---|---|---|---|
| 0 | 7182 | - | 51581124 |
| 1 | 5811 | 0,5811 | 33767721 |
| 2 | 7677 | 0,7677 | 58936329 |
| 3 | 9363 | 0,9363 | 87665769 |
| 4 | 6657 | 0,6657 | 44315649 |
| 5 | 3156 | 0,3156 | 09960336 |
| 6 | 9603 | 0,9603 | 92217609 |
| 7 | 2176 | 0,2176 | 04734976 |
| 8 | 7349 | 0,7349 | 54007801 |
| 9 | 78 | 0,0078 | 00006084 |
| 10 | 60 | 0,006 | 00003600 |
| 11 | 36 | 0,0036 | 00001296 |
| 12 | 12 | 0,0012 | 00000144 |
| 13 | 1 | 0,0001 | 00000001 |
| 14 | 0 | 0 | 00000000 |
| 15 | 0 | 0 | 00000000 |


... random numbers should not be generated with a method chosen at random. Some theory should be used.
Donald E. Knuth, The Art of Computer Programming, Vol. 2

# Linear Congruential Method

- To produce a sequence of integers $X_1, X_2, \ldots$ between $0$ and $m$-$1$ by following a recursive relationship:

$$X_{i+1} = (aX_i + c) \bmod m, \quad i = 0,1,2,\ldots$$

The multiplier

The increment

The modulus

- Assumption: $m > 0$ and $a < m$, $c < m$, $X_0 < m$
- The selection of the values for $a$, $c$, $m$, and $X_0$ drastically affects the statistical properties and the cycle length
- The random integers $X_i$ are being generated in $[0, m\text{-}1]$

# Linear Congruential Method

- Convert the integers $X_i$ to random numbers

$$R_i = \frac{X_i}{m}, \quad i = 1,2,\ldots$$

- Note:
  - $X_i \in \{0, 1, \ldots, m\text{-}1\}$
  - $R_i \in [0, (m\text{-}1)/m]$

# Linear Congruential Method: Example

- Use $X_0 = 27$, $a = 17$, $c = 43$, and $m = 100$.
- The $X_i$ and $R_i$ values are:

$$X_1 = (17 \times 27 + 43) \bmod 100 = 502 \bmod 100 = 2 \quad \Rightarrow \quad R_1 = 0.02$$
$$X_2 = (17 \times 2 + 43) \bmod 100 = 77 \quad \Rightarrow \quad R_2 = 0.77$$
$$X_3 = (17 \times 77 + 43) \bmod 100 = 52 \quad \Rightarrow \quad R_3 = 0.52$$
$$X_4 = (17 \times 52 + 43) \bmod 100 = 27 \quad \Rightarrow \quad R_3 = 0.27$$
$$\ldots$$

# Linear Congruential Method: Example

- Use $a = 13, c = 0$, and $m = 64$
- The period of the generator is very low
- Seed $X_0$ influences the sequence

| $i$ | $X_i$ $X_0$=1 | $X_i$ $X_0$=2 | $X_i$ $X_0$=3 | $X_i$ $X_0$=4 |
|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 |
| 1 | 13 | 26 | 39 | 52 |
| 2 | 41 | 18 | 59 | 36 |
| 3 | 21 | 42 | 63 | 20 |
| 4 | 17 | 34 | 51 | 4 |
| 5 | 29 | 58 | 23 | |
| 6 | 57 | 50 | 43 | |
| 7 | 37 | 10 | 47 | |
| 8 | 33 | 2 | 35 | |
| 9 | 45 | | 7 | |
| 10 | 9 | | 27 | |
| 11 | 53 | | 31 | |
| 12 | 49 | | 19 | |
| 13 | 61 | | 55 | |
| 14 | 25 | | 11 | |
| 15 | 5 | | 15 | |
| 16 | 1 | | 3 | |

# Linear Congruential Method:
## Characteristics of a good Generator

- Maximum Density
  - The values assumed by $R_i$, $i=1,2,\ldots$ leave no large gaps on [0,1]
  - Problem: Instead of continuous, each $R_i$ is discrete
  - Solution: a very large integer for modulus $m$
    - Approximation appears to be of little consequence

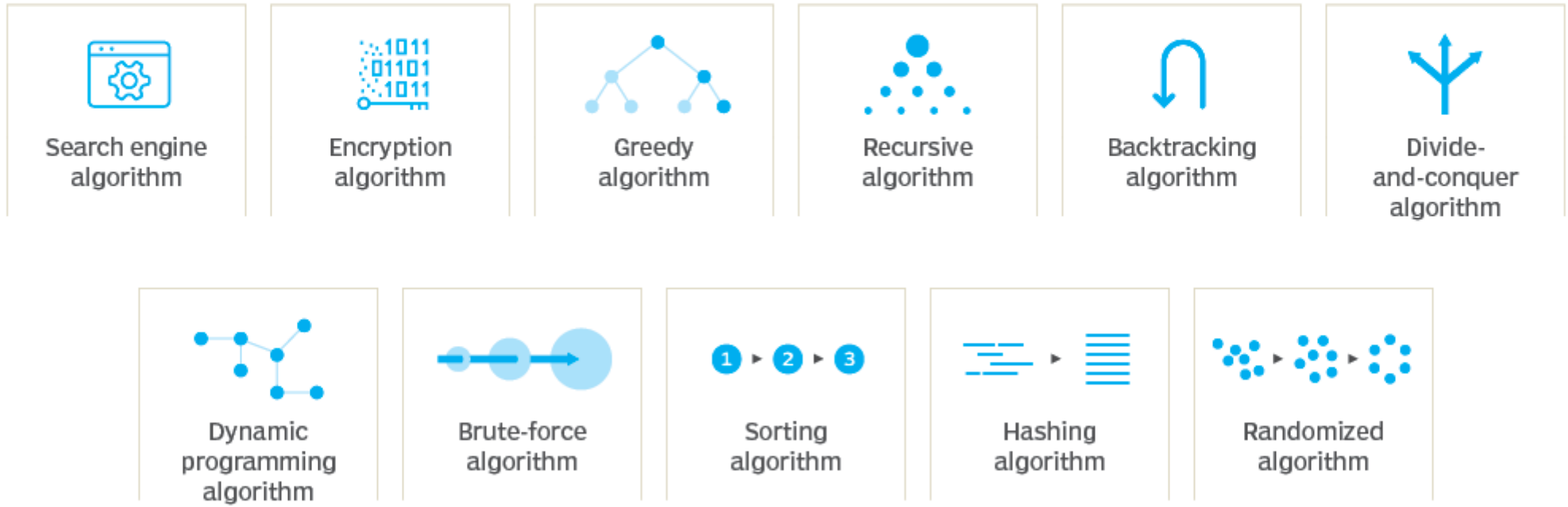
- Maximum Period
  - To achieve maximum density and avoid cycling
  - Achieved by proper choice of $a, c, m$, and $X_0$

- Most digital computers use a binary representation of numbers
  - Speed and efficiency are aided by a modulus, $m$, to be (or close to) a power of 2.

# Linear Congruential Method:
## Characteristics of a good Generator

- The LCG has full period if and only if the following three conditions hold (Hull and Dobell, 1962):
  1. The only positive integer that (exactly) divides both $m$ and $c$ is 1
  2. If $q$ is a prime number that divides $m$, then $q$ divides $a$-1
  3. If 4 divides $m$, then 4 divides $a$-1

# Types of algorithms



# Outline

- Properties of Random Numbers
- Pseudo-Random Numbers
- Generating Random Numbers
  - Linear Congruential Method
  - Combined Linear Congruential Method