How to Keep Up with Mathematics

Third Kenneth C. Schraut Memorial Lecture

University of Dayton

November 2, 2002

Paul J. Campbell '67

Beloit College

Beloit, WI 53511

campbell@beloit.edu

Outline

- SEE . . .
- SEEK . . .
- SPEAK . . .

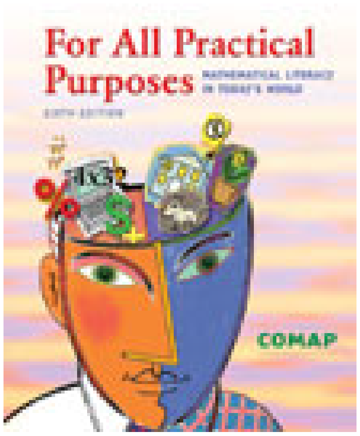
mathematics

SEE mathematics . . . :

serious applications

- Error-correcting codes (book numbers, UPCs, driver's licenses, banknotes)
- Benford's law
- CDs and MP3s
- Web searching

ISBN Numbers



Example:

$$\underbrace{0}_{\textit{country}} \quad - \quad \underbrace{7167}_{\textit{publisher}} - \underbrace{4782}_{\textit{book}} - \underbrace{\qquad}_{\textit{check digit}}$$

For a correct ISBN $a_1 a_2 \cdots a_{10}$,

$$10a_1 + 9a_2 + 8a_3 + 7a_4 + 6a_5 + 5a_6 + 4a_7 + 3a_8 + 2a_9 + 1a_{10} \equiv 0 \bmod 11, \quad \text{or}$$

$$\begin{aligned} a_{10} &\equiv 1a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 \\ &\quad + 6a_6 + 7a_7 + 8a_8 + 9a_9 \bmod 11 \\ &= \text{remainder after division by 11} \\ &= 1 \cdot 0 + 2 \cdot 7 + 3 \cdot 1 + 4 \cdot 6 + 5 \cdot 7 \\ &\quad + 6 \cdot 4 + 7 \cdot 7 + 8 \cdot 8 + 9 \cdot 2 \\ &= 0 + 14 + 3 + \ldots + 18 \\ &= 231 = 0 + 21 \cdot 11 \equiv 0 \bmod 11 \end{aligned}$$

When the remainder is 10 . . .

The ISBN detects all transposition errors, all single errors—90% of all errors.

Benford's Law

Frequency of leading digit $d = 1, 2, \ldots, 9$:
you expect $1/9$, but for many data sets get:

$$\log_{10}(1 + 1/d)$$

$$\log_{10}\left(1 + \tfrac{1}{1}\right) \approx .3, \ldots, \log_{10}\left(1 + \tfrac{1}{9}\right) \approx .05$$

!!!

1881: Simon Newcomb: log tables
1938: Frank Benford: areas of rivers,
atomic weights, electric bills, . . .
—analogous laws for second, third, etc.
significant digits

Why?

1995: Ted Hill (Georgia Tech) proved
—unique base-invariant distribution
—unique scale-invariant distribution
—random samples from random
"neutral" data sets

—applications: fraud detection;
computer design

CDs and MP3 Players

Pulse Code Modulation:
—samples sound (voltage) 44,100 / sec
—gives 16-bit digital result
—for one-minute of stereo: 8.1 MB
—CD: 650 MB / 8.1 MB / min ≈ 80 min
—error-correction overhead
(Reed-Solomon code) ⇒ 74 min

Compression:
—PK-ZIP, Stuffit: reduce only 10%
—special algorithms: reduce 60%

Lossless ⟹ "lossy" compression
1992: MP3 (= MPEG3): compression ratio from 8 to 12
—divides frequency range into 32 bands
—modified discrete cosine transform gives 18 constituents / band
—*where the loss comes in:* "masking," removing redundancy in each band
—Huffman coding compression (shorter codes for frequent values)

Web Searching

Web site: important if others link to it

$x_1, x_2, \ldots, x_n$ the importances

Idea: importance of site $\propto$ sum of
importances of sites that link to it

e.g.,

$$x_1 = k(x_{14} + x_{97} + x_{541})$$

Construct $n \times n$ matrix A:

$$a_{i,j} = \begin{cases} 1, & \text{if site } j \text{ links to site } i; \\ 0, & \text{else.} \end{cases}$$

$$x_i = k \sum_{j=1}^{n} a_{i,j} x_j \qquad (i = 1, \ldots, n)$$

as matrices: $X = kAX$, $AX = \frac{1}{k} X$

—an eigenvalue / eigenvector problem!

—Google uses a variant: $n = 2.7$ billion!!!

eigenvalue ranking: Kendall, Wei 1950s

another application: “top ten” teams

SEE mathematics . . . :

fun examples

- Matching problems
  (gift, card game problems)
- Dynamic programming
  (Farmer Klaus, football)
- The Wave
- Markov chain models
  (baseball, Monopoly, . . .

Matching Problems

The Secret Santa, or Guy Gift, Problem
(aka The Hat-Check Problem)

$$P(\text{no match}) \approx \left(1 - \frac{1}{n}\right)^n \to e^{-1} \approx .37$$

The Guy Gift Problem with Couples

$$P(\text{no match}) \to e^{-2} \approx .14$$

The Bridge Couples Problem

$$P(\text{no match}) \to e^{-1/2} \approx .61$$

$$p(k) \to \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{Poisson, } \lambda = 1/2$$

The Card Game of War

$$P(\text{no “battles”}) \to e^{-3/2} \approx .22$$

$$p(k) \to ?? \text{ Poisson, } \lambda = 3/2$$

$$P(\text{“annihilation”}) = e^{-3/2}/2^{26} \approx 3 \times 10^{-9}$$

A Different Matching Problem

Algorithm for Matching
Medical Residents to Hospitals
(aka The Marriage Problem)

Farmer Klaus and the Mouse

Goal: Get all the grain in (6 sacks each of 4 kinds) before 6 mice come out of their holes.

6-sided die:
each kind of grain, a mouse, and a generic sack

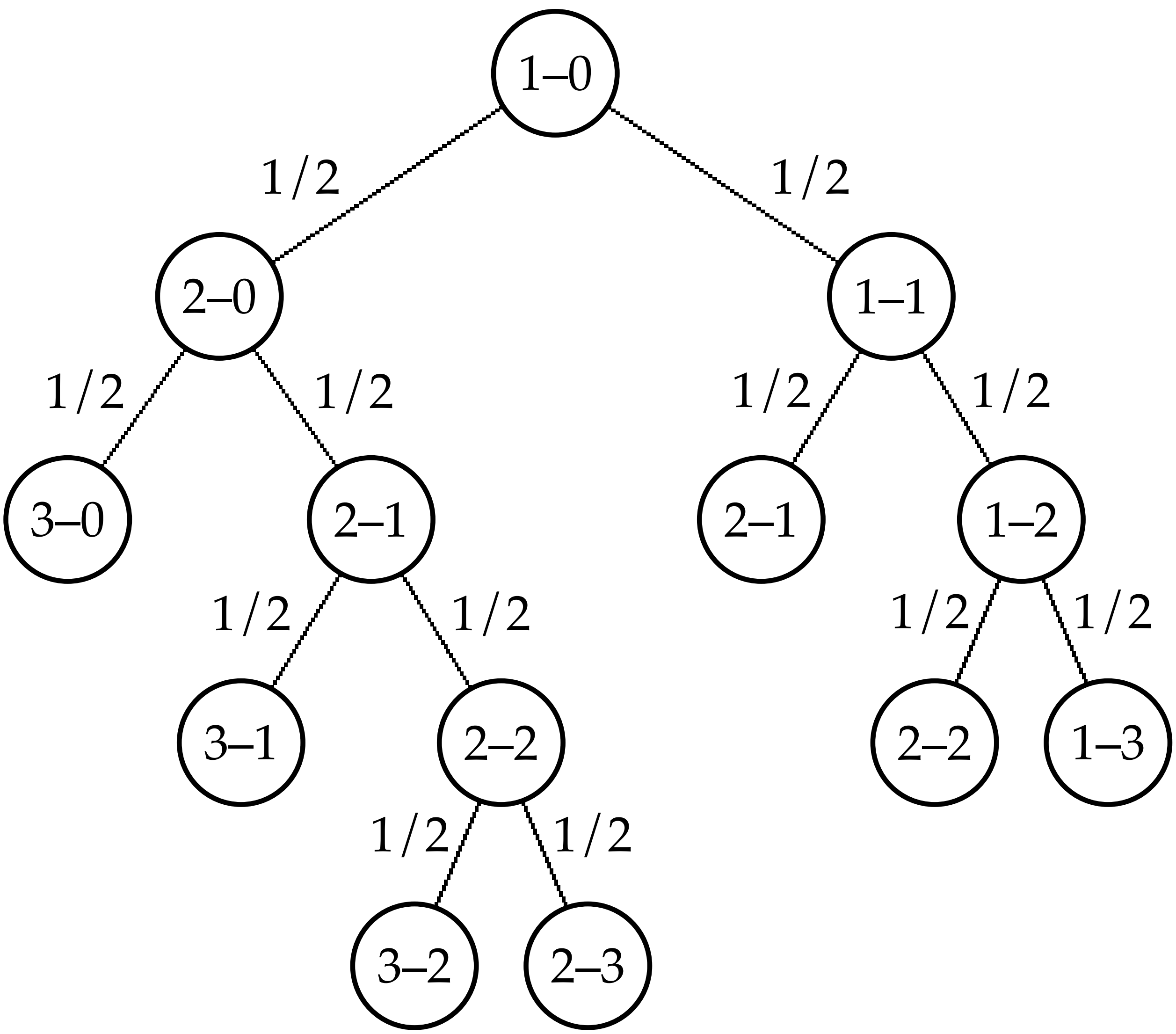
What is P(children win)?

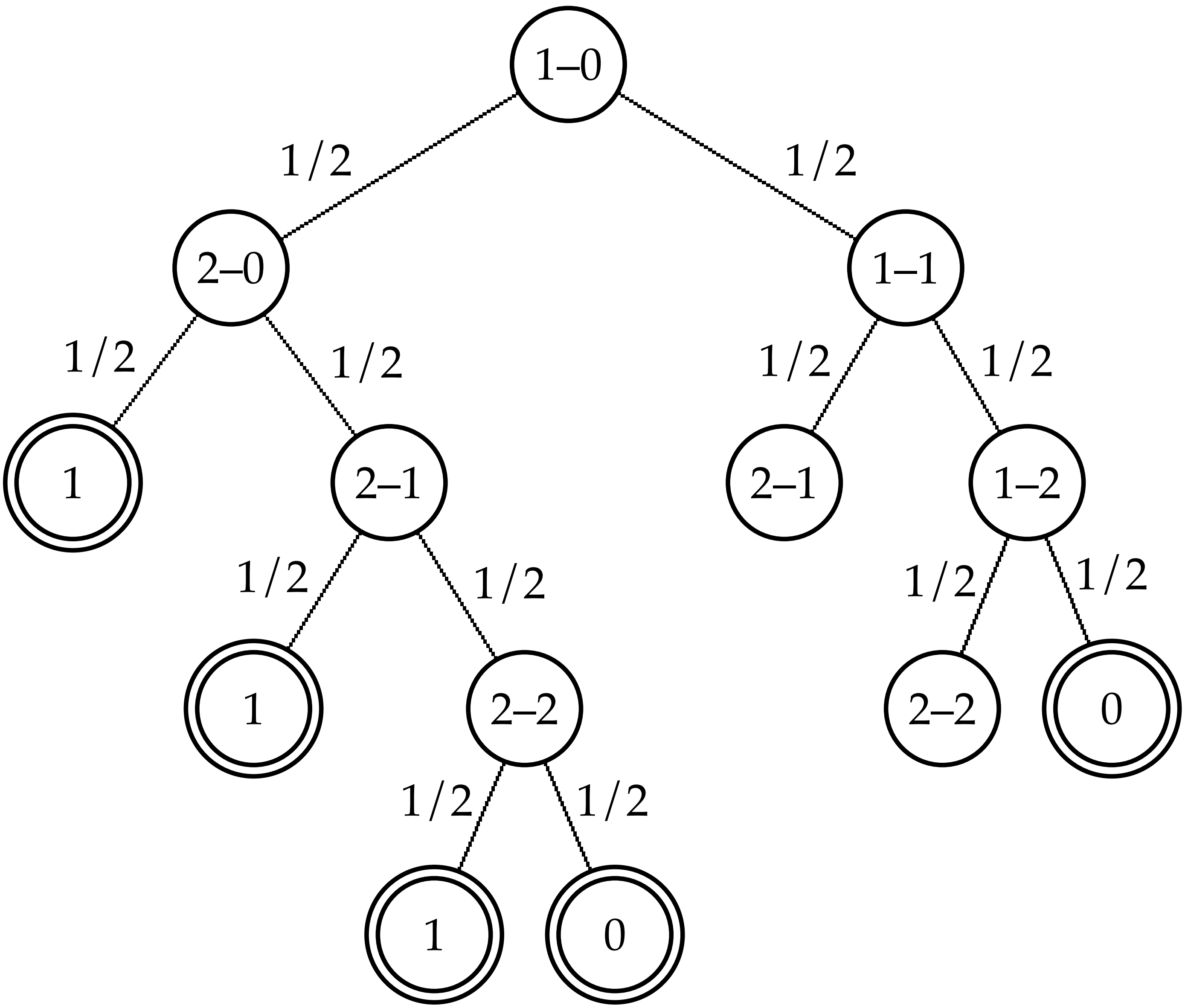
Method: Monte Carlo simulation?
Better here: *dynamic programming*,
using symmetry to reduce size

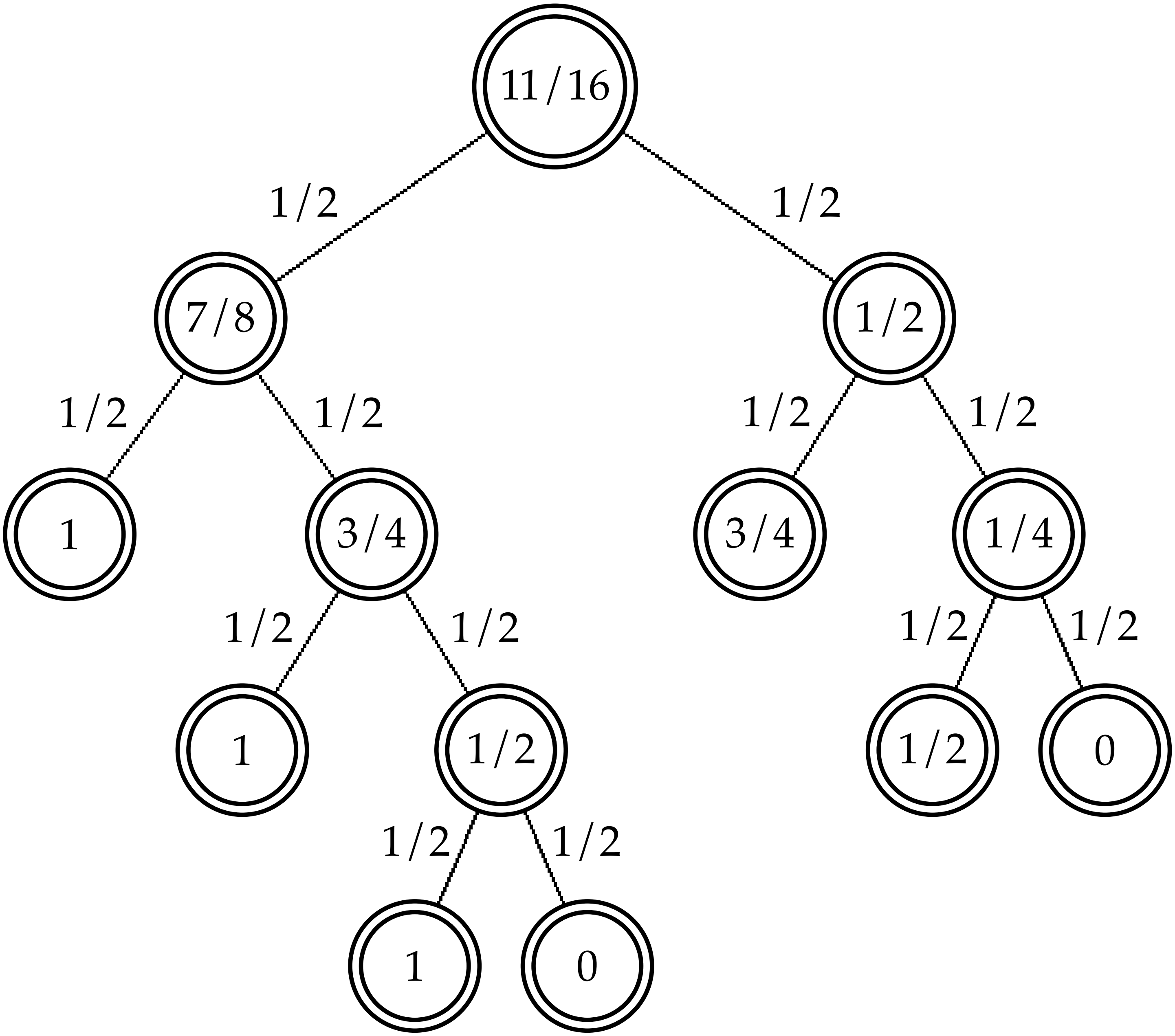
- Expand out along the game tree
- Solve each simple case only once, remember result
- Apply that result as the simple case arises over and over again
- Propagate back the results from simple cases

Coin Flipping

You're ahead 1–0.
What is your chance of winning?




1–0
1/2
1/2
2–0
1–1
1/2
1/2
1/2
1/2
1
2–1
2–1
1–2
1/2
1/2
1/2
1/2
1
2–2
2–2
0
1/2
1/2
1
0


11/16
1/2
1/2
7/8
1/2
1/2
1/2
1/2
1/2
1
3/4
3/4
1/4
1/2
1/2
1/2
1/2
1
1/2
1/2
0
1/2
1/2
1
0

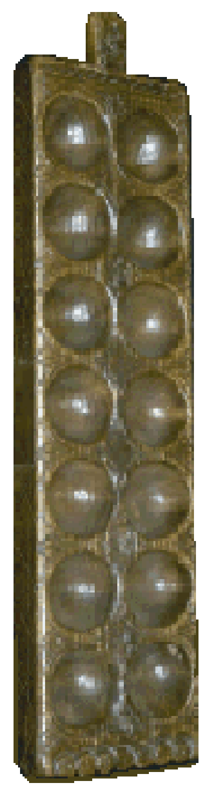
Simplified versions: Pascal's
"problem of the points"

With best play, P(children win) =

$$\frac{\mathbf{256937772392134220582251920344 89}}{\mathbf{50209695374400000000000000000000}} \approx \mathbf{.512}$$

Other Examples of Dynamic Programming

Obstgarten, . . .
Shut the Box



variants on Blackjack
mancala games
football
—point-after-touchdown strategy
—fourth-down strategy (August 2002)

The Wave

1981: first(?) appearance, at
Oakland Athletics game

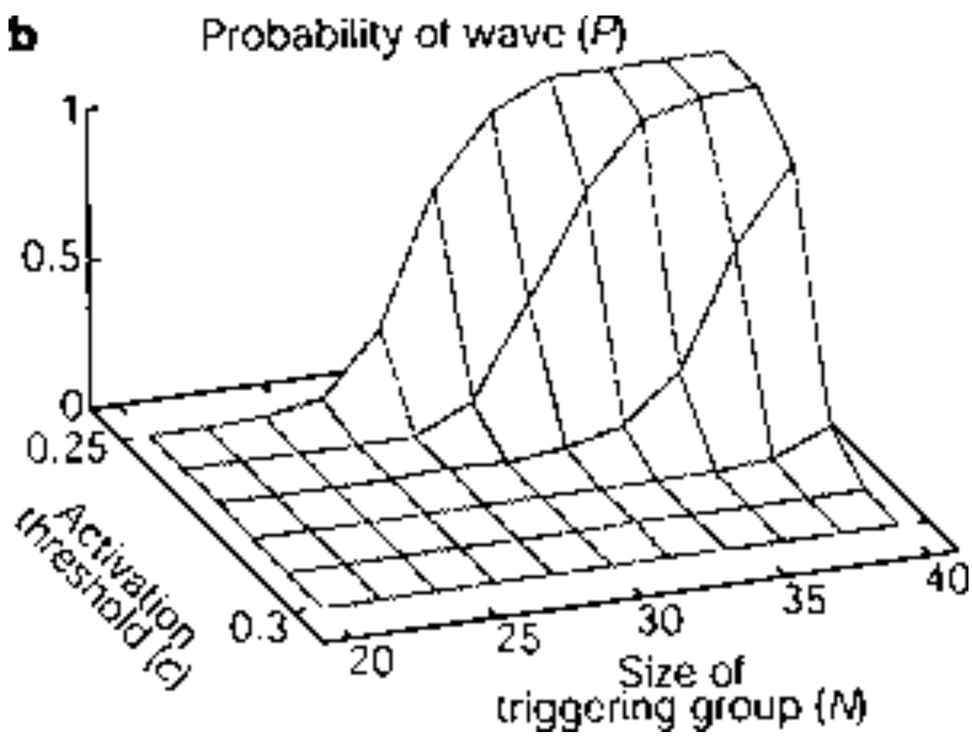
20 seats / sec with ave. width of 15 seats

majority right-handed $\Rightarrow$ clockwise

can apply mathematical models for
—spread of forest fires
—propagation of electrical impulses
in heart tissue

September, 2002: Model of the wave

activation probability
—decreases exponentially with distance
—decreaseslinearly with cos(direction)
(so people in front, on left have more influence)



affiliations of authors:
—Dept. of Biological Physics
—Institute for Economics and Traffic

Baseball and Markov Chains

How much difference does
batting order make?

The *state space* of baseball:
—18 half innings
—in each half inning, 24 states:

$\{0, 1, \text{ or } 2 \text{ outs}\} \times \{8 \text{ base-runner situations}\}$

—for each batter, 12 possible ball counts

Associate probabilities with transitions;
get *Markov chain* model

Different results for different lineups

$9! = 362,880$ lineups of 9 batters

Findings:
—Place best batter 2nd, 3rd, or 4th.
—Either the second-best or third-best
should go just before or just after
the best batter
—Best lineup vs. worst: 4 to 5.5 games
per season.

Markov Chain Theory

—involves matrix algebra, eigenvectors
—determines average duration of game
—determines average time spent in a particular state

Other Examples of Markov Chains

games

—Monopoly

—Chutes and Ladders

—Bingo

—Cootie

number of games in a World Series

. . . plus multitudes of serious applications
(e.g., to social mobility)

SEEK mathematics . . . :

- serious mathematics
  - mathematical organizations
  - general press
  - mathematics indexes
  - mathematics repositories
  - 2002 examples:
    - Catalan conjecture
    - primality

How to Stay in Touch

—passively: mathematics comes to you

—math'l organizations and their publ'ns
Math'l Ass'n of America, Amer. Math. Soc.,
Ntl Council of Tchrs of Math,
Soc. Ind'l and Appl. Math. (SIAM)

—general press and general science press
your local paper, *New York Times*, *Discover*,
Science, *Scientific American*,
Science News, *Nature*

—more actively, for serendipity:

—"columns" at Web sites of MAA, etc.

—when you're hungry for more:

—books in the *library*
—mathematics indexes
MathSciNet, MathDI Database, Zentralblatt
—mathematics repositories
Mathematics ArXiv
JSTOR

Heuristic for Finding Articles

"Ask, and you shall be answered;
seek, and you shall find;
knock, and it shall be opened unto you"

–Sermon on the Mount

1. Input author and topic into
—general search engine (Google, NYT, . . .)
—your library's electronic journals index
—mathematics indexes
—mathematics repositories

2. Follow leads to
—article summary or full text
—reference to holdings in local libraries
—author's home page
—journal's home page

3. If no success:
—email / write / fax / call author, locating from home pages, *Combined Membership List*, *World Directory of Mathematicians*
—consult a librarian

Mathematical Advances in 2002:

Polynomial-Time Primality Test

Why important?
contemporary cryptography requires primes with *hundreds* of digits

Testing whether an integer n is prime:
—brute force: try every prime $\leq \sqrt{n}$; takes time exponential in $\log n$ (the number of digits)—
i.e., time proportional to n
notation: $\mathcal{O}(n)$, "big-oh of n"
—ideal: polynomial-time algorithm "in $\mathcal{P}$"
in the number of digits: $\mathcal{O}\big((\log n)^k\big)$
—conditional algorithms: "If RH is true . . . "
—used in practice: probabilistic algorithms;
fast, but could be wrong

August, 2002: Agrawal with two undergrads
(Indian Institute of Technology)
show in brief paper that "PRIMES is in $\mathcal{P}$":
almost $\mathcal{O}\left((\log n)^{12}\right)$, cond'ly $\mathcal{O}\left((\log n)^{6}\right)$

Speed up their algorithm, find faster ones?

Mathematical Advances in 2002:

Catalan's Conjecture Proved

Why important?
it's not, except to alleviate curiosity

Eugène Charles Catalan (1814–1894) conjectured in 1844 that:
the only solution to $x^m - y^n = 1$
in integers $x, y, m, n \geq 2$ is $3^2 - 2^3 = 1$

Main approach to both Catalan's conjecture and Fermat's Last Theorem *abstract algebra*
—Ernst Kummer (1810–1893) used algebra to prove FLT for various classes of exponents

—20th century: new paradigm; both problems lead to elliptic curves (algebraic geometry)
—1995: elliptic curves are the key to FLT . . .
— . . . but are dead end for Catalan's conjecture

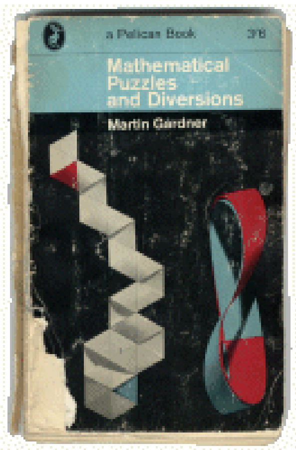
April, 2002: Preda Mihailescu (University of Paderborn, Germany) proves conjecture by going back to Kummer's algebraic theory

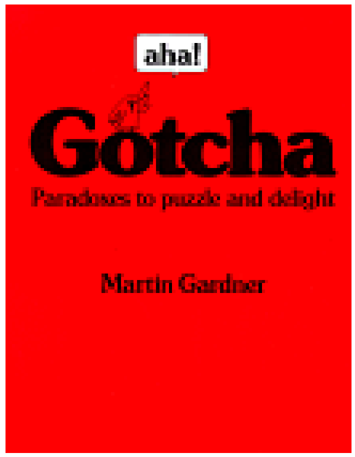
SEEK mathematics . . . :

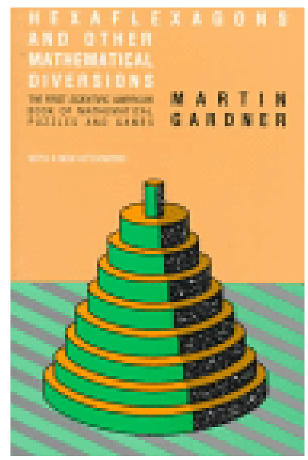
- fun mathematics
  - Martin Gardner:
    books, indexes
  - other books on recreational math
  - current sources

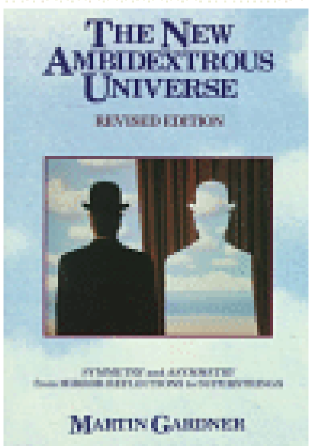
Mathematics for Fun

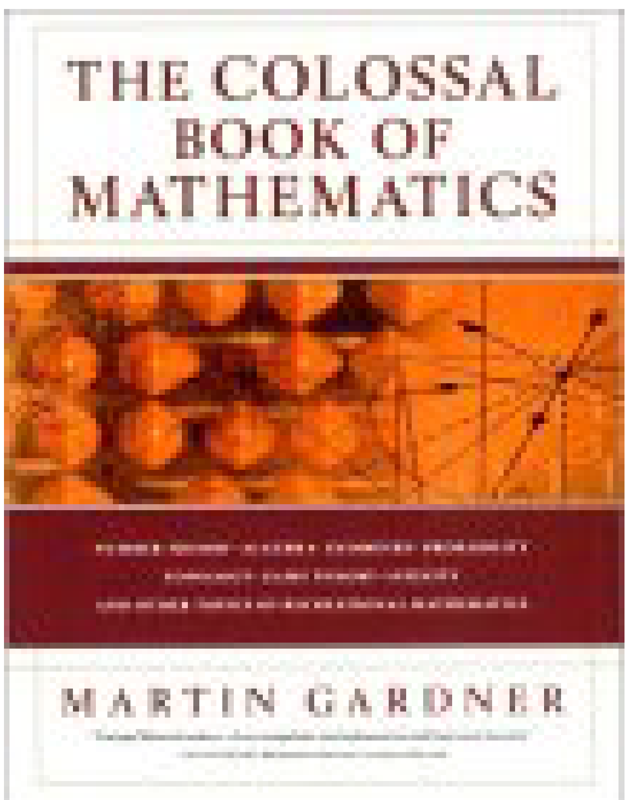
—books by Martin Gardner











Sample Research Project: *Optimizing Bingo*

B	I	N	G	O
12	17	35	59	63
2	24	41	51	72
15	16	Free	60	67
10	20	45	57	69
9	25	43	50	61

- SEEing Bingo mathematically
  - P(card wins)
  - P(card wins on draw k)
  - average duration: E [draws until a win when n cards in play]
  - P(more than one winner when n cards in play)

- SEEKing Bingo
  - at the public library
  - on the Web
    - history of Bingo
    - practical questions
    - mathematical research questions
  - at the bingo hall
  - in the math literature

Mathematical Questions

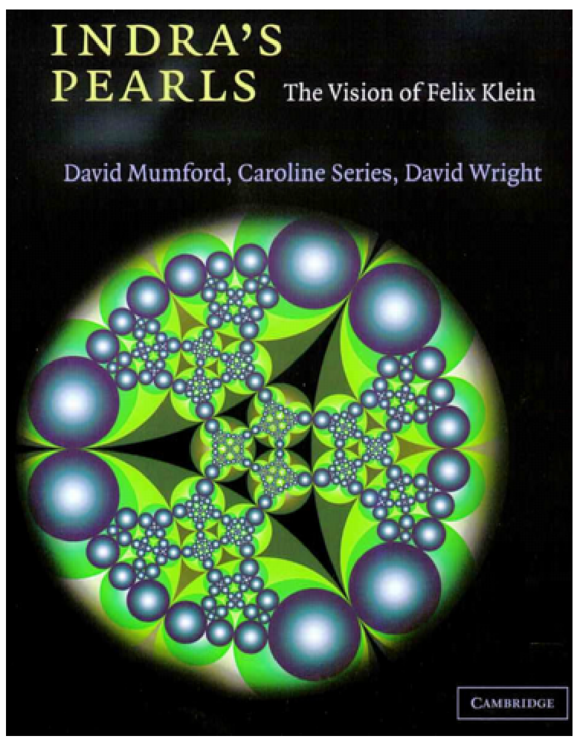
- How large can a set of cards be that always has a single winner?
- How can you construct such a set?
- With actual cards used, what is the probability of a single winner with n cards in play?

Some Approaches

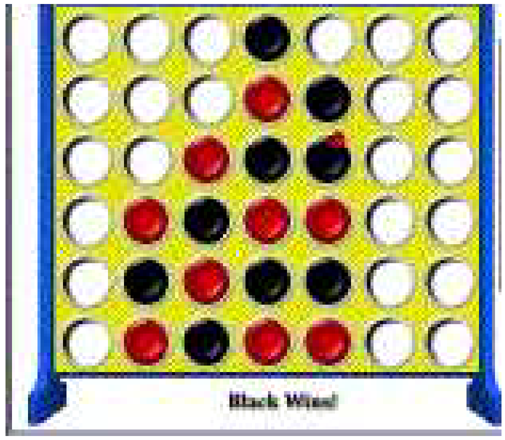
- Heuristic: "Solve a simpler problem": 3×3 bingo
- Heuristic: "Translate the problem into various branches of mathematics"
  - geom.: intersecting lines?
  - algebra: equations?
  - comb.: Latin squares?

SPEAK mathematics . . . :

- don't keep what you see to yourself!
- don't talk technical
- get a good coffee-table math book



How to Make Big Money from Math



Theorem (Allis 1988, master's thesis).
The first player in Connect-4 can always win.

1. Learn the strategy (and teach it to your computer).
2. Learn German.
3. Travel to Augsburg, Germany.
4. Tune in the Munich TV station that lets you play the game for thousands of dollars.
5. Dial in fast and often (use your computer) (costs $1 per call).
6. Be lucky and get through.
7. You get to play first, so beat the host.
8. Stay on the line to tell them where to send the money.

Accompanying Handout

`ftp://cs.beloit.edu/math-cs/Faculty/`

`Paul%20Campbell/Public/Schraut/`