Asides on Curious and Interesting Things

Complex Networks | @networksvox
CSYS/MATH 303, Spring, 2016

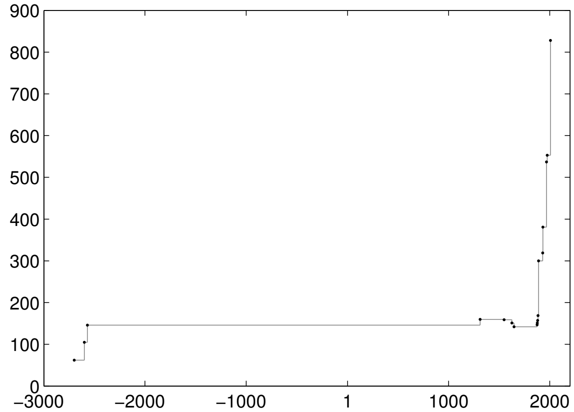
Prof. Peter Dodds | @peterdodds

Dept. of Mathematics & Statistics | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

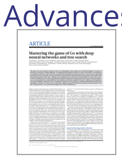
These slides are brought to you by:

Outline

What's this?



Advances in sociotechnical algorithms:



"Mastering the game of Go with deep neural networks and tree search"
Silver and Silver,
Nature, **529**, 484–489, 2016. [6]

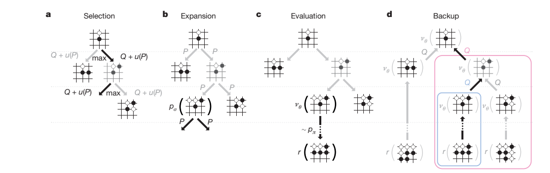
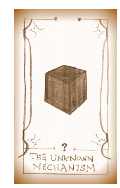
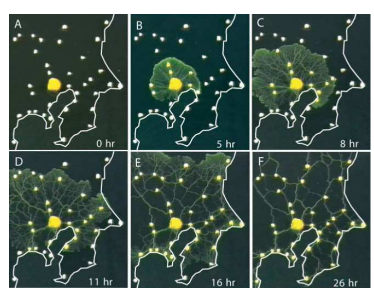


Figure 3 | Monte Carlo tree search in AlphaGo. a, Each simulation traverses the tree by selecting the edge with maximum action value Q, plus a bonus $u(P)$ that depends on a stored prior probability P for that edge. **b**, The leaf node may be expanded; the new node is processed once by the policy network p_σ and the output probabilities are stored as prior probabilities P for each action. **c**, At the end of a simulation, the leaf node is evaluated in two ways: using the value network v_θ and by running a rollout to the end of the game with the fast rollout policy p_π, then computing the winner with function r. **d**, Action values Q are updated to track the mean value of all evaluations $r(\cdot)$ and $v_\theta(\cdot)$ in the subtree below that action.

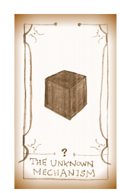
- Nature News (2016): Digital Intuition
- Wired (2012): Network Science of the game of Go



"Rules for Biologically Inspired Adaptive Network Design"
Tero et al.,
Science, **327**, 439-442, 2010. [7]



Urban deslime in action:
https://www.youtube.com/watch?v=GwKuFREOgmo



"Citations to articles citing Benford's law: A Benford analysis"
Tariq Ahmad Mir,
Preprint available at
http://arxiv.org/abs/1602.01205,
2016. [4]

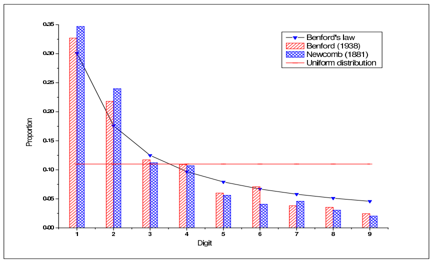
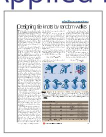


Fig. 1: The observed proportions of first digits of citations received by the articles citing FB and SN on September 30, 2012. For comparison the proportions expected from BL and uniform distributions are also shown.

Applied knot theory:



"Designing tie knots by random walks"
Fink and Mao,
Nature, **398**, 31–32, 1999. [1]

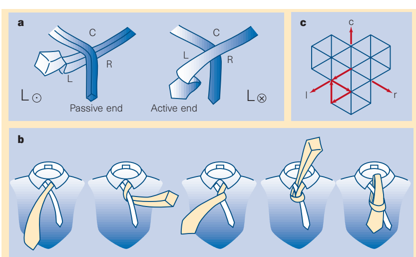
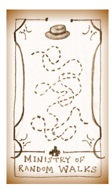


Figure 1 All diagrams are drawn in the frame of reference of the mirror image of the actual tie. **a**, The two ways of beginning a knot, $L_\odot$ and $L_\otimes$. For knots beginning with $L_\otimes$, the tie must begin inside-out. **b**, The four-in-hand, denoted by the sequence $L_\otimes R_\odot L_\otimes C_\odot T$. **c**, A knot may be represented by a persistent random walk on a triangular lattice. The example shown is the four-in-hand, indicated by the walk ↑↓↑↓.



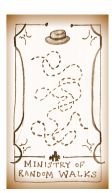
Applied knot theory:

Table 1 **Aesthetic tie knots**

h	γ	γ/h	$K(h,\gamma)$	s	b	Name	Sequence
3	1	0.33	1	0	0		$L_\odot R_\otimes C_\odot T$
4	1	0.25	1	−1	1	Four-in-hand	$L_\otimes R_\odot L_\otimes C_\odot T$
5	2	0.40	2	−1	0	Pratt knot	$L_\odot C_\otimes R_\odot L_\otimes C_\odot T$
6	2	0.33	4	0	0	Half-Windsor	$L_\otimes R_\odot C_\otimes L_\odot R_\otimes C_\odot T$
7	2	0.29	6	−1	1		$L_\odot R_\otimes L_\odot C_\otimes R_\odot L_\otimes C_\odot T$
7	3	0.43	4	0	1		$L_\odot C_\otimes R_\odot C_\otimes L_\odot R_\otimes C_\odot T$
8	2	0.25	8	0	2		$L_\otimes R_\odot L_\otimes C_\odot R_\otimes L_\odot R_\otimes C_\odot T$
8	3	0.38	12	−1	0	Windsor	$L_\otimes C_\odot R_\otimes L_\odot C_\otimes R_\odot L_\otimes C_\odot T$
9	3	0.33	24	0	0		$L_\odot R_\otimes C_\odot L_\otimes R_\odot C_\otimes L_\odot R_\otimes C_\odot T$
9	4	0.44	8	−1	2		$L_\odot C_\otimes R_\odot C_\otimes L_\odot C_\otimes R_\odot L_\otimes C_\odot T$

Knots are characterized by half-winding number h, centre number γ, centre fraction γ/h, knots per class $K(h,\gamma)$, symmetry s, balance b, name and sequence.

- h = number of moves
- γ = number of center moves
- $K(h,\gamma) = 2^{\gamma-1}\binom{h-\gamma-2}{\gamma-1}$
- $s = \sum_{i=1}^{h} x_i$ where x = -1 for L and +1 for R.
- $b = \frac{1}{2}\sum_{i=2}^{h-1} |\omega_i + \omega_{i-1}|$ where $\omega = \pm 1$ represents winding direction.

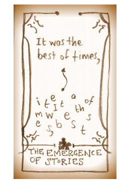


Irregular verbs

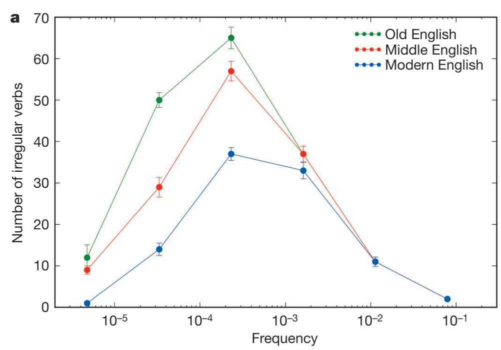
Cleaning up the code that is English:

"Quantifying the evolutionary dynamics of language"
Lieberman et al.,
Nature, **449**, 713–716, 2007. [2]

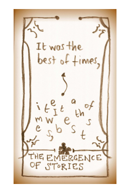
- Exploration of how verbs with irregular conjugation gradually become regular over time.
- Comparison of verb behavior in Old, Middle, and Modern English.



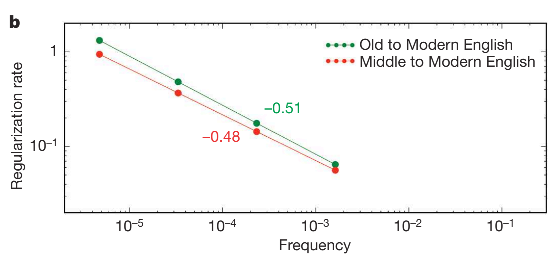
Irregular verbs



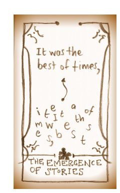
- Universal tendency towards regular conjugation
- Rare verbs tend to be regular in the first place



Irregular verbs



- Rates are relative.
- The more common a verb is, the more resilient it is to change.



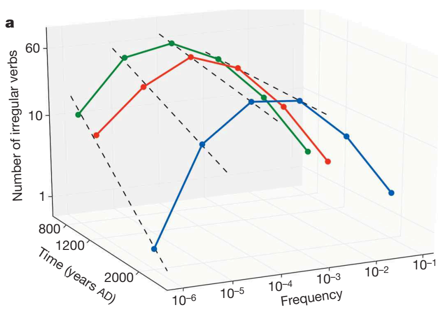
Irregular verbs

Table 1 | The 177 irregular verbs studied

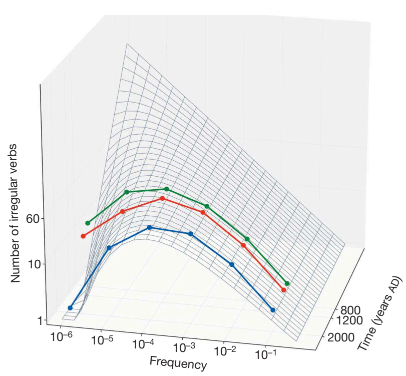
Frequency	Verbs	Regularization (%)	Half-life (yr)
10^{-1}–1	be, have	0	38,800
10^{-2}–10^{-1}	come, do, find, get, give, go, know, say, see, take, think	0	14,400
10^{-3}–10^{-2}	begin, break, bring, buy, choose, draw, drink, drive, eat, fall, fight, forget, grow, hang, help, hold, leave, let, lie, lose, reach, rise, run, seek, set, shake, sit, sleep, speak, stand, teach, throw, understand, walk, win, work, write	10	5,400
10^{-4}–10^{-3}	arise, bake, bear, beat, bind, bite, blow, bow, burn, burst, carve, chew, climb, cling, creep, dare, dig, drag, flee, float, flow, fly, fold, freeze, grind, leap, lend, lock, melt, reckon, ride, rush, shape, shine, shoot, shrink, sigh, sing, sink, slide, slip, smoke, spin, spring, starve, steal, step, stretch, strike, stroke, suck, swallow, swear, sweep, swim, swing, tear, wake, wash, weave, weep, weigh, wind, yell, yield	43	2,000
10^{-5}–10^{-4}	bark, bellow, bid, blend, braid, brew, cleave, cringe, crow, dive, drip, fare, fret, glide, gnaw, grip, heave, knead, low, milk, mourn, mow, prescribe, redden, reek, row, scrape, seethe, shear, shed, shove, slay, slit, smite, sow, span, spurn, sting, stink, strew, stride, swell, tread, uproot, wade, warp, wax, wield, wring, writhe	72	700
10^{-6}–10^{-5}	bide, chide, delve, flay, hew, rue, shrive, slink, snip, spew, sup, wreak	91	300

177 Old English irregular verbs were compiled for this study. These are arranged according to frequency bin, and in alphabetical order within each bin. Also shown is the percentage of verbs in each bin that have regularized. The half-life is shown in years. Verbs that have regularized are indicated in red. As we move down the list, an increasingly large fraction of the verbs are red; the frequency-dependent regularization of irregular verbs becomes immediately apparent.

- Red = regularized
- Estimates of half-life for regularization ($\propto f^{1/2}$)

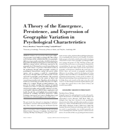


- 'Wed' is next to go.
- -ed is the winning rule...
- But 'snuck' is sneaking up on sneaked.[3]



- Projecting back in time to proto-Zipf story of many tools.

Personality distributions:



"A Theory of the Emergence, Persistence, and Expression of Geographic Variation in Psychological Characteristics"
Rentfrow, Gosling, and Potter,
Perspectives on Psychological Science, **3**, 339–369, 2008. [5]

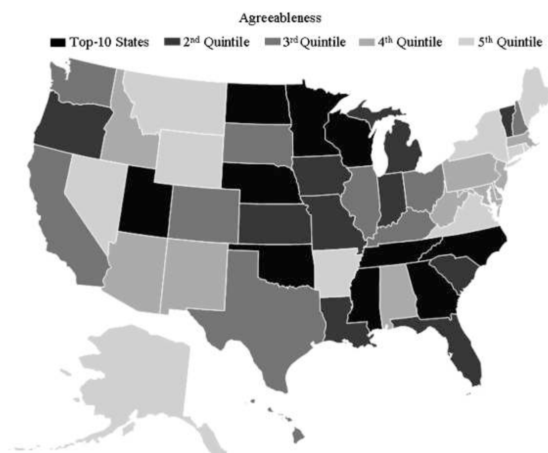
Five Factor Model (FFM):

- Extraversion [E]
- Agreeableness [A]
- Conscientiousness [C]
- Neuroticism [N]
- Openness [O]

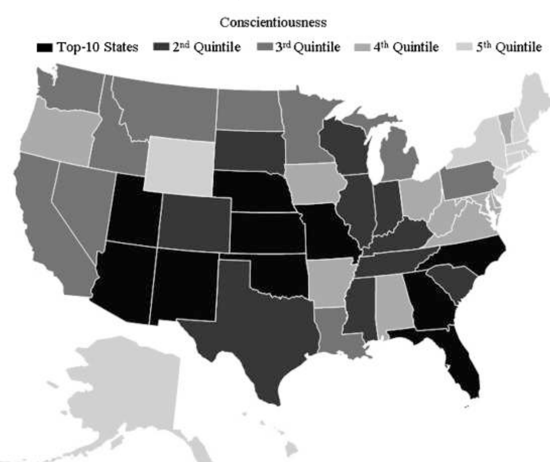
"...a robust and widely accepted framework for conceptualizing the structure of personality... Although the FFM is not universally accepted in the field..." [5]

A concern: self-reported data.

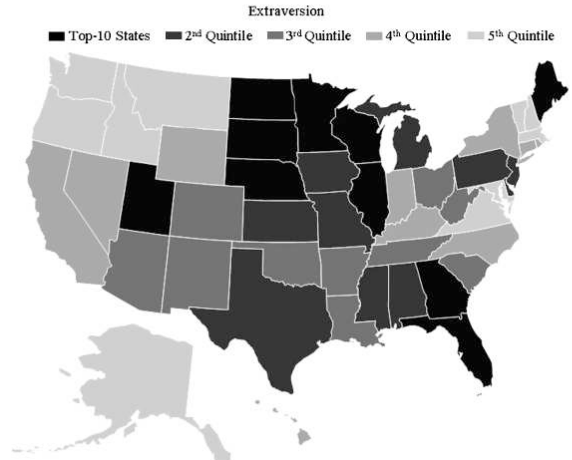
Agreeableness:



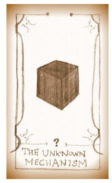
Conscientiousness:



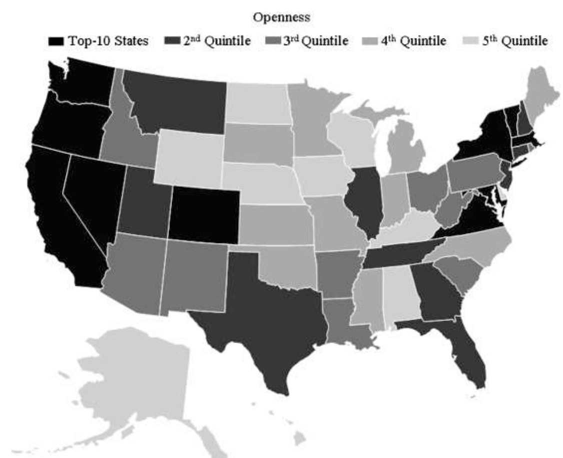
Extraversion:



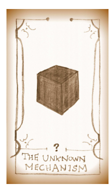
Random
Randomness
References



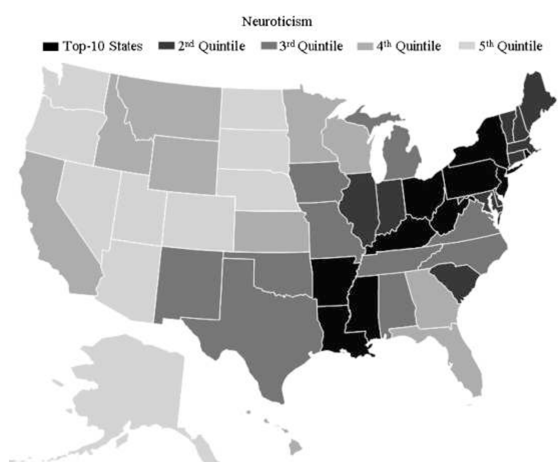
Openness



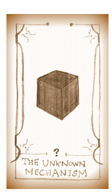
Random
Randomness
References



Neuroticism:

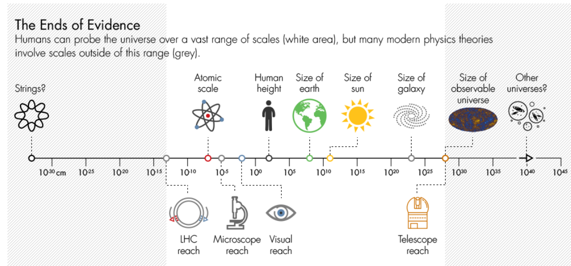


Random
Randomness
References

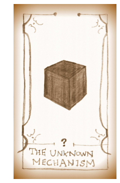


Limits of testability and happiness in Science:

From A Fight for the soul of Science in Quanta Magazine (2016/02):



Random
Randomness
References



Europe:

- Many errors called out in comments. Why hasn't this been done well?

Random
Randomness
References

John Conway's Doomsday rule for determining a date's day of the week:

Doomsdays for the Gregorian calendar

Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.
1898	1899	1900	1901	1902	1903	→	1904	1905	1906	1907	→	1908	1909
1910	1911	→	1912	1913	1914	1915	→	1916	1917	1918	1919	→	1920
1921	1922	1923	→	1924	1925	1926	1927	→	1928	1929	1930	1931	→
1932	1933	1934	1935	→	1936	1937	1938	1939	→	1940	1941	1942	1943
→	1944	1945	1946	1947	→	1948	1949	1950	1951	→	1952	1953	1954
1955	→	1956	1957	1958	1959	→	1960	1961	1962	1963	→	1964	1965
1966	1967	→	1968	1969	1970	1971	→	1972	1973	1974	1975	→	1976
1977	1978	1979	→	1980	1981	1982	1983	→	1984	1985	1986	1987	→
1988	1989	1990	1991	→	1992	1993	1994	1995	→	1996	1997	1998	1999
→	2000	2001	2002	2003	→	2004	2005	2006	2007	→	2008	2009	2010
2011	→	2012	2013	2014	2015	→	2016	2017	2018	2019	→	2020	2021
2022	2023	→	2024	2025	2026	2027	→	2028	2029	2030	2031	→	2032
2033	2034	2035	→	2036	2037	2038	2039	→	2040	2041	2042	2043	→
2044	2045	2046	2047	→	2048	2049	2050	2051	→	2052	2053	2054	2055
→	2056	2057	2058	2059	→	2060	2061	2062	2063	→	2064	2065	2066
2067	→	2068	2069	2070	2071	→	2072	2073	2074	2075	→	2076	2077
2078	2079	→	2080	2081	2082	2083	→	2084	2085	2086	2087	→	2088
2089	2090	2091	→	2092	2093	2094	2095	→	2096	2097	2098	2099	2100

- Works for Gregorian (1582–, haphazardly) and the increasingly inaccurate Julian calendars (400 and 28 years cycles).
- Apparently inspired by Lewis Carroll's work on a perpetual calendar.

Random
Randomness
References

Outline:

- Determine “anchor day” for a given century, then find Doomsday for a given year in that century.
- Remember special Doomsday dates and work from there.
- Naturally: Load this year’s Doomsday into brain.

Century’s anchor day (Gregorian, Sunday $\equiv$ 0):

$$5 \times \left(\left\lfloor \frac{YYYY}{100} \right\rfloor \mod 4 \right) \mod 7 + \text{Tuesday}$$

Offset:

$$\left(365YY + \left\lfloor \frac{YY}{4} \right\rfloor \right) \mod 7 = \left(YY + \left\lfloor \frac{YY}{4} \right\rfloor \right) \mod 7$$

Memorable Doomsdays:

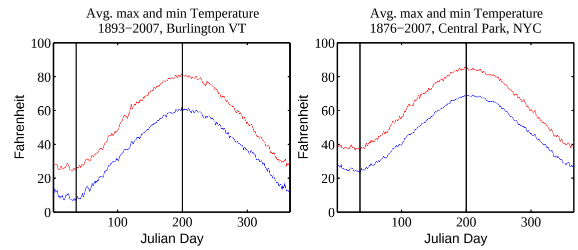
Month	Memorable date	Month/Day	Mnemonic[6]
January	January 3 (common years), January 4 (leap years)	1/3 or 1/4	the 3rd **3** years in 4 and the 4th in the **4**th
February	February 28 (common years), February 29 (leap years)	2/28 or 2/29	last day of February
March	"March 0"	3/0	last day of February
April	April 4	4/4	**4/4**, 6/6, 8/8, 10/10, 12/12
May	May 9	5/9	**9-to-5** at 7-11
June	June 6	6/6	4/4, **6/6**, 8/8, 10/10, 12/12
July	July 11	7/11	9-to-5 at **7-11**
August	August 8	8/8	4/4, 6/6, **8/8**, 10/10, 12/12
September	September 5	9/5	**9-to-5** at 7-11
October	October 10	10/10	4/4, 6/6, 8/8, **10/10**, 12/12
November	November 7	11/7	9-to-5 at **7-11**
December	December 12	12/12	4/4, 6/6, 8/8, 10/10, **12/12**

- Pi day (March 14), July 4, Halloween, and Boxing Day are always Doomsdays.

The bissextile year

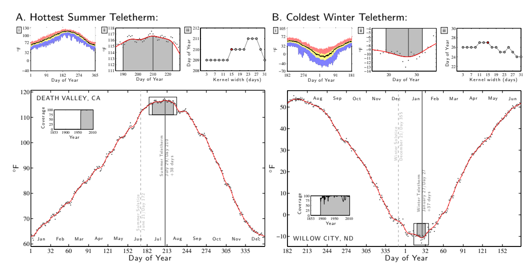
“The Julian calendar, which was developed in 46 BC by Julius Caesar, and became effective in 45 BC, distributed an extra ten days among the months of the Roman Republican calendar. Caesar also replaced the intercalary month by a single intercalary day, located where the intercalary month used to be. To create the intercalary day, the existing ante diem sextum Kalendas Martias (February 24) was doubled, producing ante diem bis sextum Kalendas Martias. Hence, the year containing the doubled day was a bissextile (bis sextum, “twice sixth”) year. For legal purposes, the two days of the bis sextum were considered to be a single day, with the second half being intercalated; but in common practice by 238, when Censorinus wrote, the intercalary day was followed by the last five days of February, a. d. VI, V, IV, III and pridie Kal. Mart. (the days numbered 24, 25, 26, 27, and 28 from the beginning of February in a common year), so that the intercalated day was the first half of the doubled day. Thus the intercalated day was effectively inserted between the 23rd and 24th days of February.”

The Teletherm, an early conception:

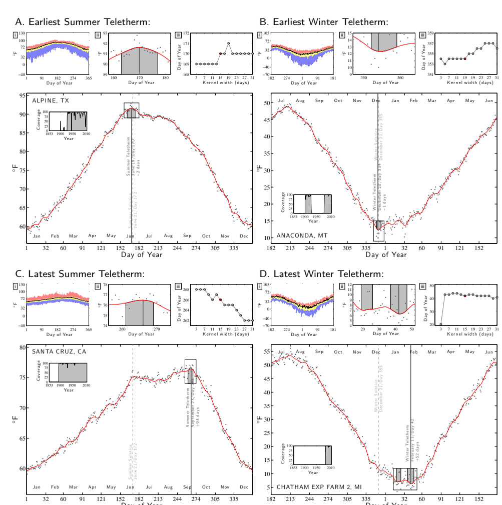


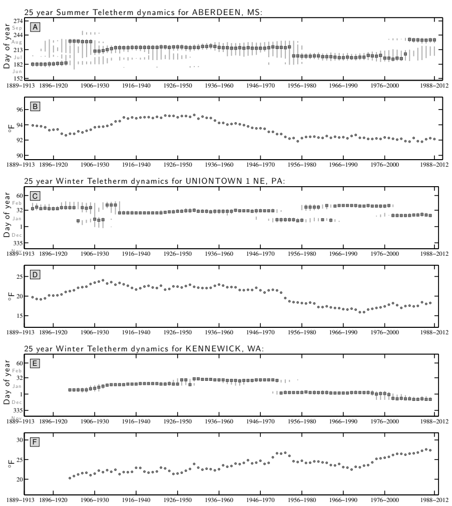
- Hibernal Teletherm $\approx$ February 4.
- Halfway between Winter Solstice and Spring Equinox
- Bonus: Groundhog Day, Imbolc, ...
- Aesteval Teletherm $\approx$ July 19 (164 days later).

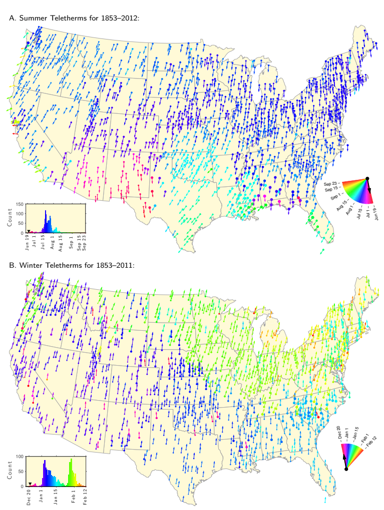
In review: “Tracking the Teletherms: The spatiotemporal dynamics of the hottest and coldest days of the year”, Dodds, Mitchell, Reagan, and Danforth.

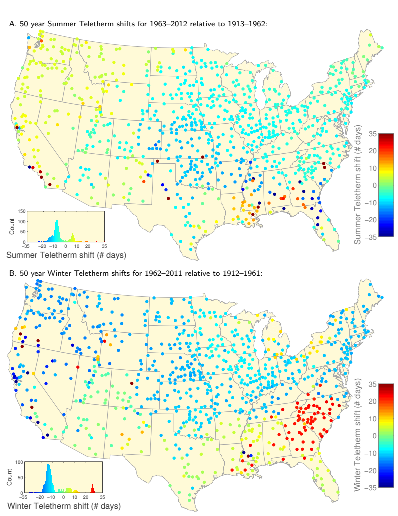


- 2×1218 similar figures for the US.
- 6000ish pages of Supplementary Information (all figures)
- Interactive website.





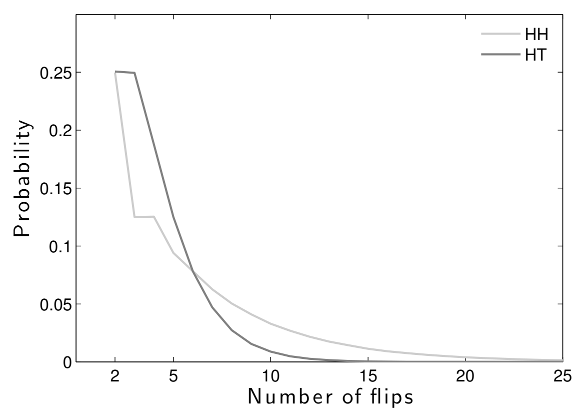




Homo nonprobabilisticus, continued:

- Important detour: The final digits of primes are not entirely random (how did we not know this?).
- Start flipping a coin ...
- Two tosses: What are the probabilities of flipping (1) HH and (2) HT?
- Flip a coin $n \geq 2$ times: What are the probabilities that the last two tosses are (1) HH or (2) HT?
- Estimate: On average, how many flips does it take to first see the sequence HT?
- Estimate: On average, how many flips does it take to first see the sequence HH?
- What's the probability of first flipping a HT sequence on the $n-1$th and nth flips?
- What's the probability of first flipping two heads in a row (HH) on the $(n-1)$th and nth flips?

Homo nonprobabilisticus, continued:



Average number of flips: 4 and 6.

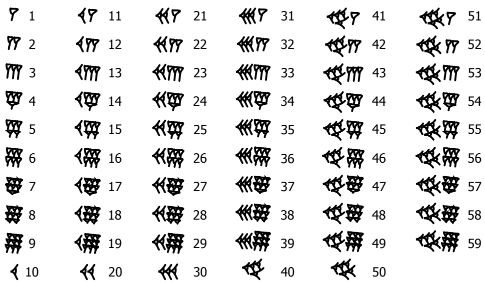
Universal numbers

From here.

- Accidents of evolution[1] give us 5 + 5 = 10 fingers and hence base 10.
- We could be happy with base 6, 8, 12, ...
- We like these:
  - 60 seconds in a minute
  - 60 minutes in an hour.
  - 2 × 12 = 24 hours in a day.
  - 360 degrees in a circle.

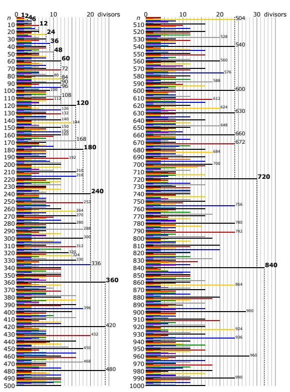
[1] Maybe 5 fingers are not an accident

We've liked these kinds of numbers for a long time:



- 2000 BC: Babylonian base 60/Sexagesimal system.
- Other bases (or radices): 2, 10, 12 (duodecimal/dozenal), 6 (senary), 8, 16, 20 (vigesimal), 60.

Highly composite numbers:



- HCN = natural number with more divisors than any smaller natural number.
- 2, 4, 6, 12, 24, 36, 48, 60, 120, 180, 240, 360, 720, 840, 1260, 1680, 2520, 5040 (Plato's optimal city population), ...
- OEIS sequence A002182

Superior highly composite numbers:

# prime factors	SHCN n	prime factorization	prime exponents	# divisors d(n)		primorial factorization
1	2	2	1	2	2	2
2	6	$2 \cdot 3$	1,1	2^2	4	6
3	12	$2^2 \cdot 3$	2,1	3x2	6	$2 \cdot 6$
4	60	$2^2 \cdot 3 \cdot 5$	2,1,1	3×2^2	12	$2 \cdot 30$
5	120	$2^3 \cdot 3 \cdot 5$	3,1,1	4×2^2	16	$2^2 \cdot 30$
6	360	$2^3 \cdot 3^2 \cdot 5$	3,2,1	4x3x2	24	$2 \cdot 6 \cdot 30$
7	2520	$2^3 \cdot 3^2 \cdot 5 \cdot 7$	3,2,1,1	$4\times3\times2^2$	48	$2 \cdot 6 \cdot 210$
8	5040	$2^4 \cdot 3^2 \cdot 5 \cdot 7$	4,2,1,1	$5\times3\times2^2$	60	$2^2 \cdot 6 \cdot 210$
9	55440	$2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$	4,2,1,1,1	$5\times3\times2^3$	120	$2^2 \cdot 6 \cdot 2310$
10	720720	$2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13$	4,2,1,1,1,1	$5\times3\times2^4$	240	$2^2 \cdot 6 \cdot 30030$

- SHCN = natural number n whose number of divisors exceeds that of any other number when scaled relative to itself in a sneaky way:

$$\frac{d(n)}{n^\epsilon} \geq \frac{d(j)}{j^\epsilon} \text{ and } \frac{d(n)}{n^\epsilon} > \frac{d(k)}{k^\epsilon}$$

for $j < n < k$ and some $\epsilon > 0$.

There's more: Superabundant numbers

- n is superabundant if:

$$\frac{\sigma_1(n)}{n} > \frac{\sigma_1(j)}{j}$$

for $j < n$ and where $\sigma_x(n) = \sum_{d|n} d^x$ is the divisor function.

- 449 numbers are both superabundant and highly composite.

Yet more: Colossally abundant numbers:

- n is colossally abundant if for all j and some $\epsilon > 0$:

$$\frac{\sigma_1(n)}{n^{1+\epsilon}} \geq \frac{\sigma_1(j)}{j^{1+\epsilon}}$$

- Infinitely many but only 22 less than 10^{18}.

Some very, very silly units of measurement courtesy of the Imperial system:

- 22 yards in a chain = 1 cricket pitch, 100 links in a chain, 10 chains in a furlong, 80 chains in a mile.
- 1 acre = 1 furlong $\times$ 1 chain = 43,560 square feet.
- 160 fluid ounces in a gallon.
- 14 pounds in a stone.
- Hundredweight = 112 pounds.

Also:

- Fahrenheit, Celcius, and Kelvin.
- The entire metric system.

Training with stories as fuel:

Randomness:

References I

[1] T. M. Fink and Y. Mao.
Designing tie knots by random walks.
Nature, 398:31–32, 1999. pdf

[2] E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, and M. A. Nowak.
Quantifying the evolutionary dynamics of language.
Nature, 449:713–716, 2007. pdf

[3] J.-B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, T. G. B. Team, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M. A. Nowak, and E. A. Lieberman.
Quantitative analysis of culture using millions of digitized books.
Science Magazine, 2010. pdf

References II

[4] T. A. Mir.
Citations to articles citing Benford's law: A Benford analysis, 2016.
Preprint available at `http://arxiv.org/abs/1602.01205`. pdf

[5] P. J. Rentfrow, S. D. Gosling, and J. Potter.
A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics.
Perspectives on Psychological Science, 3:339–369, 2008. pdf

[6] D. Silver et al.
Mastering the game of Go with deep neural networks and tree search.
Nature, 529:484–489, 2016. pdf

References III

[7] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki.
Rules for biologically inspired adaptive network design.
Science, 327(5964):439–442, 2010. pdf