Complexity, Big Data Science, and Happiness

Discrete Days, St. Michael's College, 2011

Peter Dodds

Department of Mathematics & Statistics
Center for Complex Systems
Vermont Advanced Computing Center
University of Vermont

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Outline

Complexity

- Introduction
- Emergence
- Universality
- Symmetry Breaking
- The Big Theory
- Revolution: Big Data & Complex Networks
- Nutshell

Measuring Happiness

- Tweetage
- Mechanical Turk

References

Definitions

A meaningful definition of a Complex System:

- Distributed possibly networked system of many interrelated parts with no centralized control exhibiting emergent behavior—'More is Different' [2]

A few optional features:

- Nonlinear relationships
- Presence of feedback loops
- Being open or driven
- Presence of memory
- Modular (nested)/multiscale structure
- Opaque boundaries

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Examples of Complex Systems:

- human societies
- cells
- organisms
- power systems
- weather systems
- ecosystems
- animal societies
- disease ecologies
- brains
- social insects
- geophysical systems
- the world wide web

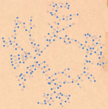
- i.e., everything that's interesting...

Relevant fields:

- Physics
- Economics
- Sociology
- Psychology
- Information Sciences
- Cognitive Sciences
- Biology
- Ecology
- Geosciences
- Geography
- Medical Sciences
- Systems Engineering
- Computer Science
- ...

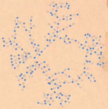
- i.e., everything that's interesting...

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

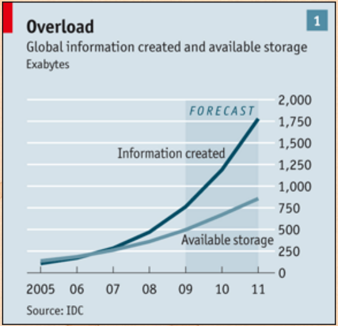


Complexity Manifesto:

1. Systems are ubiquitous and systems matter.
2. Consequently, much of science is about understanding how pieces dynamically fit together.
3. 1700 to 2000 = Golden Age of Reductionism.
   - Atoms!, sub-atomic particles, DNA, genes, people, ...
4. Understanding and creating systems (including new 'atoms') is the greater part of science and engineering.
5. Universality: systems with quantitatively different micro details exhibit qualitatively similar macro behavior.
6. Computing advances make the Science of Complexity possible:
   6.1 We can measure and record enormous amounts of data, research areas continue to transition from data scarce to data rich.
   6.2 We can simulate, model, and create complex systems in extraordinary detail.



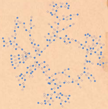
Data, Data, Everywhere—the Economist, Feb 25, 2010 (⊞)



- Exponential growth: $\sim$ 60% per year.

Big Data Science:

- 2013: year traffic on Internet estimate to reach 2/3 Zettabytes ($1ZB = 10^3 EB = 10^6 PB = 10^9 TB$)
- Large Hadron Collider: 40 TB/second.
- 2016—Large Synoptic Survey Telescope: 140 TB every 5 days.
- Facebook: $\sim$ 100 billion photos
- Twitter: $\sim$ 5 billion tweets



No really, that's a lot of data

Data inflation

2

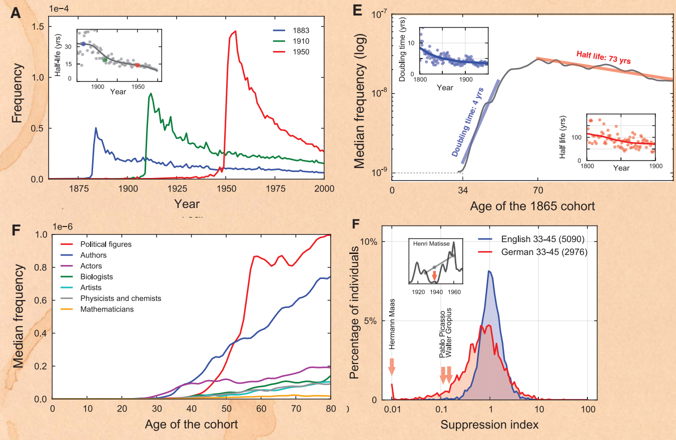
Unit	Size	What it means
Bit (b)	1 or 0	Short for "binary digit", after the binary code (1 or 0) computers use to store and process data
Byte (B)	8 bits	Enough information to create an English letter or number in computer code. It is the basic unit of computing
Kilobyte (KB)	1,000, or 2^{10}, bytes	From "thousand" in Greek. One page of typed text is 2KB
Megabyte (MB)	1,000KB; 2^{20} bytes	From "large" in Greek. The complete works of Shakespeare total 5MB. A typical pop song is about 4MB
Gigabyte (GB)	1,000MB; 2^{30} bytes	From "giant" in Greek. A two-hour film can be compressed into 1-2GB
Terabyte (TB)	1,000GB; 2^{40} bytes	From "monster" in Greek. All the catalogued books in America's Library of Congress total 15TB
Petabyte (PB)	1,000TB; 2^{50} bytes	All letters delivered by America's postal service this year will amount to around 5PB. Google processes around 1PB every hour
Exabyte (EB)	1,000PB; 2^{60} bytes	Equivalent to 10 billion copies of *The Economist*
Zettabyte (ZB)	1,000EB; 2^{70} bytes	The total amount of information in existence this year is forecast to be around 1.2ZB
Yottabyte (YB)	1,000ZB; 2^{80} bytes	Currently too big to imagine

The prefixes are set by an intergovernmental group, the International Bureau of Weights and Measures. Yotta and Zetta were added in 1991; terms for larger amounts have yet to be established.

Source: *The Economist*

Big Data—Culturomics:

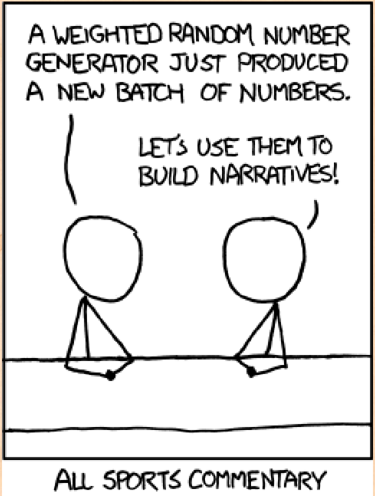
“Quantitative analysis of culture using millions of digitized books” by Michel et al., Science, 2011 [11]



http://www.culturomics.org/ (⊞)
Google Books ngram viewer (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Homo narrativus:



http://xkcd.com/904/ (⊞)

- Mechanisms = Evolution equations, algorithms, stories, ...
- Rollover zing: “Also, all financial analysis. And, more directly, D&D.”

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Basic Science $\simeq$ Describe + Explain:

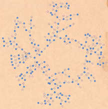
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Lord Kelvin (possibly):

- "To measure is to know."
- "If you cannot measure it, you cannot improve it."

Bonus:

- "X-rays will prove to be a hoax."
- "There is nothing new to be discovered in physics now, All that remains is more and more precise measurement."



Emergence:

Tornadoes, financial collapses, human emotion aren't found in water molecules, dollar bills, or carbon atoms.

Examples:

- Fundamental particles $\rightarrow$ Life, the Universe, and Everything
- Genes $\rightarrow$ Organisms
- Brains $\rightarrow$ Thoughts
- People $\rightarrow$ The Web
- People $\rightarrow$ Religion
- People $\rightarrow$ Language, and rules in language (e.g., -ed, -s).
- ? $\rightarrow$ time; ? $\rightarrow$ gravity; ? $\rightarrow$ reality.

"The whole is more than the sum of its parts" –Aristotle

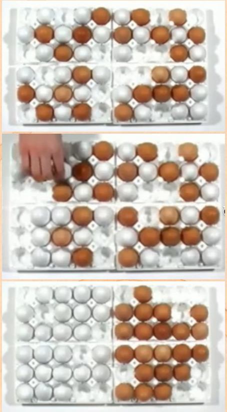
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Toast + Capers + Almonds = Something Different:

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Emergence—Mechanism

Thomas Schelling (⊞) (Economist/Nobelist):



[youtube] (⊞)

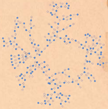
- "Micromotives and Macrobehavior" [14]
  - Segregation
  - Wearing hockey helmets
  - Seating choices

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Reductionism

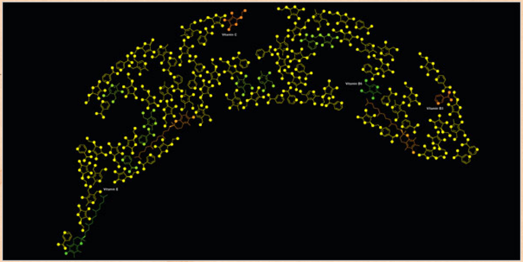
- Complex Systems enthusiasts often decry reductionist approaches . . .
- But reductionism seems to be misunderstood.
- Reductionist techniques can explain weak emergence (e.g., phase transitions).
- 'A Miracle Occurs' explains strong emergence.
- But: maybe miracle should be interpreted as an inscrutable yet real mechanism that cannot be simply described. Gulp.
- Listen to Steve Strogatz and Hod Lipson (Cornell) in the last piece on Radiolab's show 'Limits' (51:40): `http://blogs.wnyc.org/radiolab/2010/04/05/limits/`

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



The emergence of taste:

- Molecules → Ingredients → Taste/Nutrition/Health
- See Michael Pollan's article on nutritionism (⊞) in the New York Times, January 28, 2007.



nytimes.com

- See also: bumblebees.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Limits to what is possible:

Universality (⊞):

- The property that the macroscopic aspects of a system do not depend sensitively on the system's details.
- Key figure: Leo Kadanoff (⊞).

Examples:

- The Central Limit Theorem:

$$P(x;\mu,\sigma)\mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}\mathrm{d}x\,.$$

- Nature of phase transitions in statistical mechanics.
- Navier Stokes equation for fluids.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

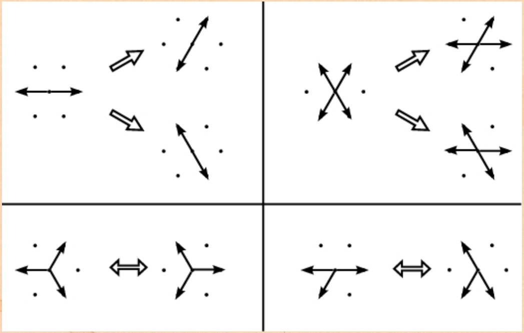
Fluids mechanics

- Fluid mechanics = One of the great successes of understanding complex systems.
- Navier-Stokes equations: micro-macro system evolution.
- The big three: Experiment + Theory + Simulations.
- Works for many very different ‘fluids’:
  - the atmosphere,
  - oceans,
  - blood,
  - galaxies,
  - the earth’s mantle...
  - and ball bearings on lattices...?

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Lattice gas models

Collision rules in 2-d on a hexagonal lattice:

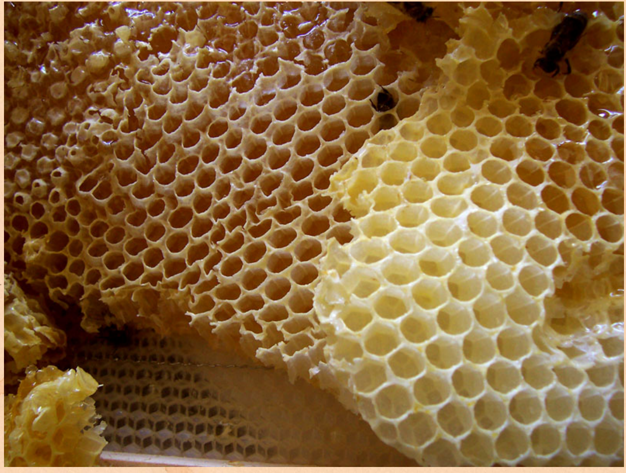


- Lattice matters... Only hexagonal lattice works in 2-d.
- No ‘good’ lattice in 3-d.
- Upshot: play with ‘particles’ of a system to obtain new or specific macro behaviours.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Hexagons—Honeycomb: (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



- Orchestrated? Or an accident of bees working hard?
- See "On Growth and Form" by D'Arcy Wentworth Thompson (⊞).[16, 17]

Hexagons—Giant's Causeway: (⊞)

http://newdesktopwallpapers.info

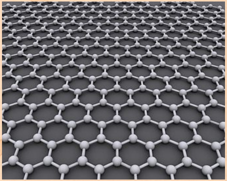
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Hexagons—Giant's Causeway: (⊞)

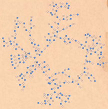
http://www.physics.utoronto.ca/

Hexagons run amok:

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

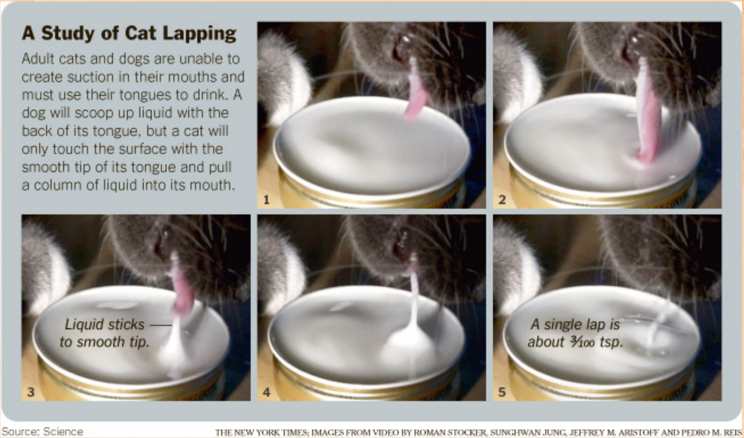


- Graphene (⊞): single layer of carbon molecules in a perfect hexagonal lattice (super strong).
- Chicken wire (⊞) ...



Whimsical but great example of real science:

"How Cats Lap: Water Uptake by *Felis catus*" (⊞)
Reis et al., *Science*, 2010.

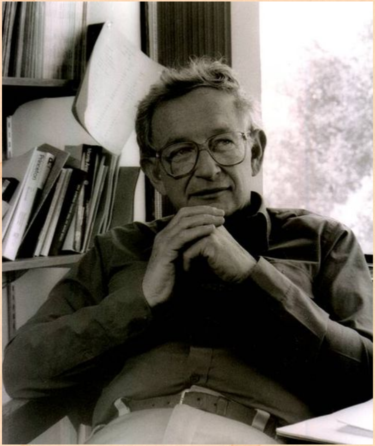


Amusing interview here (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Symmetry Breaking

Philip Anderson (⊞)—"More is Different," Science, 1972 [2]



- Argues against idea that the only real scientists are those working on the fundamental laws.
- Symmetry breaking → different laws/rules at different scales...

(2006 study → "most creative physicist in the world" (⊞))

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Symmetry Breaking

“Elementary entities of science X obey the laws of science Y”

- X
- solid state or many-body physics
- chemistry
- molecular biology
- cell biology

vdots

- psychology
- social sciences

- Y
- elementary particle physics
- solid state many-body physics
- chemistry
- molecular biology

⋮

- physiology
- psychology

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Symmetry Breaking

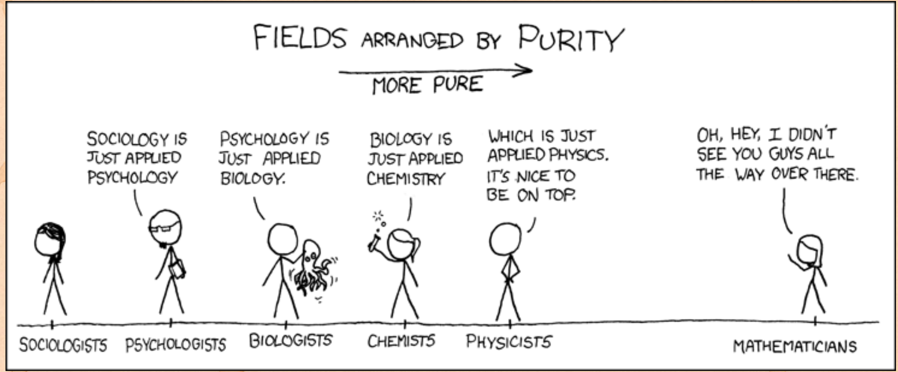
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Anderson:
[the more we know about] “fundamental laws, the less relevance they seem to have to the very real problems of the rest of science.”

Scale and complexity thwart the constructionist hypothesis.

Accidents of history and path dependence (⊞) matter.

More is different:



http://xkcd.com/435/ (⊞)

A real science of complexity:

A real theory of ~~everything~~ anything:

1. Is not just about the ridiculously small stuff...
2. It's about the increase of complexity

Symmetry breaking/ Accidents of history	vs.	Universality

- Second law of thermodynamics: we're toast in the long run.
- So how likely is the local complexification of structure we enjoy?
- How likely are the Big Transitions?

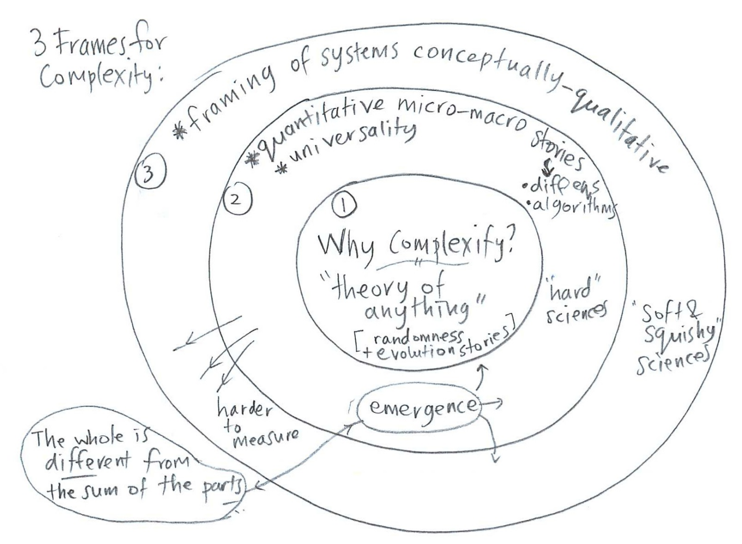
Complexification—the Big Transitions:

- Big Bang.
- Big Randomness.
- Big Replicate.
- Big Life.
- Big Evolve.

- Big Word.
- Big Story.
- Big Number.
- Big God.
- Big Make.

- Big Science.
- Big Data.
- Big Information.
- Big Algorithm.
- Big Connection.
- Big Social.
- Big Awareness.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

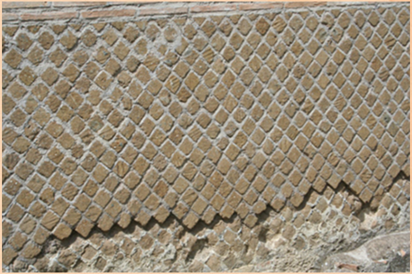


Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Ancestry:

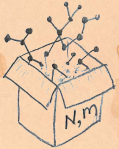
From Keith Briggs's excellent etymological investigation: (⊞)

- Opus reticulatum:
- A Latin origin?



[http://serialconsign.com/2007/11/we-put-net-network]

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

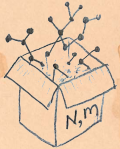


Key Observation:

- Many complex systems can be viewed as complex networks of physical or abstract interactions.
- Opens door to mathematical and numerical analysis.
- Dominant approach of last decade of a theoretical-physics/stat-mechish/combinatorics flavor.
- Mindboggling amount of work published on complex networks since 1998...
- ... largely due to your typical theoretical physicist:

- *Piranha physicus*
- Hunt in packs.
- Feast on new and interesting ideas (see chaos, cellular automata, ...)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

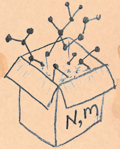


More observations

- But surely networks aren't new...
- Graph theory is well established...
- Study of social networks started in the 1930's...
- So why all this 'new' research on networks?
- Answer (to repeat): Oodles of Easily Accessible Data.
- We can now inform (alas) our theories with a much more measurable reality.*
- Crucial observation: Real networks occupy a tiny, low entropy part of all network space and require specific attention.
- A central goal: establish mechanistic explanations.
- What kinds of dynamics lead to these real networks?

** If this is upsetting, maybe string theory is for you...*

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Popularity (according to ISI)

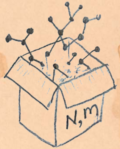
“Collective dynamics of ‘small-world’ networks” [20]

- Watts and Strogatz
  Nature, 1998
- ≈ 4677 citations (as of January 18, 2011)
- Over 1100 citations in 2008 alone.

“Emergence of scaling in random networks” [3]

- Barabási and Albert
  Science, 1999
- ≈ 5270 citations (as of January 18, 2011)
- Over 1100 citations in 2008 alone.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

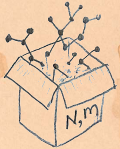


Models

1. generalized random networks:

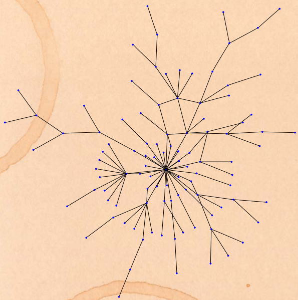
- Arbitrary degree distribution P_k.
- Wire nodes together randomly.
- Create ensemble to test deviations from randomness.
- Interesting, applicable, rich mathematically, very important.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Models

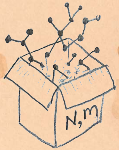
2. 'scale-free networks':



γ = 2.5
$\langle k \rangle$ = 1.8
$N = 150$

- Introduced by Barabasi and Albert [3]
- Generative, mechanistic model
- Ancestory: Herbert Simon's model for Zipf's law [15]
- Preferential attachment model with growth:
- P[attachment to node i] $\propto \mathrm{k}_\mathrm{i}^{\alpha}$.
- Produces $P_k \sim k^{-\gamma}$ when $\alpha = 1$.
- Trickiness: other models generate skewed degree distributions.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Models

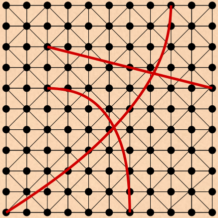
3. small-world networks

- Introduced by Watts and Strogatz [20]

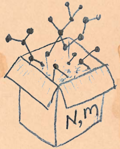
Two scales:

- local regularity (an individual's friends know each other)
- global randomness (shortcuts).

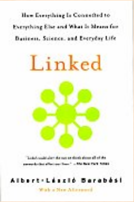
- Shortcuts allow disease to jump
- Number of infectives increases exponentially in time
- Facilitates synchronization



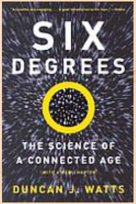
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Popularity according to books:

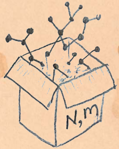


Linked: How Everything Is Connected to Everything Else and What It Means—Albert-Laszlo Barabási



Six Degrees: The Science of a Connected Age—Duncan Watts [19]

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



More observations

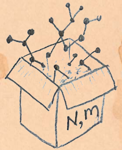
- Web-scale data sets can be overly exciting.

Witness:

- The End of Theory: The Data Deluge Makes the Scientific Theory Obsolete (Anderson, Wired) (⊞)
- "The Unreasonable Effectiveness of Data," Halevy et al. [9]
- c.f. Wigner's "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" [21]

But:

- For scientists, description is only part of the battle.
- We still need to understand.

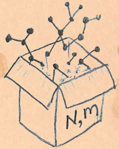


Examples

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

What passes for a complex network?

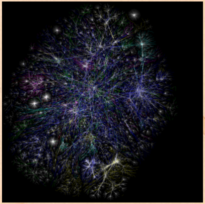
- Complex networks are large (in node number)
- Complex networks are sparse (low edge to node ratio)
- Complex networks are usually dynamic and evolving
- Complex networks can be social, economic, natural, informational, abstract, ...

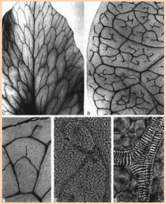


Examples

Physical networks

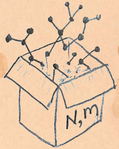
- River networks
- Neural networks
- Trees and leaves
- Blood networks
- The Internet
- Road networks
- Power grids





- Distribution (branching) versus redistribution (cyclical)

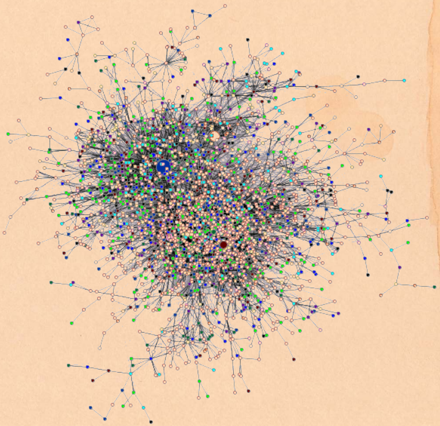
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Examples

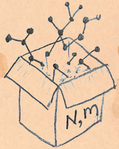
Interaction networks

- The Blogosphere
- Biochemical networks
- Gene-protein networks
- Food webs: who eats whom
- The World Wide Web (?)
- Airline networks
- Call networks (AT&T)
- The Media



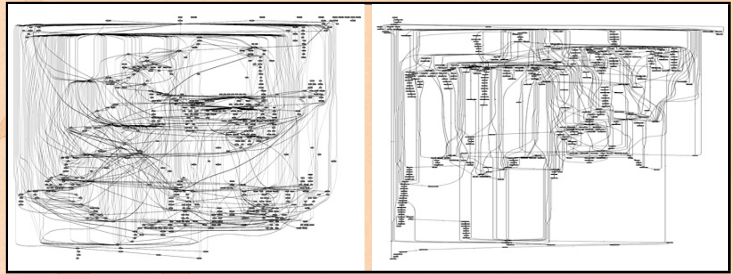
datamining.typepad.com (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Dynamic networks: Server security

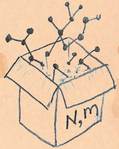
Serving one html page with an image:



- Map of system calls made by a Linux server running Apache and Windows server running IIS. Which is which?

Taken from http://www.visualcomplexity.com (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

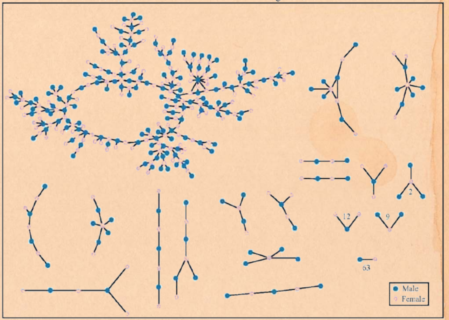


Examples

Interaction networks: social networks

- Snogging
- Friendships
- Acquaintances
- Boards and directors
- Organizations
- twitter.com (⊞) facebook.com (⊞),

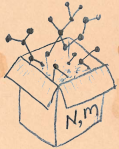
The Structure of Romantic and Sexual Relations at "Jefferson High School"



Each circle represents a student and lines connecting students represent romantic relations occuring within the 6 months preceding the interview. Numbers under the figure count the number of times that pattern was observed (i.e. we found 63 pairs unconnected to anyone else)

(Bearman *et al.*, 2004)

- 'Remotely sensed' by: tweets (open), instant messaging, Facebook posts, emails, phone logs (*cough*).



Examples

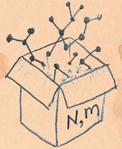
Relational networks

- Consumer purchases (⊞)
  (Wal-Mart: $\approx$ 2.5 petabyte $= 2.5 \times 10^{15}$ bytes)
- Thesauri: Networks of words generated by meanings
- Knowledge/Databases/Ideas
- Metadata—Tagging: delicious (⊞), flickr (⊞)

common tags cloud | list

community daily dictionary education **encyclopedia**
english free imported info information internet knowledge
learning news **reference** research resource
resources search tools useful web web2.0 **wiki**
wikipedia

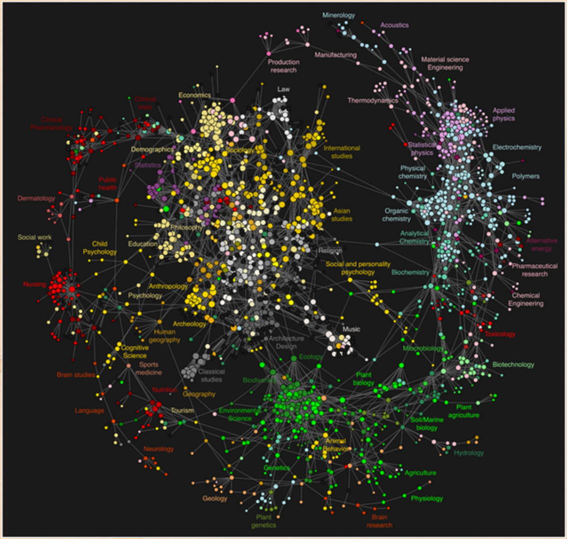
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



The UNIVERSITY of VERMONT

Clickworthy Science:

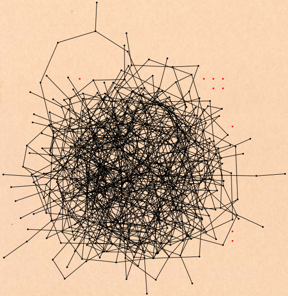
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Bollen et al. [5]; a higher resolution figure is here (⊞)

A notable feature of large-scale networks:

- Graphical renderings are often just a big mess.

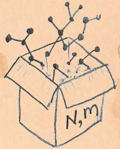


⇐ Typical hairball

- number of nodes $N = 500$
- number of edges $m = 1000$
- average degree $\langle k \rangle = 4$

- And even when renderings somehow look good:
  "That is a very graphic analogy which aids understanding wonderfully while being, strictly speaking, wrong in every possible way"
  said Ponder [Stibbons] —*Making Money*, T. Pratchett.
- We need to extract digestible, meaningful aspects.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Properties

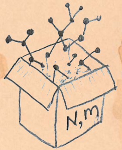
Some key aspects of real complex networks:

- degree distribution P_k*
- assortativity
- homophily
- clustering
- motifs
- modularity
- concurrency
- hierarchical scaling
- network distances
- centrality
- efficiency
- robustness

- Plus coevolution of network structure and processes on networks.

* Degree distribution is the elephant in the room that we are now all very aware of...

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

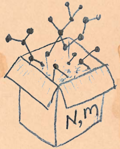


Nutshell:

Overview Key Points:

- The field of complex networks came into existence in the late 1990s.
- Explosion of papers and interest since 1998/99.
- Hardened up much thinking about complex systems.
- Specific focus on networks that are large-scale, sparse, natural or man-made, evolving and dynamic, and (crucially) measurable.
- Three main (blurred) categories:
  1. Physical (e.g., river networks),
  2. Interactional (e.g., social networks),
  3. Abstract (e.g., thesauri).

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

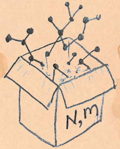


Nutshell:

Overview Key Points (cont.):

- Obvious connections with the vast extant field of graph theory.
- But focus on dynamics is more of a physics/stat-mech/comp-sci flavor.
- Two main areas of focus:
  1. Description: Characterizing very large networks
  2. Explanation: Micro story $\rightarrow$ Macro features
- Some essential structural aspects are understood: degree distribution, clustering, assortativity, group structure, overall structure,...
- Still much work to be done, especially with respect to dynamics... exciting!

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



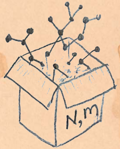
Bonus materials:

Graduate Course Websites:

- Principles of Complex Systems (⊞), University of Vermont
- Complex Networks (⊞), University of Vermont

Textbooks:

- David Easley and Jon Kleinberg (Economics and Computer Science, Cornell)
  "Networks, Crowds, and Markets: Reasoning About a Highly Connected World" (⊞)
- Mark Newman (Physics, Michigan)
  "Networks: An Introduction" (⊞)

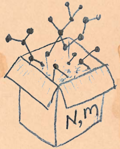


Bonus materials:

Review articles:

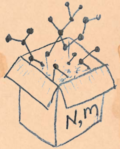
- S. Boccaletti et al.
  "Complex networks: structure and dynamics" [4]
  Times cited: 1,028 (as of June 7, 2010)
- M. Newman
  "The structure and function of complex networks" [12]
  Times cited: 2,559 (as of June 7, 2010)
- R. Albert and A.-L. Barabási
  "Statistical mechanics of complex networks" [1]
  Times cited: 3,995 (as of June 7, 2010)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References



Bonus materials:

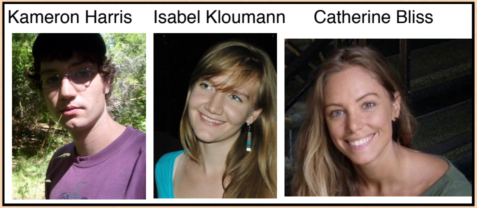
- Complex Social Networks—F. Vega-Redondo [18]
- Fractal River Basins: Chance and Self-Organization—I. Rodríguez-Iturbe and A. Rinaldo [13]
- Random Graph Dynamics—R. Durette
- Scale-Free Networks—Guido Caldarelli
- Evolution and Structure of the Internet: A Statistical Physics Approach—Romu Pastor-Satorras and Alessandro Vespignani
- Complex Graphs and Networks—Fan Chung
- Social Network Analysis—Stanley Wasserman and Kathleen Faust
- Handbook of Graphs and Networks—Eds: Stefan Bornholdt and H. G. Schuster [6]
- Evolution of Networks—S. N. Dorogovtsev and J. F. F. Mendes [8]



The Team:

1. People:

Chris Danforth



2. Machines:

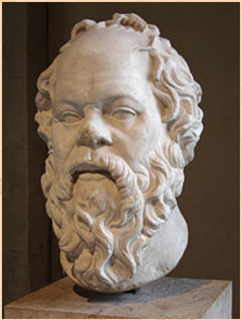
- 3000 processors + storage at the Vermont Advanced Computing Center
- 40 TB of storage in Danforth's office.

3. Support:

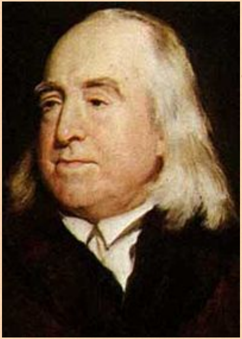
NSF and NASA.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

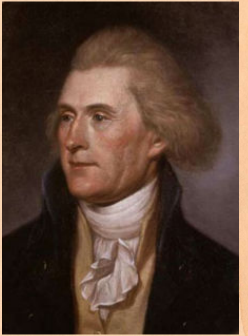
Happiness:



Socrates et al.: eudaimonia [10]



Bentham: hedonistic calculus



Jefferson: . . . the pursuit of happiness

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

Early drafts:

that among these are:

Life, ✓

Liberty, ✓ and ?? ~~Money~~?

~~Libations~~ ... ~~Alcohol~~

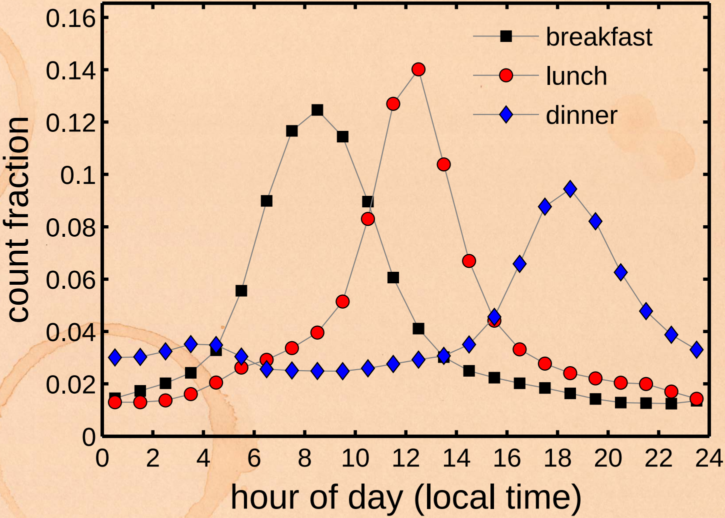
~~Property~~ ~~Foot-the-ball~~

~~Beer~~

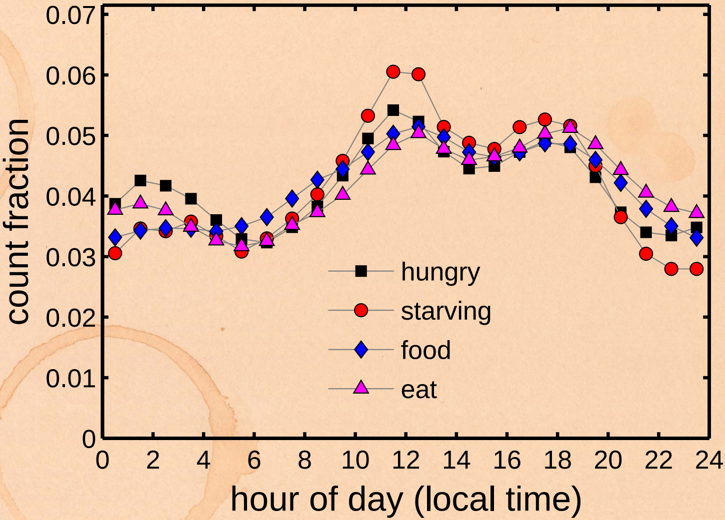
Happiness ✓✓✓

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

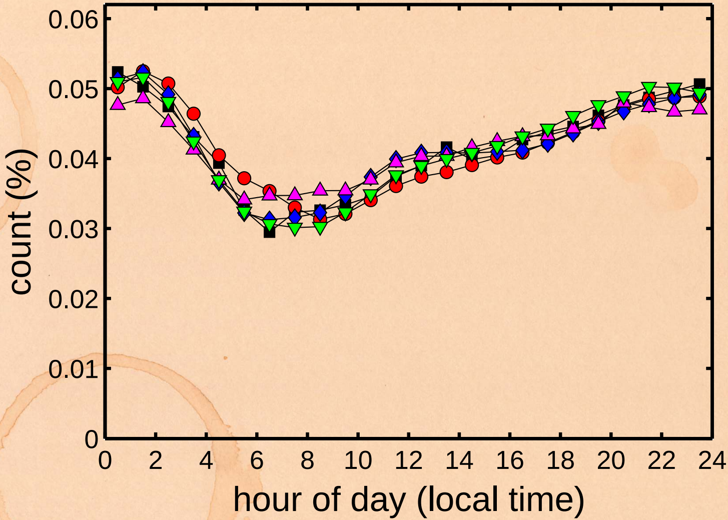
Twitter—living in the now:



Twitter—living in the now:



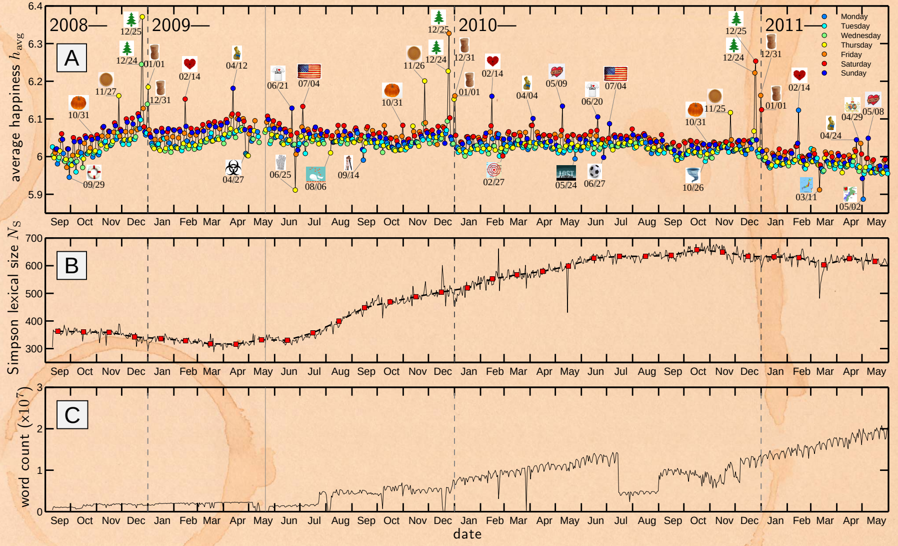
Twitter—living in the now:

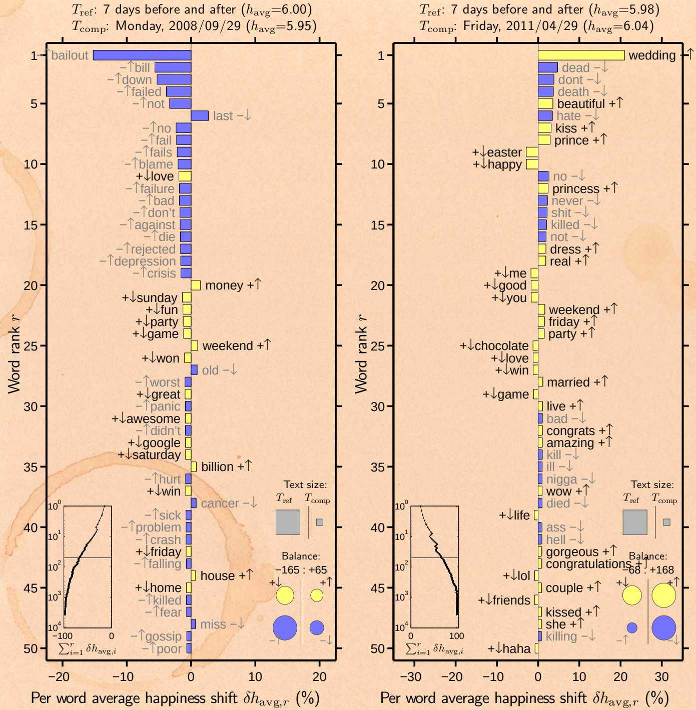


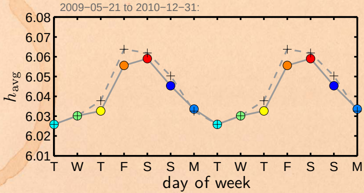
A few words you can't say on television.

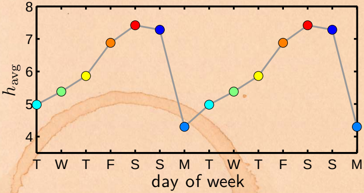
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

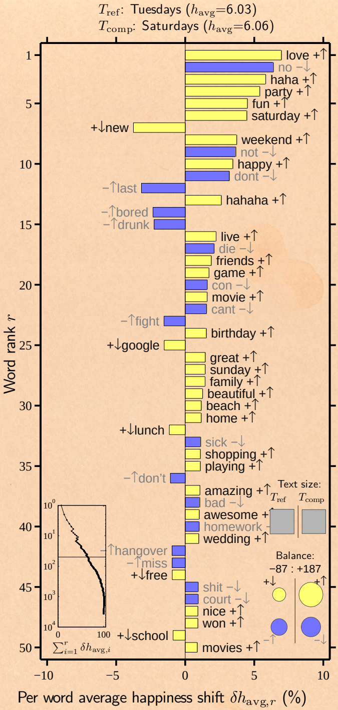
Twitter—overall time series:




$T_{\rm ref}$: 7 days before and after ($h_{\rm avg}$=6.00)
$T_{\rm comp}$: Monday, 2008/09/29 ($h_{\rm avg}$=5.95)
bailout, bill, down, failed, not, last, no, fail, fails, blame, love, failure, bad, don't, against, die, rejected, depression, crisis, money, sunday, fun, party, game, weekend, won, old, worst, great, panic, awesome, didn't, google, saturday, billion, hurt, win, cancer, sick, problem, crash, friday, falling, house, home, killed, fear, miss, gossip, poor
Word rank r
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)
Text size:
Balance:
−165 : +65
$\sum_{i=1}^{r} \delta h_{{\rm avg},i}$
$T_{\rm ref}$: 7 days before and after ($h_{\rm avg}$=5.98)
$T_{\rm comp}$: Friday, 2011/04/29 ($h_{\rm avg}$=6.04)
wedding, dead, dont, death, beautiful, hate, kiss, prince, easter, happy, no, princess, never, shit, killed, not, dress, real, me, good, you, weekend, friday, party, chocolate, love, win, married, game, live, bad, congrats, amazing, kill, ill, nigga, wow, died, life, ass, hell, gorgeous, congratulations, lol, couple, friends, kissed, she, killing, haha
Word rank r
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)
Text size:
Balance:
−68 : +168
$\sum_{i=1}^{r} \delta h_{{\rm avg},i}$







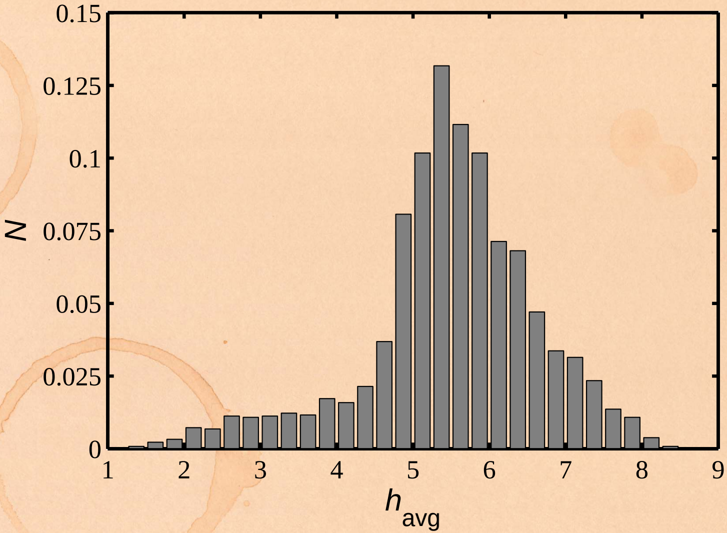
Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

valence rank	word	valence	std dev	twitter rank	g-books rank	nyt rank	lyrics rank
1	laughter	8.50	0.93	3600	–	–	1728
2	happiness	8.44	0.97	1853	2458	–	1230
3	love	8.42	1.11	25	317	328	23
4	happy	8.30	0.99	65	1372	1313	375
5	laughed	8.26	1.16	3334	3542	–	2332
6	laugh	8.22	1.37	1002	3998	4488	647
7	laughing	8.20	1.11	1579	–	–	1122
8	excellent	8.18	1.10	1496	1756	3155	–
9	laughs	8.18	1.16	3554	–	–	2856
10	joy	8.16	1.06	988	2336	2723	809
11	successful	8.16	1.08	2176	1198	1565	–
12	win	8.12	1.08	154	3031	776	694
13	rainbow	8.10	0.99	2726	–	–	1723
14	smile	8.10	1.02	925	2666	2898	349
15	won	8.10	1.22	810	1167	439	1493
16	pleasure	8.08	0.97	1497	1526	4253	1398
17	smiled	8.08	1.07	–	3537	–	2248
18	rainbows	8.06	1.36	–	–	–	4216
19	winning	8.04	1.05	1876	–	1426	3646
20	celebration	8.02	1.53	3306	–	2762	4070
21	enjoyed	8.02	1.53	1530	2908	3502	–
22	healthy	8.02	1.06	1393	3200	3292	4619
23	music	8.02	1.12	132	875	167	374
24	celebrating	8.00	1.14	2550	–	–	–
25	congratulations	8.00	1.63	2246	–	–	–
26	weekend	8.00	1.29	317	–	833	2256
27	celebrate	7.98	1.15	1606	–	3574	2108
28	comedy	7.98	1.15	1444	–	2566	–
29	jokes	7.98	0.98	2812	–	–	3808
30	rich	7.98	1.32	1625	1221	1469	890
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

valence rank	word	valence	std dev	twitter rank	g-books rank	nyt rank	lyrics rank
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10193	violence	1.86	1.05	4299	1724	1238	2016
10194	cruel	1.84	1.15	2963	–	–	1447
10195	cry	1.84	1.28	1028	3075	–	226
10196	failed	1.84	1.00	2645	1618	1276	2920
10197	sickness	1.84	1.18	4735	–	–	3782
10198	abused	1.83	1.31	–	–	–	4589
10199	tortured	1.82	1.42	–	–	–	4693
10200	fatal	1.80	1.53	–	4089	–	3724
10201	killings	1.80	1.54	–	–	4914	–
10202	murdered	1.80	1.63	–	–	–	4796
10203	war	1.80	1.41	468	175	291	462
10204	kills	1.78	1.23	2459	–	–	2857
10205	jail	1.76	1.02	1642	–	2573	1619
10206	terror	1.76	1.00	4625	4117	4048	2370
10207	die	1.74	1.19	418	730	2605	143
10208	killing	1.70	1.36	1507	4428	1672	998
10209	arrested	1.64	1.01	2435	4474	1435	–
10210	deaths	1.64	1.14	–	–	2974	–
10211	raped	1.64	1.43	–	–	–	4528
10212	torture	1.58	1.05	3175	–	–	3126
10213	died	1.56	1.20	1223	866	208	826
10214	kill	1.56	1.05	798	2727	2572	430
10215	killed	1.56	1.23	1137	1603	814	1273
10216	cancer	1.54	1.07	946	1884	796	3802
10217	death	1.54	1.28	509	307	373	433
10218	murder	1.48	1.01	2762	3110	1541	1059
10219	terrorism	1.48	0.91	–	–	3192	–
10220	rape	1.44	0.79	3133	–	4115	2977
10221	suicide	1.30	0.84	2124	4707	3319	2107
10222	terrorist	1.30	0.91	3576	–	3026	–

std dev rank	word	valence	std dev	twitter rank	g-books rank	nyt rank	lyrics rank
1	f♯@king	4.64	2.93	448	–	–	620
2	f☆☆kin	3.86	2.74	1077	–	–	688
3	f☆☆ked	3.56	2.71	1840	–	–	904
4	pussy	4.80	2.66	2019	–	–	949
5	whiskey	5.72	2.64	–	–	–	2208
6	slut	3.57	2.63	–	–	–	4071
7	cigarettes	3.31	2.60	–	–	–	3279
8	f☆☆k	4.14	2.58	322	–	–	185
9	mortality	4.38	2.55	–	3960	–	–
10	cigarette	3.09	2.52	–	–	–	2678
11	motherf☆☆kers	2.51	2.47	–	–	–	1466
12	churches	5.70	2.46	–	2281	–	–
13	motherf☆☆king	2.64	2.46	–	–	–	2910
14	capitalism	5.16	2.45	–	4648	–	–
15	porn	4.18	2.43	1801	–	–	–
16	summer	6.40	2.39	896	1226	721	590
17	beer	5.92	2.39	839	4924	3960	1413
18	execution	3.10	2.39	–	2975	–	–
19	wines	6.28	2.37	–	–	3316	–
20	zombies	4.00	2.37	4708	–	–	–
21	aids	4.28	2.35	2983	3996	1197	–
22	capitalist	4.84	2.34	–	4694	–	–
23	revenge	3.71	2.34	–	–	–	2766
24	mcdonalds	5.98	2.33	3831	–	–	–
25	beatles	6.44	2.33	3797	–	–	–
26	islam	4.68	2.33	–	4514	–	–
27	pay	5.30	2.32	627	769	460	499
28	alcohol	5.20	2.32	2787	2617	3752	3600
29	muthaf☆☆kin	3.00	2.31	–	–	–	4107
30	christ	6.16	2.31	2509	909	4238	1526
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Positive bias in the English language:



Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

For more...

- PSD, KDH, IMK, CAB, and CMD
  "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter."
  http://arxiv.org/abs/1101.5120 (⊞)
- P. S. Dodds and C. M. Danforth
  "Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents." [7]
  Journal of Happiness Studies, 2009.
- http://www.uvm.edu/~pdodds/research/ (⊞)
- http://www.onehappybird.com (⊞)

- "Does a Nation's Mood Lurk in Its Songs and Blogs?" by Benedict Carey
  New York Times, August 2009. (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

References I

[1] R. Albert and A.-L. Barabási.
Statistical mechanics of complex networks.
Rev. Mod. Phys., 74:47–97, 2002. pdf (⊞)

[2] P. W. Anderson.
More is different.
Science, 177(4047):393–396, 1972. pdf (⊞)

[3] A.-L. Barabási and R. Albert.
Emergence of scaling in random networks.
Science, 286:509–511, 1999. pdf (⊞)

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang.
Complex networks: Structure and dynamics.
Physics Reports, 424:175–308, 2006. pdf (⊞)

References II

[5] J. Bollen, H. Van de Sompel, A. Hagberg, L. Bettencourt, R. Chute, M. A. Rodriguez, and B. Lyudmila.
Clickstream data yields high-resolution maps of science.
PLoS ONE, 4:e4803, 2009. pdf (⊞)

[6] S. Bornholdt and H. G. Schuster, editors.
Handbook of Graphs and Networks.
Wiley-VCH, Berlin, 2003.

[7] P. S. Dodds and C. M. Danforth.
Measuring the happiness of large-scale written expression: Songs, blogs, and presidents.
Journal of Happiness Studies, 2009. doi:10.1007/s10902-009-9150-9. pdf (⊞)

References III

[8] S. N. Dorogovtsev and J. F. F. Mendes.
Evolution of Networks.
Oxford University Press, Oxford, UK, 2003.

[9] A. Halevy, P. Norvig, and F. Pereira.
The unreasonable effectiveness of data.
IEEE Intelligent Systems, 24:8–12, 2009. pdf (⊞)

[10] W. T. Jones.
The Classical Mind.
Harcourt, Brace, Jovanovich, New York, 1970.

[11] J.-B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, The Google Books 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.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

References IV

Science Magazine, 331:176–182, 2011. pdf (⊞)

[12] M. E. J. Newman.
The structure and function of complex networks.
SIAM Review, 45(2):167–256, 2003. pdf (⊞)

[13] I. Rodríguez-Iturbe and A. Rinaldo.
Fractal River Basins: Chance and Self-Organization.
Cambridge University Press, Cambrigde, UK, 1997.

[14] T. C. Schelling.
Micromotives and Macrobehavior.
Norton, New York, 1978.

[15] H. A. Simon.
On a class of skew distribution functions.
Biometrika, 42:425–440, 1955. pdf (⊞)

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

References V

[16] D. W. Thompson.
On Growth and From.
Cambridge University Pres, Great Britain, 2nd edition, 1952.

[17] D. W. Thompson.
On Growth and Form — Abridged Edition.
Cambridge University Press, Great Britain, 1961.

[18] F. Vega-Redondo.
Complex Social Networks.
Cambridge University Press, 2007.

[19] D. J. Watts.
Six Degrees.
Norton, New York, 2003.

Complexity
Introduction
Emergence
Universality
Symmetry Breaking
The Big Theory
Revolution: Big Data & Complex Networks
Nutshell
Measuring Happiness
Tweetage
Mechanical Turk
References

References VI

[20] D. J. Watts and S. J. Strogatz.
Collective dynamics of ‘small-world’ networks.
Nature, 393:440–442, 1998. pdf (⊞)

[21] E. Wigner.
The unreasonable effectivenss of mathematics in the natural sciences.
Communications on Pure and Applied Mathematics, 13:1–14, 1960. pdf (⊞)