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
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
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 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.
Big Data Science:

- 2013: year traffic on Internet estimate to reach 2/3 Zettabytes
  \((1\text{ZB} = 10^3\text{EB} = 10^6\text{PB} = 10^9\text{TB})\)
- 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

Exponential growth:
\(\sim 60\%\) per year.
No really, that’s a lot of data

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

Source: *The Economist*  
Yotta and Zetta were added in 1991; terms for larger amounts have yet to be established.
Big Data—Culturomics:


http://www.culturomics.org/ (田)
Google Books ngram viewer (田)
Homo narrativus:

A weighted random number generator just produced a new batch of numbers.

Let's use them to build narratives!

All sports commentary

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

http://xkcd.com/904/ (❖)
Basic Science $\simeq$ Describe + Explain:

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 → Life, the Universe, and Everything
- Genes → Organisms
- Brains → Thoughts
- People → The Web
- People → Religion
- People → Language, and rules in language (e.g., -ed, -s).

“The whole is more than the sum of its parts” – Aristotle
Complexity, Big Data Science, and Happiness

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:
Emergence—Mechanism

Thomas Schelling ( Economist/Nobelist ):

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

[youtube] ( )
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/
The emergence of taste:

- Molecules → Ingredients → Taste/Nutrition/Health

See also: bumblebees.
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:
  \[ P(x; \mu, \sigma) dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx. \]
- Nature of phase transitions in statistical mechanics.
- Navier Stokes equation for fluids.
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...?
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.
Hexagons—Honeycomb: (_hexagons)

- Orchestrated? Or an accident of bees working hard?
Complexity, Big Data Science, and Happiness

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

Measuring Happiness
Tweetage
Mechanical Turk

References

Hexagons—Giant’s Causeway: (田)

http://newdesktopwallpapers.info
Complexity, Big Data Science, and Happiness

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

Measuring Happiness
Tweetage
Mechanical Turk

References

http://www.physics.utoronto.ca/
Hexagons run amok:

- **Graphene (_hexagon):** single layer of carbon molecules in a perfect hexagonal lattice (super strong).
- **Chicken wire (hexagon) ...**
Whimsical but great example of real science:

“How Cats Lap: Water Uptake by *Felis catus*” ( refunded )

A Study of Cat Lapping
Adult cats and dogs are unable to create suction in their mouths and must use their tongues to drink. A dog will scoop up liquid with the back of its tongue, but a cat will only touch the surface with the smooth tip of its tongue and pull a column of liquid into its mouth.

Liquid sticks —— to smooth tip.

A single lap is about 3/100 tsp.

Source: *Science*

THE NEW YORK TIMES; IMAGES FROM VIDEO BY ROMAN STOCKER, SUNGHWAN JUNG, JEFFREY M. ARISTOFF AND PEDRO M. REIS

Amusing interview here ( refunded )
Symmetry Breaking

Philip Anderson — “More is Different,” Science, 1972

- 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”)
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

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

$\ldots$
Symmetry Breaking

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 ( Plainsight ) matter.
More is different:

FIELDS ARRANGED BY PURITY

MORE PURE

SOCIOLGY IS JUST APPLIED PSYCHOLOGY
PSYCHOLOGY IS JUST APPLIED BIOLOGY
BIOLOGY IS JUST APPLIED CHEMISTRY
WHICH IS JUST APPLIED PHYSICS.
IT'S NICE TO BE ON TOP.

OH, HEY, I DIDN'T SEE YOU GUYS ALL THE WAY OVER THERE.

SOCIOLISTS  PSYCHOLOGISTS  BIOLOGISTS  CHEMISTS  PHYSICISTS  MATHEMATICIANS

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, Big Data Science, and Happiness

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?

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, ...).
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...*
“Collective dynamics of ‘small-world’ networks” [20]

- Watts and Strogatz
- \approx 4677 \text{ citations} \ (\text{as of January 18, 2011})
- Over 1100 citations in 2008 alone.

“Emergence of scaling in random networks” [3]

- Barabási and Albert
  Science, 1999
- \approx 5270 \text{ citations} \ (\text{as of January 18, 2011})
- Over 1100 citations in 2008 alone.
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.
2. ‘scale-free networks’:

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

$\gamma = 2.5$
$\langle k \rangle = 1.8$
$N = 150$
3. small-world networks

- Introduced by Watts and Strogatz\footnote{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, Big Data Science, and Happiness

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]
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) \[\text{(Allen)}\]
- “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

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)
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](http://datamining.typepad.com) (田)
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](http://www.visualcomplexity.com)
Examples

Interaction networks: social networks

- Snogging
- Friendships
- Acquaintances
- Boards and directors
- Organizations

- twitter.com
- facebook.com

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

(Bearman et al., 2004)
Examples

Relational networks

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

**common tags**
- cloud | list
- encyclopedia
- english free imported info information internet knowledge learning news reference research resource resources search tools useful web web2.0 wiki wikipedia
Clickworthy Science:

Bollen et al. [5]; a higher resolution figure is here (払い)
A notable feature of large-scale networks:

- Graphical renderings are often just a big mess.

\[ N = 500 \]
\[ m = 1000 \]
\[ \langle k \rangle = 4 \]

- Typical hairball

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**.
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...
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).
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 → 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, Big Data Science, and Happiness

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 ( لما ), University of Vermont
- Complex Networks ( لما ), University of Vermont

Textbooks:

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

References

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

Measuring Happiness
- Tweetage
- Mechanical Turk

The Team:

1. People:

   Chris Danforth

   Kameron Harris  Isabel Kloumann  Catherine Bliss

The Team:

2. Machines:

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

The Team:

3. Support:

   NSF and NASA.
Happiness:

Socrates et al.: eudaimonia\textsuperscript{[10]}

Bentham: hedonistic calculus

Jefferson: \ldots the pursuit of happiness
that among these are:

Life, Liberty, and ?? Money?

Property

Beer

Happiness
Twitter—living in the now:

- breakfast
- lunch
- dinner

count fraction

hour of day (local time)
Twitter—living in the now:

![Graph showing count fraction by hour of day for different keywords like hungry, starving, food, and eat.](image)

- **Complexity, Big Data Science, and 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:

A few words you can’t say on television.
Twitter—overall time series:

A

average happiness $h_{avg}$


Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday

B

Simpson lexical size $N_S$

C

word count ($x 10^7$)

date

C

word count ($x 10^7$)
Complexity, Big Data Science, and Happiness

Introduction

Emergence

Universality

Symmetry Breaking

The Big Theory Revolution: Big Data & Complex Networks

Nutshell

Measuring Happiness

Tweetage

Mechanical Turk

References
<table>
<thead>
<tr>
<th>rank</th>
<th>word</th>
<th>valence</th>
<th>std dev</th>
<th>twitter rank</th>
<th>g-books rank</th>
<th>nyt rank</th>
<th>lyrics rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>laughter</td>
<td>8.50</td>
<td>0.93</td>
<td>3600</td>
<td>–</td>
<td>–</td>
<td>1728</td>
</tr>
<tr>
<td>2</td>
<td>happiness</td>
<td>8.44</td>
<td>0.97</td>
<td>1853</td>
<td>2458</td>
<td>328</td>
<td>1230</td>
</tr>
<tr>
<td>3</td>
<td>love</td>
<td>8.42</td>
<td>1.11</td>
<td>25</td>
<td>317</td>
<td>1313</td>
<td>375</td>
</tr>
<tr>
<td>4</td>
<td>happy</td>
<td>8.30</td>
<td>0.99</td>
<td>65</td>
<td>1372</td>
<td>1469</td>
<td>890</td>
</tr>
<tr>
<td>5</td>
<td>laughed</td>
<td>8.26</td>
<td>1.16</td>
<td>3334</td>
<td>3542</td>
<td>–</td>
<td>2332</td>
</tr>
<tr>
<td>6</td>
<td>laugh</td>
<td>8.22</td>
<td>1.37</td>
<td>1002</td>
<td>3998</td>
<td>4488</td>
<td>647</td>
</tr>
<tr>
<td>7</td>
<td>laughing</td>
<td>8.20</td>
<td>1.11</td>
<td>1579</td>
<td>–</td>
<td>–</td>
<td>1122</td>
</tr>
<tr>
<td>8</td>
<td>excellent</td>
<td>8.18</td>
<td>1.10</td>
<td>1496</td>
<td>1756</td>
<td>3155</td>
<td>–</td>
</tr>
<tr>
<td>9</td>
<td>laughs</td>
<td>8.18</td>
<td>1.16</td>
<td>3554</td>
<td>–</td>
<td>–</td>
<td>2856</td>
</tr>
<tr>
<td>10</td>
<td>joy</td>
<td>8.16</td>
<td>1.06</td>
<td>988</td>
<td>2336</td>
<td>2723</td>
<td>809</td>
</tr>
<tr>
<td>11</td>
<td>successful</td>
<td>8.16</td>
<td>1.08</td>
<td>2176</td>
<td>1198</td>
<td>1565</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>win</td>
<td>8.12</td>
<td>1.08</td>
<td>154</td>
<td>3031</td>
<td>776</td>
<td>694</td>
</tr>
<tr>
<td>13</td>
<td>rainbow</td>
<td>8.10</td>
<td>0.99</td>
<td>2726</td>
<td>–</td>
<td>–</td>
<td>1723</td>
</tr>
<tr>
<td>14</td>
<td>smile</td>
<td>8.10</td>
<td>1.02</td>
<td>925</td>
<td>2666</td>
<td>2898</td>
<td>349</td>
</tr>
<tr>
<td>15</td>
<td>won</td>
<td>8.10</td>
<td>1.22</td>
<td>810</td>
<td>1167</td>
<td>439</td>
<td>1493</td>
</tr>
<tr>
<td>16</td>
<td>pleasure</td>
<td>8.08</td>
<td>0.97</td>
<td>1497</td>
<td>1526</td>
<td>4253</td>
<td>1398</td>
</tr>
<tr>
<td>17</td>
<td>smiled</td>
<td>8.08</td>
<td>1.07</td>
<td>–</td>
<td>3537</td>
<td>–</td>
<td>2248</td>
</tr>
<tr>
<td>18</td>
<td>rainbows</td>
<td>8.06</td>
<td>1.36</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4216</td>
</tr>
<tr>
<td>19</td>
<td>winning</td>
<td>8.04</td>
<td>1.05</td>
<td>1876</td>
<td>–</td>
<td>1426</td>
<td>3646</td>
</tr>
<tr>
<td>20</td>
<td>celebration</td>
<td>8.02</td>
<td>1.53</td>
<td>3306</td>
<td>–</td>
<td>2762</td>
<td>4070</td>
</tr>
<tr>
<td>21</td>
<td>enjoyed</td>
<td>8.02</td>
<td>1.53</td>
<td>1530</td>
<td>2908</td>
<td>3502</td>
<td>–</td>
</tr>
<tr>
<td>22</td>
<td>healthy</td>
<td>8.02</td>
<td>1.06</td>
<td>1393</td>
<td>3200</td>
<td>3292</td>
<td>4619</td>
</tr>
<tr>
<td>23</td>
<td>music</td>
<td>8.02</td>
<td>1.12</td>
<td>132</td>
<td>875</td>
<td>167</td>
<td>374</td>
</tr>
<tr>
<td>24</td>
<td>celebrating</td>
<td>8.00</td>
<td>1.14</td>
<td>2550</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>25</td>
<td>congratulations</td>
<td>8.00</td>
<td>1.63</td>
<td>2246</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>26</td>
<td>weekend</td>
<td>8.00</td>
<td>1.29</td>
<td>317</td>
<td>–</td>
<td>833</td>
<td>2256</td>
</tr>
<tr>
<td>27</td>
<td>celebrate</td>
<td>7.98</td>
<td>1.15</td>
<td>1606</td>
<td>–</td>
<td>3574</td>
<td>2108</td>
</tr>
<tr>
<td>28</td>
<td>comedy</td>
<td>7.98</td>
<td>1.15</td>
<td>1444</td>
<td>–</td>
<td>2566</td>
<td>–</td>
</tr>
<tr>
<td>29</td>
<td>jokes</td>
<td>7.98</td>
<td>0.98</td>
<td>2812</td>
<td>–</td>
<td>–</td>
<td>3808</td>
</tr>
<tr>
<td>30</td>
<td>rich</td>
<td>7.98</td>
<td>1.32</td>
<td>1625</td>
<td>1221</td>
<td>1469</td>
<td>890</td>
</tr>
<tr>
<td>valence rank</td>
<td>word</td>
<td>valence</td>
<td>std dev</td>
<td>twitter rank</td>
<td>g-books rank</td>
<td>nyt rank</td>
<td>lyrics rank</td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>10193</td>
<td>violence</td>
<td>1.86</td>
<td>1.05</td>
<td>4299</td>
<td>1724</td>
<td>1238</td>
<td>2016</td>
</tr>
<tr>
<td>10194</td>
<td>cruel</td>
<td>1.84</td>
<td>1.15</td>
<td>2963</td>
<td>–</td>
<td>–</td>
<td>1447</td>
</tr>
<tr>
<td>10195</td>
<td>cry</td>
<td>1.84</td>
<td>1.28</td>
<td>1028</td>
<td>3075</td>
<td>–</td>
<td>226</td>
</tr>
<tr>
<td>10196</td>
<td>failed</td>
<td>1.84</td>
<td>1.00</td>
<td>2645</td>
<td>1618</td>
<td>1276</td>
<td>2920</td>
</tr>
<tr>
<td>10197</td>
<td>sickness</td>
<td>1.84</td>
<td>1.18</td>
<td>4735</td>
<td>–</td>
<td>–</td>
<td>3782</td>
</tr>
<tr>
<td>10198</td>
<td>abused</td>
<td>1.83</td>
<td>1.31</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4589</td>
</tr>
<tr>
<td>10199</td>
<td>tortured</td>
<td>1.82</td>
<td>1.42</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4693</td>
</tr>
<tr>
<td>10200</td>
<td>fatal</td>
<td>1.80</td>
<td>1.53</td>
<td>–</td>
<td>4089</td>
<td>–</td>
<td>3724</td>
</tr>
<tr>
<td>10201</td>
<td>killings</td>
<td>1.80</td>
<td>1.54</td>
<td>–</td>
<td>–</td>
<td>4914</td>
<td>–</td>
</tr>
<tr>
<td>10202</td>
<td>murdered</td>
<td>1.80</td>
<td>1.63</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4796</td>
</tr>
<tr>
<td>10203</td>
<td>war</td>
<td>1.80</td>
<td>1.41</td>
<td>468</td>
<td>175</td>
<td>291</td>
<td>462</td>
</tr>
<tr>
<td>10204</td>
<td>kills</td>
<td>1.78</td>
<td>1.23</td>
<td>2459</td>
<td>–</td>
<td>–</td>
<td>2857</td>
</tr>
<tr>
<td>10205</td>
<td>jail</td>
<td>1.76</td>
<td>1.02</td>
<td>1642</td>
<td>–</td>
<td>2573</td>
<td>1619</td>
</tr>
<tr>
<td>10206</td>
<td>terror</td>
<td>1.76</td>
<td>1.00</td>
<td>4625</td>
<td>4117</td>
<td>4048</td>
<td>2370</td>
</tr>
<tr>
<td>10207</td>
<td>die</td>
<td>1.74</td>
<td>1.19</td>
<td>418</td>
<td>730</td>
<td>2605</td>
<td>143</td>
</tr>
<tr>
<td>10208</td>
<td>killing</td>
<td>1.70</td>
<td>1.36</td>
<td>1507</td>
<td>4428</td>
<td>1672</td>
<td>998</td>
</tr>
<tr>
<td>10209</td>
<td>arrested</td>
<td>1.64</td>
<td>1.01</td>
<td>2435</td>
<td>4474</td>
<td>1435</td>
<td>–</td>
</tr>
<tr>
<td>10210</td>
<td>deaths</td>
<td>1.64</td>
<td>1.14</td>
<td>–</td>
<td>–</td>
<td>2974</td>
<td>–</td>
</tr>
<tr>
<td>10211</td>
<td>raped</td>
<td>1.64</td>
<td>1.43</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4528</td>
</tr>
<tr>
<td>10212</td>
<td>torture</td>
<td>1.58</td>
<td>1.05</td>
<td>3175</td>
<td>–</td>
<td>–</td>
<td>3126</td>
</tr>
<tr>
<td>10213</td>
<td>died</td>
<td>1.56</td>
<td>1.20</td>
<td>1223</td>
<td>866</td>
<td>208</td>
<td>826</td>
</tr>
<tr>
<td>10214</td>
<td>kill</td>
<td>1.56</td>
<td>1.05</td>
<td>798</td>
<td>2727</td>
<td>2572</td>
<td>430</td>
</tr>
<tr>
<td>10215</td>
<td>killed</td>
<td>1.56</td>
<td>1.23</td>
<td>1137</td>
<td>1603</td>
<td>814</td>
<td>1273</td>
</tr>
<tr>
<td>10216</td>
<td>cancer</td>
<td>1.54</td>
<td>1.07</td>
<td>946</td>
<td>1884</td>
<td>796</td>
<td>3802</td>
</tr>
<tr>
<td>10217</td>
<td>death</td>
<td>1.54</td>
<td>1.28</td>
<td>509</td>
<td>307</td>
<td>373</td>
<td>433</td>
</tr>
<tr>
<td>10218</td>
<td>murder</td>
<td>1.48</td>
<td>1.01</td>
<td>2762</td>
<td>3110</td>
<td>1541</td>
<td>1059</td>
</tr>
<tr>
<td>10219</td>
<td>terrorism</td>
<td>1.48</td>
<td>0.91</td>
<td>–</td>
<td>–</td>
<td>3192</td>
<td>–</td>
</tr>
<tr>
<td>10220</td>
<td>rape</td>
<td>1.44</td>
<td>0.79</td>
<td>3133</td>
<td>–</td>
<td>4115</td>
<td>2977</td>
</tr>
<tr>
<td>10221</td>
<td>suicide</td>
<td>1.30</td>
<td>0.84</td>
<td>2124</td>
<td>4707</td>
<td>3319</td>
<td>2107</td>
</tr>
<tr>
<td>10222</td>
<td>terrorist</td>
<td>1.30</td>
<td>0.91</td>
<td>3576</td>
<td>–</td>
<td>3026</td>
<td>–</td>
</tr>
<tr>
<td>rank</td>
<td>word</td>
<td>valence</td>
<td>std dev</td>
<td>twitter rank</td>
<td>g-books rank</td>
<td>nyt rank</td>
<td>lyrics rank</td>
</tr>
<tr>
<td>------</td>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>1</td>
<td>@king</td>
<td>4.64</td>
<td>2.93</td>
<td>448</td>
<td>-</td>
<td>-</td>
<td>620</td>
</tr>
<tr>
<td>2</td>
<td>☆☆kin</td>
<td>3.86</td>
<td>2.74</td>
<td>1077</td>
<td>-</td>
<td>-</td>
<td>688</td>
</tr>
<tr>
<td>3</td>
<td>☆☆ked</td>
<td>3.56</td>
<td>2.71</td>
<td>1840</td>
<td>-</td>
<td>-</td>
<td>904</td>
</tr>
<tr>
<td>4</td>
<td>pussy</td>
<td>4.80</td>
<td>2.66</td>
<td>2019</td>
<td>-</td>
<td>-</td>
<td>949</td>
</tr>
<tr>
<td>5</td>
<td>whiskey</td>
<td>5.72</td>
<td>2.64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2208</td>
</tr>
<tr>
<td>6</td>
<td>slut</td>
<td>3.57</td>
<td>2.63</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4071</td>
</tr>
<tr>
<td>7</td>
<td>cigarettes</td>
<td>3.31</td>
<td>2.60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3279</td>
</tr>
<tr>
<td>8</td>
<td>☆☆k</td>
<td>4.14</td>
<td>2.58</td>
<td>322</td>
<td>-</td>
<td>-</td>
<td>185</td>
</tr>
<tr>
<td>9</td>
<td>mortality</td>
<td>4.38</td>
<td>2.55</td>
<td>-</td>
<td>3960</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>cigarette</td>
<td>3.09</td>
<td>2.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2678</td>
</tr>
<tr>
<td>11</td>
<td>motherf☆☆kers</td>
<td>2.51</td>
<td>2.47</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1466</td>
</tr>
<tr>
<td>12</td>
<td>churches</td>
<td>5.70</td>
<td>2.46</td>
<td>-</td>
<td>2281</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>motherf☆☆king</td>
<td>2.64</td>
<td>2.46</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2910</td>
</tr>
<tr>
<td>14</td>
<td>capitalism</td>
<td>5.16</td>
<td>2.45</td>
<td>-</td>
<td>4648</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>porn</td>
<td>4.18</td>
<td>2.43</td>
<td>1801</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>summer</td>
<td>6.40</td>
<td>2.39</td>
<td>896</td>
<td>1226</td>
<td>721</td>
<td>590</td>
</tr>
<tr>
<td>17</td>
<td>beer</td>
<td>5.92</td>
<td>2.39</td>
<td>839</td>
<td>4924</td>
<td>3960</td>
<td>1413</td>
</tr>
<tr>
<td>18</td>
<td>execution</td>
<td>3.10</td>
<td>2.39</td>
<td>-</td>
<td>2975</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>wines</td>
<td>6.28</td>
<td>2.37</td>
<td>-</td>
<td>-</td>
<td>3316</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>zombies</td>
<td>4.00</td>
<td>2.37</td>
<td>4708</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>aids</td>
<td>4.28</td>
<td>2.35</td>
<td>2983</td>
<td>3996</td>
<td>1197</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>capitalist</td>
<td>4.84</td>
<td>2.34</td>
<td>-</td>
<td>4694</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>revenge</td>
<td>3.71</td>
<td>2.34</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2766</td>
</tr>
<tr>
<td>24</td>
<td>mcdonalds</td>
<td>5.98</td>
<td>2.33</td>
<td>3831</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>beatles</td>
<td>6.44</td>
<td>2.33</td>
<td>3797</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>islam</td>
<td>4.68</td>
<td>2.33</td>
<td>-</td>
<td>4514</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>pay</td>
<td>5.30</td>
<td>2.32</td>
<td>627</td>
<td>769</td>
<td>460</td>
<td>499</td>
</tr>
<tr>
<td>28</td>
<td>alcohol</td>
<td>5.20</td>
<td>2.32</td>
<td>2787</td>
<td>2617</td>
<td>3752</td>
<td>3600</td>
</tr>
<tr>
<td>29</td>
<td>muthaf☆☆kin</td>
<td>3.00</td>
<td>2.31</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4107</td>
</tr>
<tr>
<td>30</td>
<td>christ</td>
<td>6.16</td>
<td>2.31</td>
<td>2509</td>
<td>909</td>
<td>4238</td>
<td>1526</td>
</tr>
</tbody>
</table>
Positive bias in the English language:
For more...

- PSD, KDH, IMK, CAB, and CMD
  “Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter.”

- P. S. Dodds and C. M. Danforth
  “Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents.”
  [Journal of Happiness Studies, 2009.](http://www.uvm.edu/~pdodds/research/)

- [http://www.onehappybird.com](http://www.onehappybird.com)

- “Does a Nation’s Mood Lurk in Its Songs and Blogs?” by Benedict Carey
References


References II


References III


References IV

Science Magazine, 331:176–182, 2011. [pdf](inya)


References V

On Growth and From.  


Complex Social Networks.  

Six Degrees.  
Collective dynamics of ‘small-world’ networks. 

The unreasonable effectivenss of mathematics in the natural sciences. 