Fundamentals

Last updated: 2021/10/25, 11:55:37 EDT

Principles of Complex Systems, Vols. 1 & 2
CSYS/MATH 300 and 303, 2021–2022 | @pocsvox

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

No really, that's a lot of data

Data inflation 2

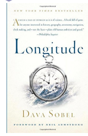
Unit	Size	What it means
Bit (b)	1 or 0	Short for "binary digit", after the binary code (1 or 0) computers use to store and process data
Byte (B)	8 bits	Enough information to create an English letter or number in computer code. It is the basic unit of computing
Kilobyte (KB)	1,000, or 2^{10}, bytes	From "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*

A brief history of measuring time:

- Megaliths for Big Time
- Sundials, 1500 BC, Egypt (solid for over 2000 years)
- Escapements (200s), Hourglasses (1300s?), Pendulum clocks (Galileo, 1500s)
- Chronometers, 1700s:



"Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time" by Dava Sobel (2007). [16]

- Billionths of a second accuracy: Atomic clocks (Lord Kelvin, 1879)

PoCS
@pocsvox
Fundamentals

Data
Measurement
Emergence
Self-Organization
Modeling
Statistical Mechanics
Nutshell
References

11 of 74

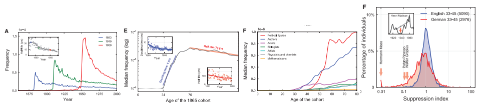
Outline

PoCS
@pocsvox
Fundamentals

2 of 74

Big Data—Culturomics:

"Quantitative analysis of culture using millions of digitized books" by Michel et al., Science, 2011 [9]



- http://www.culturomics.org/ and Google Books ngram viewer

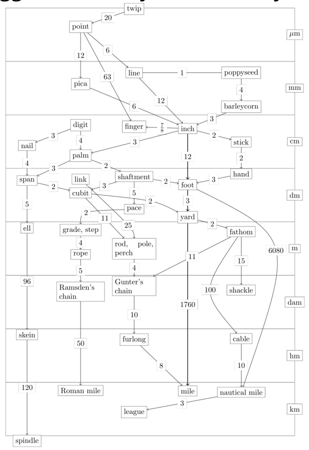
Barney Rubble:

"Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution"
Pechenick, Danforth, and Dodds, PLoS ONE, **10**, e0137041, 2015. [10]

PoCS
@pocsvox
Fundamentals

9 of 74

Our struggle to sensibly measure anything at all:

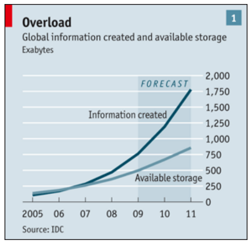


By 42CrMo4, Christoph Päper – English length units graph (PNG), CC BY-SA 4.0
https://commons.wikimedia.org/w/index.php?curid=61338012
From https://en.wikipedia.org/wiki/Barleycorn_(unit)

PoCS
@pocsvox
Fundamentals

12 of 74

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 = 10^3EB = 10^6PB = 10^9TB)
- Large Hadron Collider: 40 TB/second.
- 2016—Large Synoptic Survey Telescope: 140 TB every 5 days.
- Facebook: $\sim$ 250 billion photos (mid 2013)
- Twitter: $\sim$ 500 billion tweets (mid 2013)

PoCS
@pocsvox
Fundamentals

7 of 74

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."

PoCS
@pocsvox
Fundamentals

10 of 74

Measuring temperature was thought impossible:

The properties measured by our instruments usually begin as subjective judgments. Temperature is a good example. People were aware of variations in temperature long before there were any objective measurements of temperature. Judgments of temperature are imperfectly correlated among different persons, or even the same person at different times, depending on the humidity, the person's activity level and age, surrounding air currents, and so on. The idea that anything as subtle and complex as all the manifestations of changes in temperature could be measured and quantified on a single numerical scale was scoffed at as impossible, even by the leading philosophers of the sixteenth century.

The first thermometer invented by Galileo in 1592 did not go far in dispelling the notion that temperature was inherently unmeasurable, because the earliest thermometers, for about their first hundred years, were so imperfect as to make it possible for those who wished to do so to argue that no one could ever succeed in measuring temperature. Temperature was then confounded with all the subtleties of subjective judgment, which easily seem incompatible with a single numerical scale of measurement. How could the height of a column of mercury in a glass tube possibly reflect the rich varieties of temperature—damp cold, dank cold, frosty cold, crisp cold, humid heat, searing heat, scalding heat, dry heat, feverish heat, prickly heat, and so on?

From "Bias in Mental Testing", Arthur Jensen, 1980 [8]
per @SilverVVulpes: Also: Inventing Temperature, Hasok Chang, 2004 [3]

PoCS
@pocsvox
Fundamentals

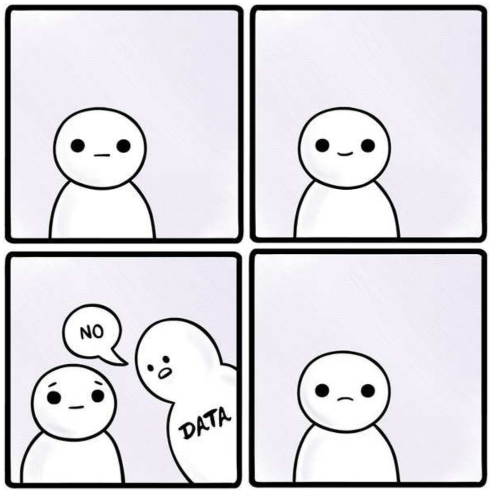
13 of 74

Measuring temperature was thought impossible:

The early thermometers were inconsistent, both with themselves and with each other. Because they consisted of open-ended glass tubes, they were sensitive to changes in barometric pressure as well as to temperature. And there were problems of calibration, such as where to locate the zero point and how to divide the column of mercury into units. It was believed, incorrectly, that all caves had the same temperature, so thermometers were calibrated in caves. The freezing and boiling points of water were also used in calibration, but, as these vary with impurities in the water and the barometric pressure, the calibration of different thermometers at different times and places resulted in thermometers that failed to correlate perfectly with one another in any given instance. They lacked reliability, as we now would say.

All the while, no one knew what temperature is in a theoretical or scientific sense. There was no theory of thermodynamics that could explain temperature phenomena and provide a scientific rationale for the construction and calibration of thermometers. Yet quite adequate and accurate thermometers, hardly differing from those we use today, were eventually developed by the middle of the eighteenth century. Thus the objective measurement of temperature considerably preceded the development of an adequate theory of temperature and heat, and necessarily so, as the science of thermodynamics could not possibly have developed without first having been able to quantify or measure the temperatures of liquids, gasses, and other substances independently of

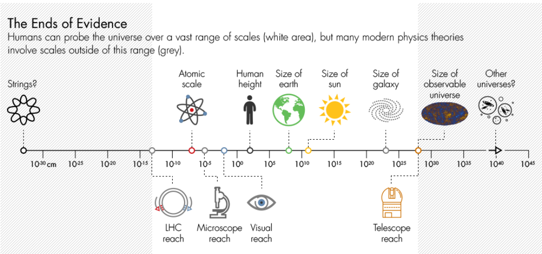
From "Bias in Mental Testing", Arthur Jensen, 1980 [8] per @SilverVVulpes: Also: Inventing Temperature, Hasok Chang, 2004 [3]



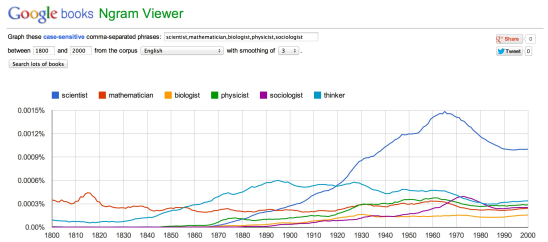
THIS COMIC MADE POSSIBLE THANKS TO ADAM LINGELBACH MRLOVENSTEIN.COM

Limits of testability and happiness in Science:

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



The Newness of being a Scientist (1833 on):



- Etymology here.
- "Scientists are the people who ask a question about a phenomenon and proceed to systematically go about answering the question themselves. They are by nature curious, creative and well organized."

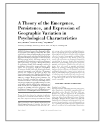
Please do not measure complex systems with one number:

- This is real—someone having some fun.
- Obtained from this tweet.
- Sadness for Buckingham (if Buckingham has no sense of humor).

The conceptual trapping pit of a single scale:

- Lure of simplicity: Comparisons and rankings are easy.
- A single scale measure is very appealing, very hard to resist ... ... and hard to push back against when widely adopted.
- Examples:
  - Grade point average (GPA)
  - College rankings, City rankings, Country rankings, Wine scores, Michelin Guide, Yelp scores, Amazon ratings, ...
  - Body Mass Index (BMI)
  - Intelligence Quotient (IQ)[1]
  - Effective temperature
  - Price for all things: One dimension of belief
  - Salary!
  - stock market valuation for corporations
  - Complexity of civilizations [17]
  - A 1-d axis for political ideologies (a spatial metaphor trap, thanks France!)

Personality distributions:



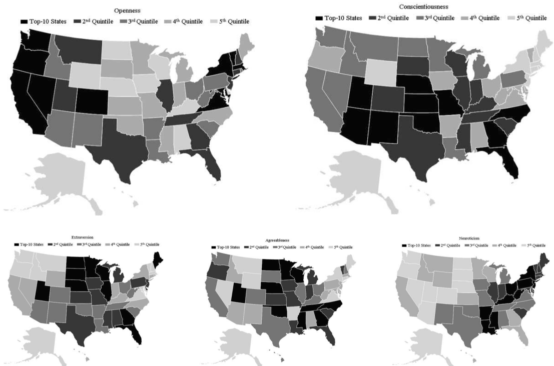
"A Theory of the Emergence, Persistence, and Expression of Geographic Variation in Psychological Characteristics"
Renfrow, Gosling, and Potter, Perspectives on Psychological Science, **3**, 339–369, 2008. [11]

Five Factor Model (FFM):

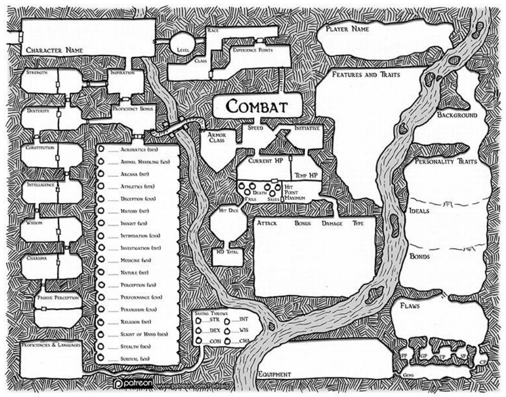
- Extraversion [E]
- Agreeableness [A]
- Conscientiousness [C]
- Neuroticism [N]
- Openness [O]

"...a robust and widely accepted framework for conceptualizing the structure of personality... Although the FFM is not universally accepted in the field..." [11]

A concern: self-reported data. Bigger concern: mass manipulation.



Dungeons & Dragons' full embrace of complexity:



From here.

Dungeons & Dragons—Two alignment axes for character:

[2]

Law–Chaos (vertical) and Good–Evil (horizontal).

[2] From this Reddit thread, where, naturally, the choices are enthusiastically debated.

Emergence:

The Wikipedia on Emergence (2006):

“In philosophy, systems theory and the sciences, emergence refers to the way complex systems and patterns arise out of a multiplicity of relatively simple interactions. ... emergence is central to the physics of complex systems and yet very controversial.”

Wikipedia, 2016:

In philosophy, systems theory, science, and art, emergence is a process whereby larger entities arise through interactions among smaller or simpler entities such that the larger entities exhibit properties the smaller/simpler entities do not exhibit.

The philosopher G. H. Lewes first used the word explicity in 1875.

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
- Neurons etc. ⇒ Brain ⇒ Thoughts
- People ⇒ Religion, Collective behaviour
- People ⇒ The Web
- People ⇒ Language, and rules of language
- ? ⇒ time; ? ⇒ gravity; ? ⇒ reality.

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

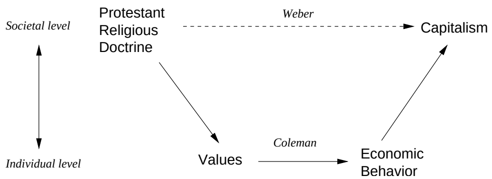
Emergence:

Friedrich Hayek (Economist/Philospher/Nobelist):

- Markets, legal systems, political systems are emergent and not designed.
- ‘Taxis’ = made order (by God, Sovereign, Government, ...)
- ‘Cosmos’ = grown order
- Archetypal limits of hierarchical and decentralized structures.
- Hierarchies arise once problems are solved. [5]
- Decentralized structures help solve problems.
- Dewey Decimal System versus tagging.

Emergence:

James Coleman in *Foundations of Social Theory*:



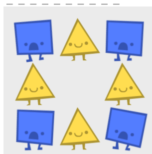
- Understand macrophenomena arises from microbehavior which in turn depends on macrophenomena. [4]
- More on Coleman here.

Emergence:

Thomas Schelling (Economist/Nobelist):

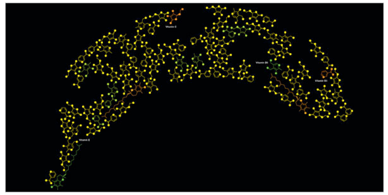
- “Micromotives and Macrobehavior” [14]
  - Segregation [12, 15]
  - Wearing hockey helmets [13]
  - Seating choices

Vi Hart and Nicky Case’s Polygon-themed visualization:



The emergence of taste:

- Molecules ⇒ Ingredients ⇒ Taste
- See Michael Pollan’s article on nutritionism in the New York Times, January 28, 2007.



nytimes.com

Reductionism

Reductionism and food:

- Pollan: “even the simplest food is a hopelessly complex thing to study, a virtual wilderness of chemical compounds, many of which exist in complex and dynamic relation to one another...”
- “So ... break the thing down into its component parts and study those one by one, even if that means ignoring complex interactions and contexts, as well as the fact that the whole may be more than, or just different from, the sum of its parts. This is what we mean by reductionist science.”

Reductionism

- “people don’t eat nutrients, they eat foods, and foods can behave very differently than the nutrients they contain.”
- Studies suggest diets high in fruits and vegetables help prevent cancer.
- So... find the nutrients responsible and eat more of them
- But “in the case of beta carotene ingested as a supplement, scientists have discovered that it actually increases the risk of certain cancers. Oops.”

Reductionism

Thyme's known antioxidants:

4-Terpineol, alanine, anethole, apigenin, ascorbic acid, beta carotene, caffeic acid, camphene, carvacrol, chlorogenic acid, chrysoeriol, eriodictyol, eugenol, ferulic acid, gallic acid, gamma-terpinene isochlorogenic acid, isoeugenol, isothymonin, kaempferol, labiatic acid, lauric acid, linalyl acetate, luteolin, methionine, myrcene, myristic acid, naringenin, oleanolic acid, p-coumoric acid, p-hydroxy-benzoic acid, palmitic acid, rosmarinic acid, selenium, tannin, thymol, tryptophan, ursolic acid, vanillic acid.

[cnn.com]

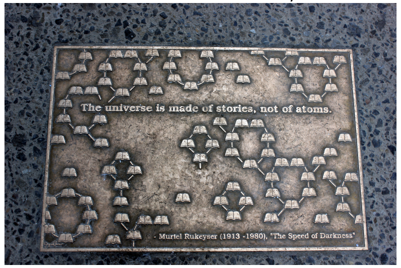
Reductionism

"It would be great to know how this all works, but in the meantime we can enjoy thyme in the knowledge that it probably doesn't do any harm (since people have been eating it forever) and that it may actually do some good (since people have been eating it forever) and that even if it does nothing, we like the way it tastes."

Gulf between theory and practice (see baseball and bumblebees).

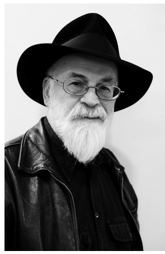
This is a Collateralized Debt Obligation:

- "The Universe is made of stories, not of atoms."



- From "The Speed of Darkness" (1968) by Muriel Rukeyser
- Quoted by Metatron in Supernatural, Meta Fiction, S9E18.

(Sir Terry) Pratchett's Narrativium:



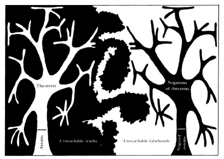
- "The most common element on the disc, although not included in the list of the standard five: earth, fire, air, water and surprise. It ensures that everything runs properly as a story."
- "A little narrativium goes a long way: the simpler the story, the better you understand it. Storytelling is the opposite of reductionism: 26 letters and some rules of grammar are no story at all."

Higher complexity:

- Many system scales (or levels) that interact with each other.
- Potentially much harder to explain/understand.

Even mathematics: [6]

Gödel's Theorem: we can't prove every theorem that's true ...



"Gödel, Escher, Bach" [7]

- Suggests a strong form of emergence: Some phenomena cannot be analytically deduced from elementary aspects of a system.

Emergence:

Roughly speaking, there are two types of emergence:

I. Weak emergence:

System-level phenomena is different from that of its constituent parts yet can be connected theoretically.

II. Strong emergence:

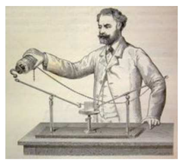
System-level phenomena fundamentally cannot be deduced from how parts interact.

Emergence:

- Reductionist techniques can explain weak emergence.
- Magic explains strong emergence. [2]
- But: maybe magic should be interpreted as an inscrutable yet real mechanism that cannot ever be simply described.
- Gulp.

Limits of Science | Radiolab

Listen to Steve Strogatz, Hod Lipson, and Michael Schmidt (Cornell) in the last piece (11:16) on Radiolab's show 'Limits' (April 5, 2010).



(El Bibliomata/flickr)

Dr. Steve Strogatz wonders if we've reached the limits of human scientific understanding, and should soon turn the reins of research over to robots. Cold, calculating robots. Then, Dr. Hod Lipson and Michael Schmidt walk us through the workings of a revolutionary computer program that they developed--a program that can deduce mathematical relationships in nature, through simple observation. The catch? As Dr. Gurol Suel explains, the program gives answers to complex biological questions that we humans have yet to ask, or even to understand.

TAGS: mind bending

Pair with some slow tv
Bonus: Mike Schmidt's talk on Eureqa at UVM's 2011 TEDx event "Big Data, Big Stories."

Definitions

"Self-organization is a process in which the internal organization of a system, normally an open system, increases in complexity without being guided or managed by an outside source." (also: Self-assembly)

Examples:

- Molecules/Atoms liking each other → Gases, liquids, and solids.
- Spin alignment → Magnetization.
- Protein folding.
- Imitation → Herding, flocking, mobs, ...

Fundamental question: how likely is 'complexification'?

Tools and techniques:

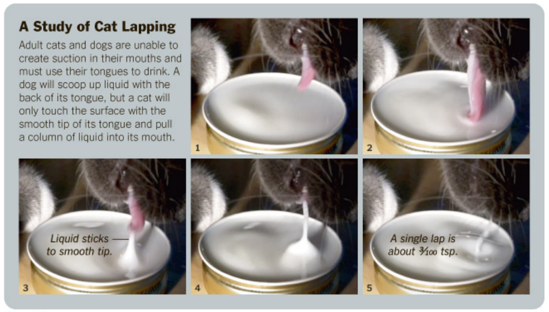
- Differential equations, difference equations, linear algebra, stochastic models.
- Statistical techniques for comparisons and descriptions.
- Methods from statistical mechanics and computer science.
- Machine learning (but beware the black box).
- Computer modeling, everything from
  - Artisanal toy models
  - to kitchen sink models.

Key advance (more soon):

- Representation of complex interaction patterns as complex networks.
- The driver: Massive amounts of Data

Rather silly but great example of real science:

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



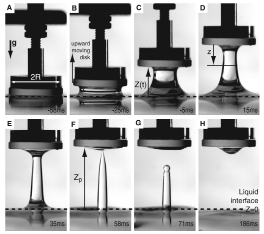
Source: Science THE NEW YORK TIMES; IMAGES FROM VIDEO BY ROMAN STOCKER, SUNGHWAN JUNG, JEFFREY M. ARISTOFF AND PEDRO M. REIS

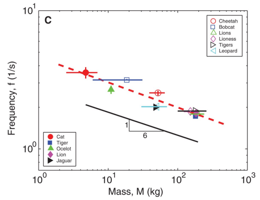
Amusing interview here

- Another great, great moment in scaling:

$$f \sim M^{-1/6}$$

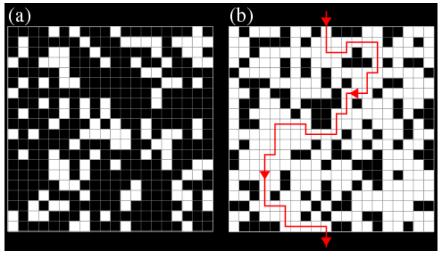
The balance of inertia and gravity yields a prediction for the lapping frequency of other felines. Assuming isometry within the Felidae family (i.e., that lapping height H scales linearly with tongue width R and animal mass M scales as R^3), the finding that Fr* is of order one translates to the prediction $f \sim R^{-1/2} \sim M^{-1/6}$. Isometry or marginally positive allomety among the Felidae has been demonstrated for skull (*20*, *21*) and limb bones (*22*). Although variability by function can lead to departures from isometry in interspecific scalings (*23*), reported variations within the Felidae (*23*, *24*) only minimally affect the predicted scaling $f \sim M^{-1/6}$. We tested this –1/6 power-law dependence by measuring the lapping frequency for eight species of felines, from videos acquired at the Zoo New England or available on YouTube (*16*). The lapping frequency was observed to decrease with animal mass as $f = 4.6\ M^{-0.181 \pm 0.024}$ (f in s^{-1}, M in kg) (Fig. 4C), close to the predicted $M^{-1/6}$. This close agreement suggests that the domestic cat's inertia- and gravity-controlled lapping mechanism is conserved among felines.





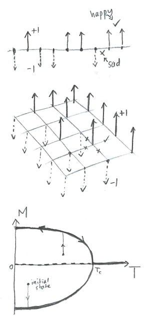
- Statistical Mechanics is "a science of collective behavior."
- Simple rules give rise to collective phenomena.

Percolation:



Snared from Michael Gastner's page on percolation [no longer online]

The Ising Model of a ferromagnet:

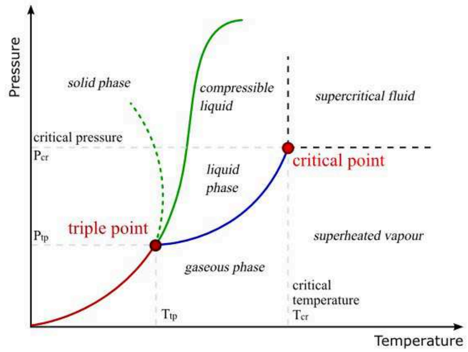


- Each atom is assumed to have a local spin that can be up or down: $S_i = \pm 1$.
- Spins are assumed to be arranged on a lattice.
- In isolation, spins like to align with each other.
- Increasing temperature breaks these alignments.
- The drosophila of statistical mechanics.
- Criticality: Power-law distributions at critical points.

Example 2-d Ising model simulation:

https://mattbierbaum.github.io/ising.js/

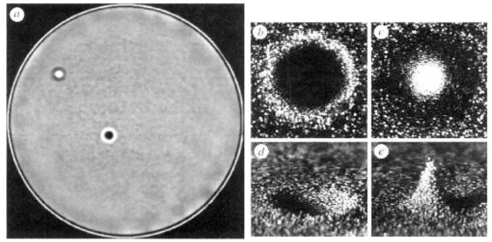
Phase diagrams



Qualitatively distinct macro states.

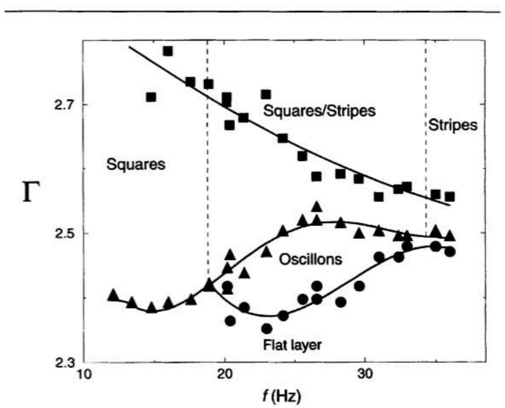
Phase diagrams

Oscillons, bacteria, traffic, snowflakes, ...

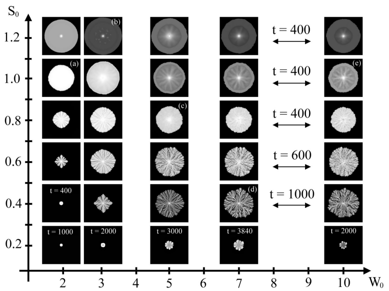


Umbanhowar et al., *Nature*, 1996 [18]

Phase diagrams



Phase diagrams



W_0 = initial wetness, S_0 = initial nutrient supply
http://math.arizona.edu/~lega/HydroBact.html

Ising model

Analytic issues:

- 1-d: simple (Ising & Lenz, 1925)
- 2-d: hard (Onsager, 1944)
- 3-d: extremely hard...
- 4-d and up: simple.

- See lower and upper critical dimension in statistical physics.
- Also: Curse and Blessing of Dimensionality

Statistics

Historical surprise:

- Origins of Statistical Mechanics are in the studies of people... (Maxwell and co.)
- Now physicists are using their techniques to study everything else including people...
- See Philip Ball's "Critical Mass"[1]

Beyond Statistical Mechanics:

- Analytic approaches have their limits, especially in evolutionary, algorithm-rich systems.
- Algorithmic methods and simulation techniques will continue to rise in importance.

Nutshell

- The central concepts Complexity and Emergence are reasonably well defined.
- There is no general theory of Complex Systems.
- But the problems exist...
  Complex (Adaptive) Systems abound...
- And the observation of Universality of dynamical systems, statistical mechanics, and other quantitative areas means not everything is special and different.
- Framing from the Manifesto: Science's focus is moving to Complex Systems because it finally can.
- We use whatever tools we need.
- Science $\simeq$ Describe + Explain.

References I

[1] P. Ball.
Critical Mass: How One Thing Leads to Another.
Farra, Straus, and Giroux, New York, 2004.

[2] M. A. Bedau.
Weak emergence.
In J. Tomberlin, editor, Philosophical Perspectives: Mind, Causation, and World, volume 11, pages 375–399. Blackwell, Malden, MA, 1997. pdf

[3] H. Chang.
Inventing temperature: Measurement and scientific progress.
Oxford University Press, 2004.

[4] J. S. Coleman.
Foundations of Social Theory.
Belknap Press, Cambridge, MA, 1994.

References II

[5] P. S. Dodds, D. J. Watts, and C. F. Sabel.
Information exchange and the robustness of organizational networks.
Proc. Natl. Acad. Sci., 100(21):12516–12521, 2003. pdf

[6] R. Foote.
Mathematics and complex systems.
Science, 318:410–412, 2007. pdf

[7] D. R. Hofstadter.
Gödel, Escher, Bach.
Vintage Books, New York, 1980.

[8] A. R. Jensen.
Bias in mental testing.
1980.

References III

[9] 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.
Science Magazine, 331:176–182, 2011. pdf

[10] E. A. Pechenick, C. M. Danforth, and P. S. Dodds.
Characterizing the google books corpus: Strong limits to inferences of socio-cultural and linguistic evolution.
PLoS ONE, 10:e0137041, 2015. pdf

References IV

[11] 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

[12] T. C. Schelling.
Dynamic models of segregation.
J. Math. Sociol., 1:143–186, 1971. pdf

[13] T. C. Schelling.
Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities.
J. Conflict Resolut., 17:381–428, 1973. pdf

References V

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

[15] T. C. Schelling.
Some fun, thirty-five years ago.
In L. Tesfatsion and K. L. Judd, editors, Handbook of Computational Economics, volume 2, pages 1639–1644. Elsevier, 2006. pdf

[16] D. Sobel.
Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time.
Bloomsbury Publishing, US, 2007.

References VI

[17] P. Turchin, T. E. Currie, H. Whitehouse, P. François, K. Feeney, D. Mullins, D. Hoyer, C. Collins, S. Grohmann, P. Savage, et al.
Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization.
Proceedings of the National Academy of Sciences, 115:E144–E151, 2018. pdf

[18] P. B. Umbanhowar, F. Melo, and H. L. Swinney.
Localized excitations in a vertically vibrated granular layer.
Nature, 382:793–6, 1996. pdf