Artificial Life & Complex Systems

Lecture 2

On Complexity and Other Stuff

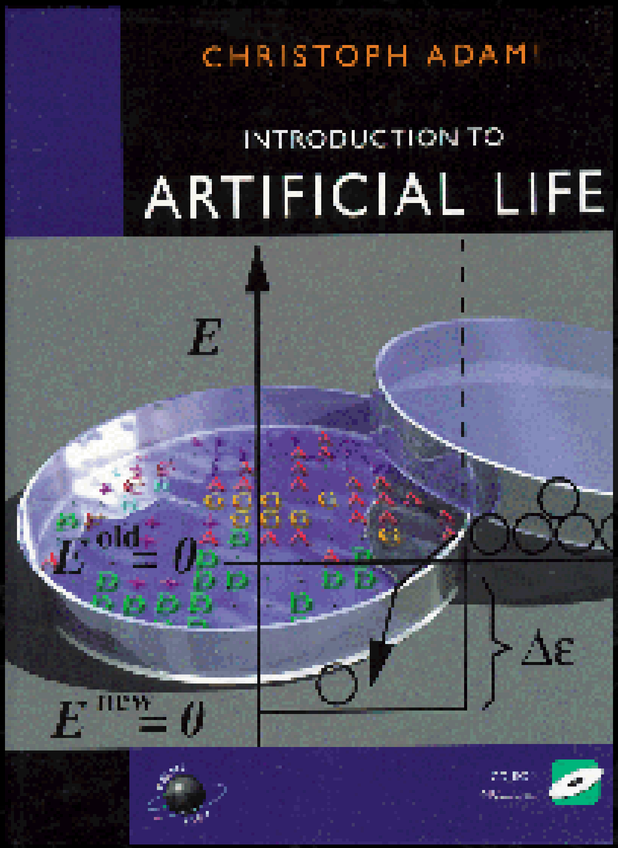
May 25

Max Lungarella

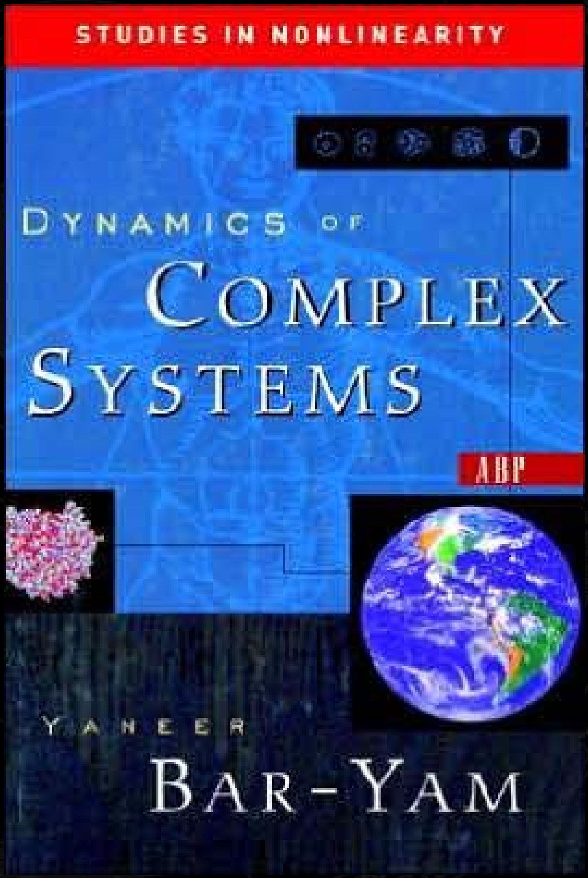
Contents

- What is complexity?
- Relationship to life and Alife
- Characterizations and examples

For The Curious Mind



Chris Adami (1997)



Yaneer Bar-Yam (1997)

Definition of Alife

From previous lecture:

Artificial Life is about the synthetic study of complex adaptive (life-like) systems (natural and artificial ones).

But What Does It Mean To Be ‘Alive’?

Is there a way to quantify ‘aliveness’?

Observation

Everything in the universe exists on a sliding scale of entropy (measure of disorder). Things very cold and still are at one end and things muddled, hot and moving rapidly are at the other.

Life sits in the middle some where.

Life is about the fight against entropy.

But …

Entropy = Measure of Disorder

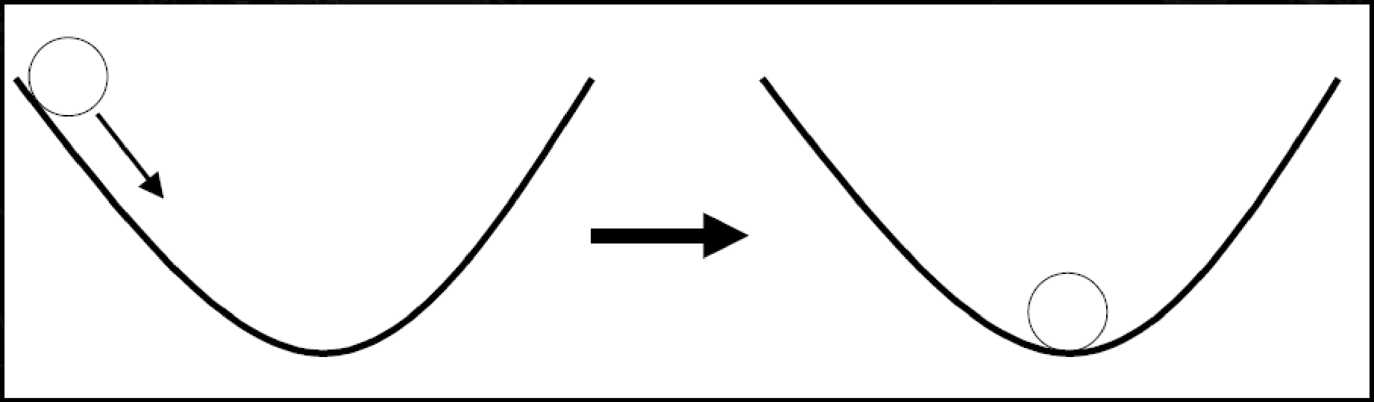
What is order?

Q: When is a System Ordered?

A: When a system is in balance at a stable of equilibrium

Main characteristics:

- Little disturbances have no consequences
- The system response is proportional to the impact
- Dramatic disturbances can cause state transitions



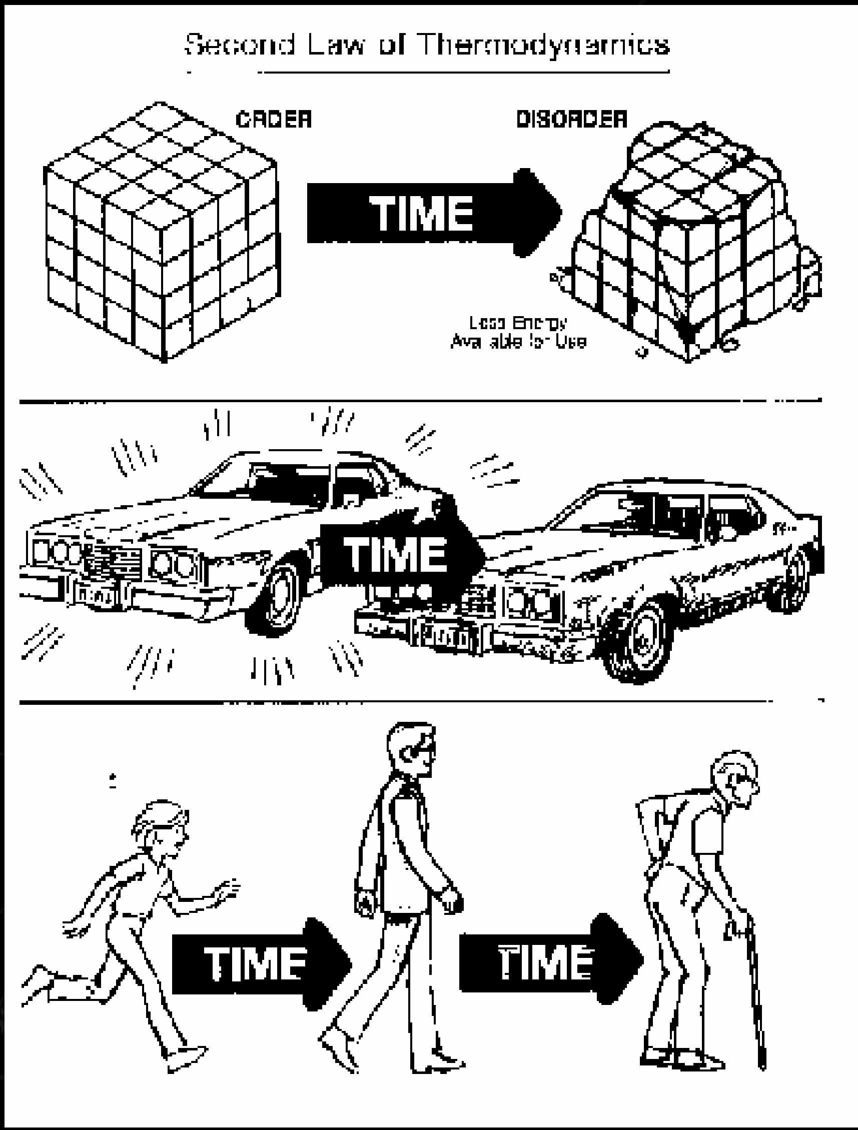
What Kind of Systems are Stable?

- Ice crystals
- Sand grain on floor
- According to economic theory: economic systems (general equilibrium theory)

2^{nd} Law of Thermodynamics

The second law of thermodynamics states that, if a system is a closed system, entropy (amount of disorder) will always increase and the whole become more disordered

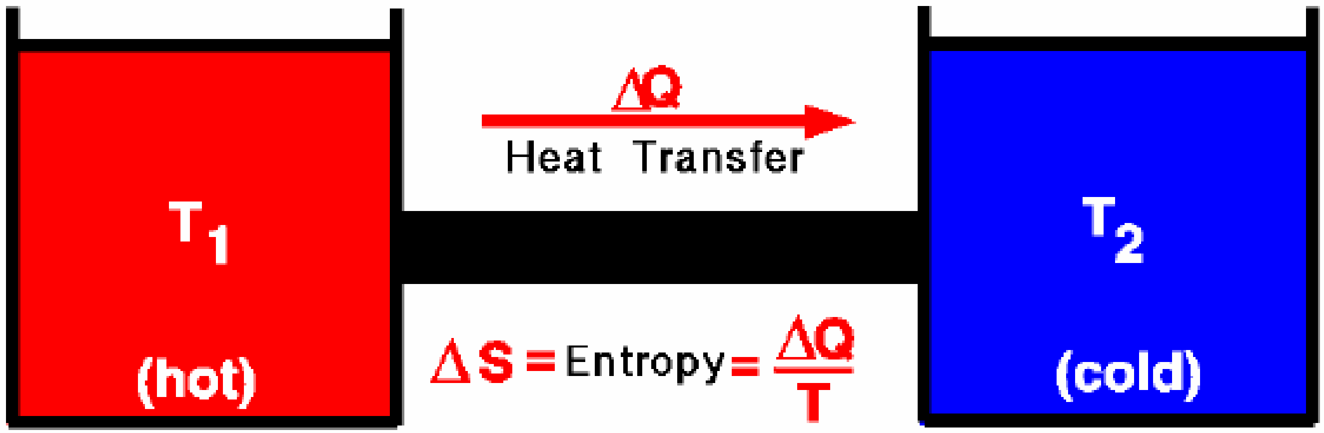
2^{nd} Law of Thermodynamics



2nd Law of Thermodynamics

Second Law of Thermodynamics

Glenn Research Center



There exists a useful thermodynamic variable called entropy (S). A natural process that starts in one equilibrium state and ends in another will go in the direction that causes the entropy of the system plus the environment to increase for an irreversible process and to remain constant for a reversible process.

$S_f = S_i$ (reversible) $S_f > S_i$ (irreversible)

2^{nd} Law of Thermodynamics and Life

If entropy had its way the whole universe would be a soup of equal density, temperature and composition.

Life is about the fight against the 2^{nd} law of thermodynamics (that is, entropy)!

It does this by absorbing, or eating, energy from the surrounding environment. The source of nearly all of Earth's energy is the sun; this keeps things going at the right pace. If Earth was too hot, we'd die, too cold and we'd freeze and also die. It appears that our version of life likes a certain temperature band.

But What Does It Mean To Be 'Alive'?

Early scientific perspective on the issue:

- Schrödinger speaks of life being characterized by and feeding on 'negative entropy' (*What Is Life?*, 1944)

Feasting on Negentropy

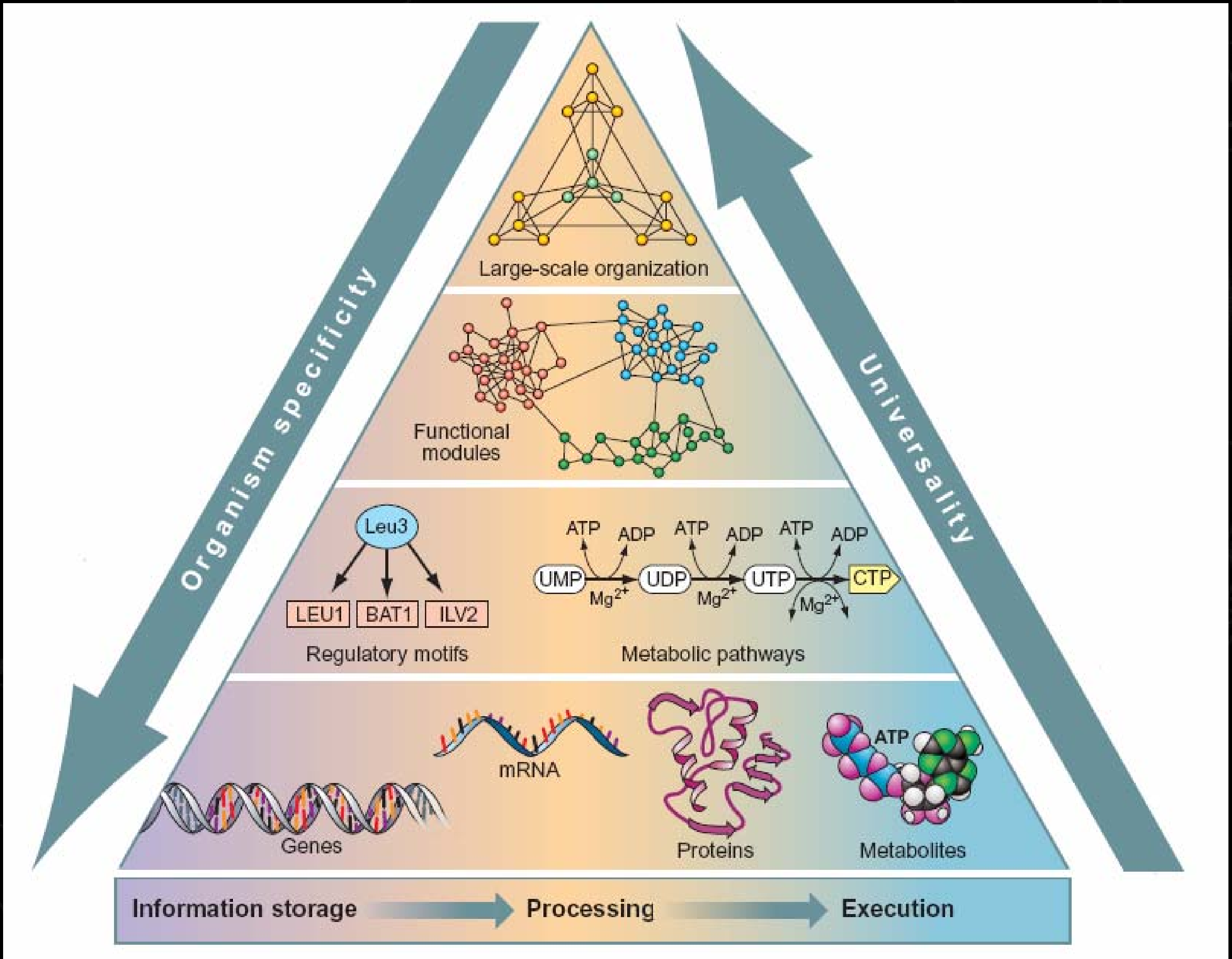
A meaningful interpretation of negative entropy ('negentropy') is that it measures the complexity of a physical structure in which quantities of energy are invested, e.g., buildings, technical devices, organisms but also atomic reactor fuel, the infrastructure of a society. In this sense organisms may be said to become more complex by feeding not on energy but on negentropy.

It's the difference of a maximally disordered system and the measured disorder of the system.

Information Flow and Pattern

- Von Neumann describes brain activity in terms of information flow (*The Computer and the Brain*, Silliman Lectures, 1958)
- Life is a (complex) pattern in space-time, rather than a specific material object *(Farmer and Belin, 1990)*

Organization: Life's Complexity Pyramid

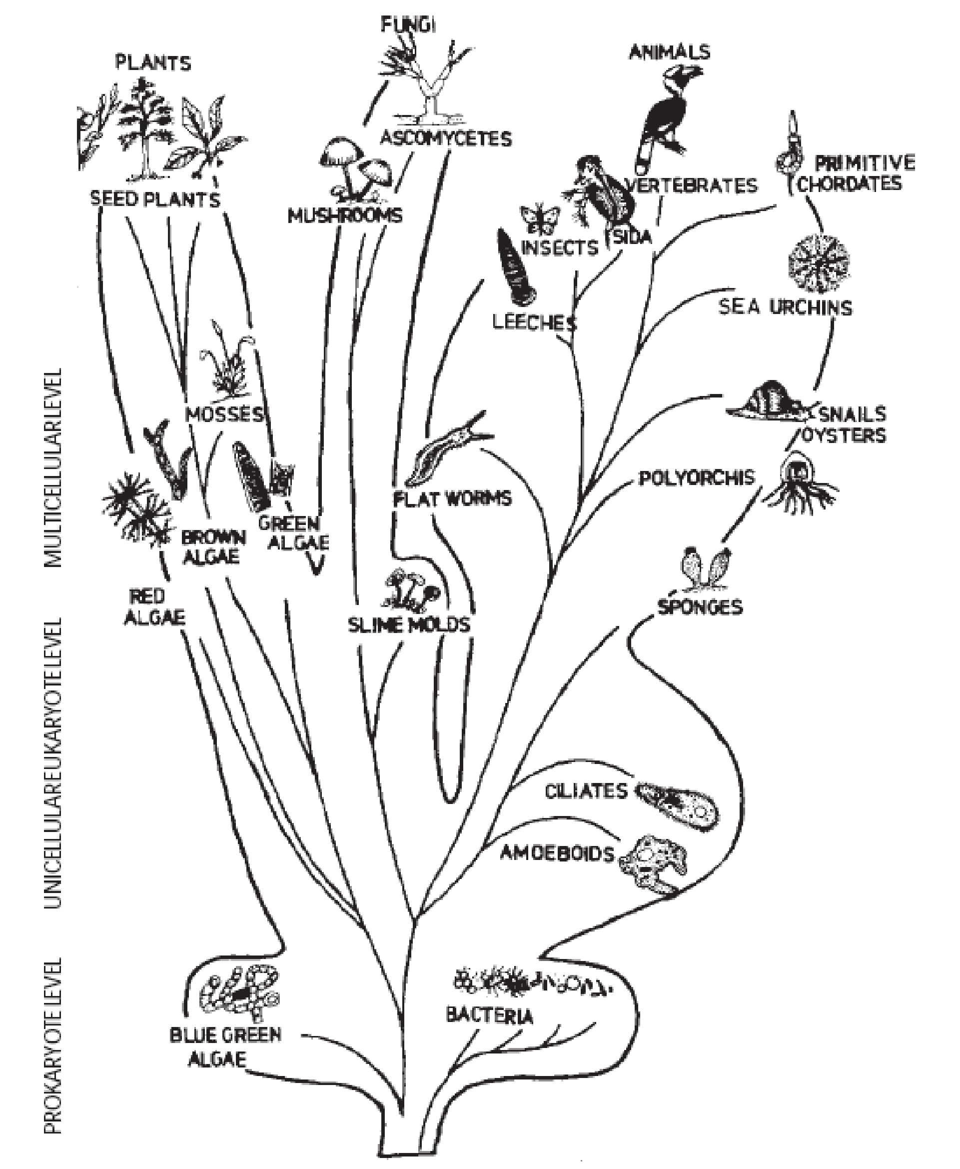


Oltvai and Barabasi, 2002

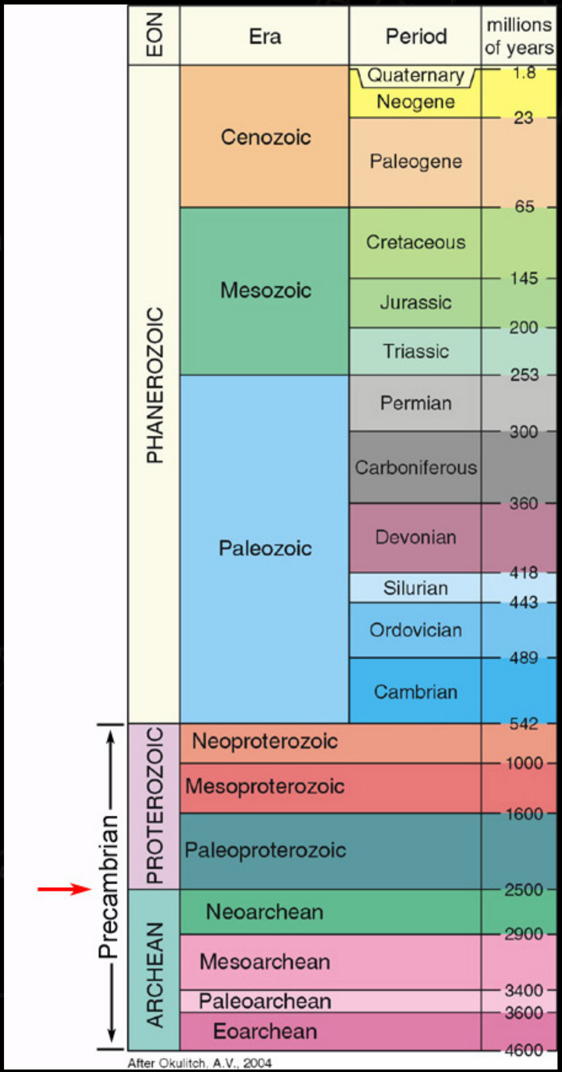
Does Complexity Measure Progress?

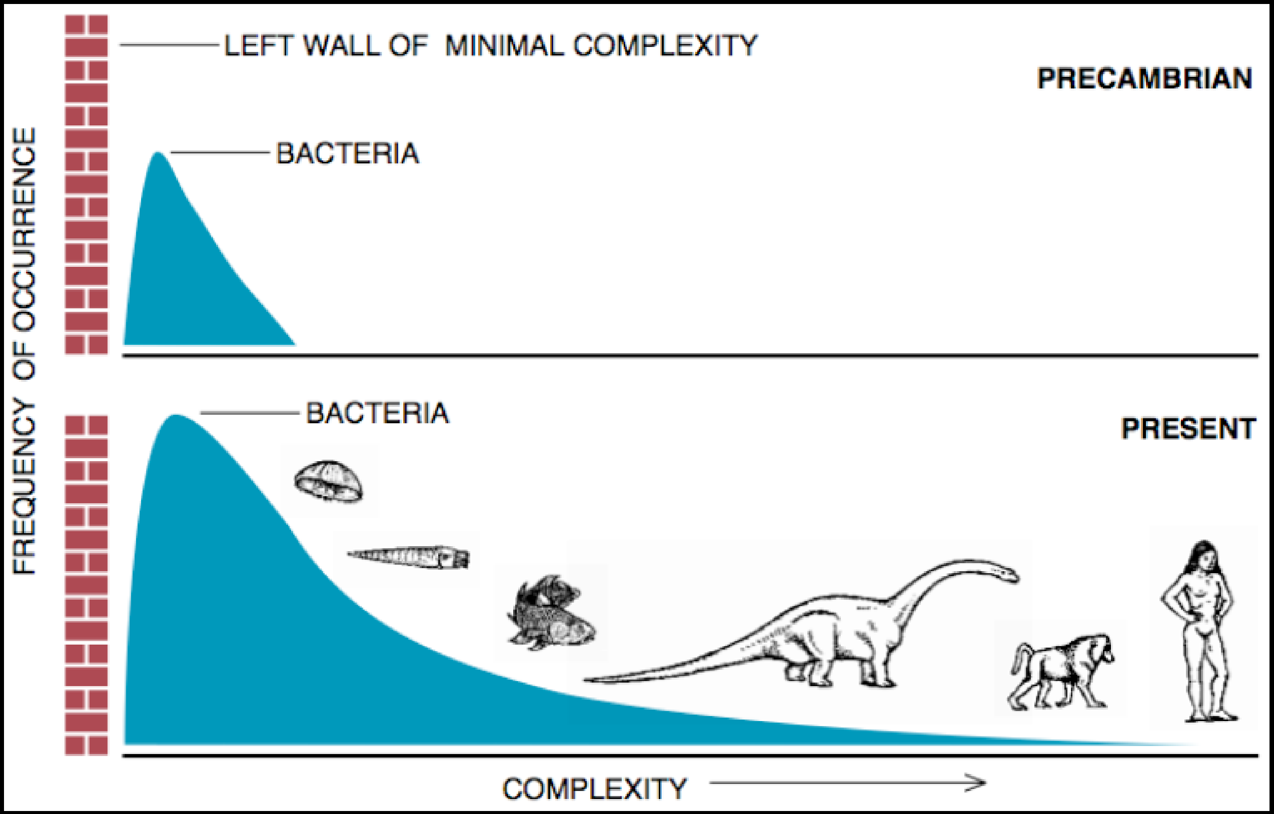
Spectrum of Life and Intelligence

From Simple Cells to Complex Cells



Evolutionary Trends in Complexity?





Evolutionary Trends in Complexity?

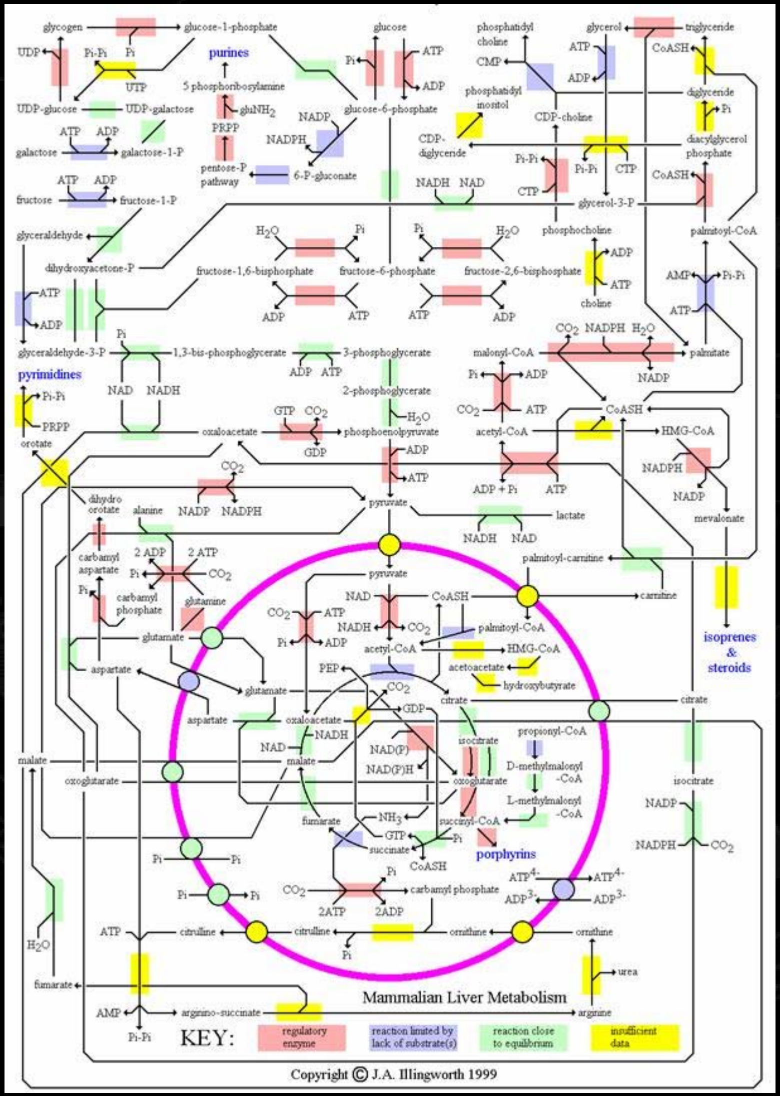
In a 1994 Scientific American article, Steven J. Gould famously argued against an evolutionary trend towards increasing complexity

However, he actually acknowledges the appearance of greater complexity over evolutionary time scales

Sources of Complexity Growth

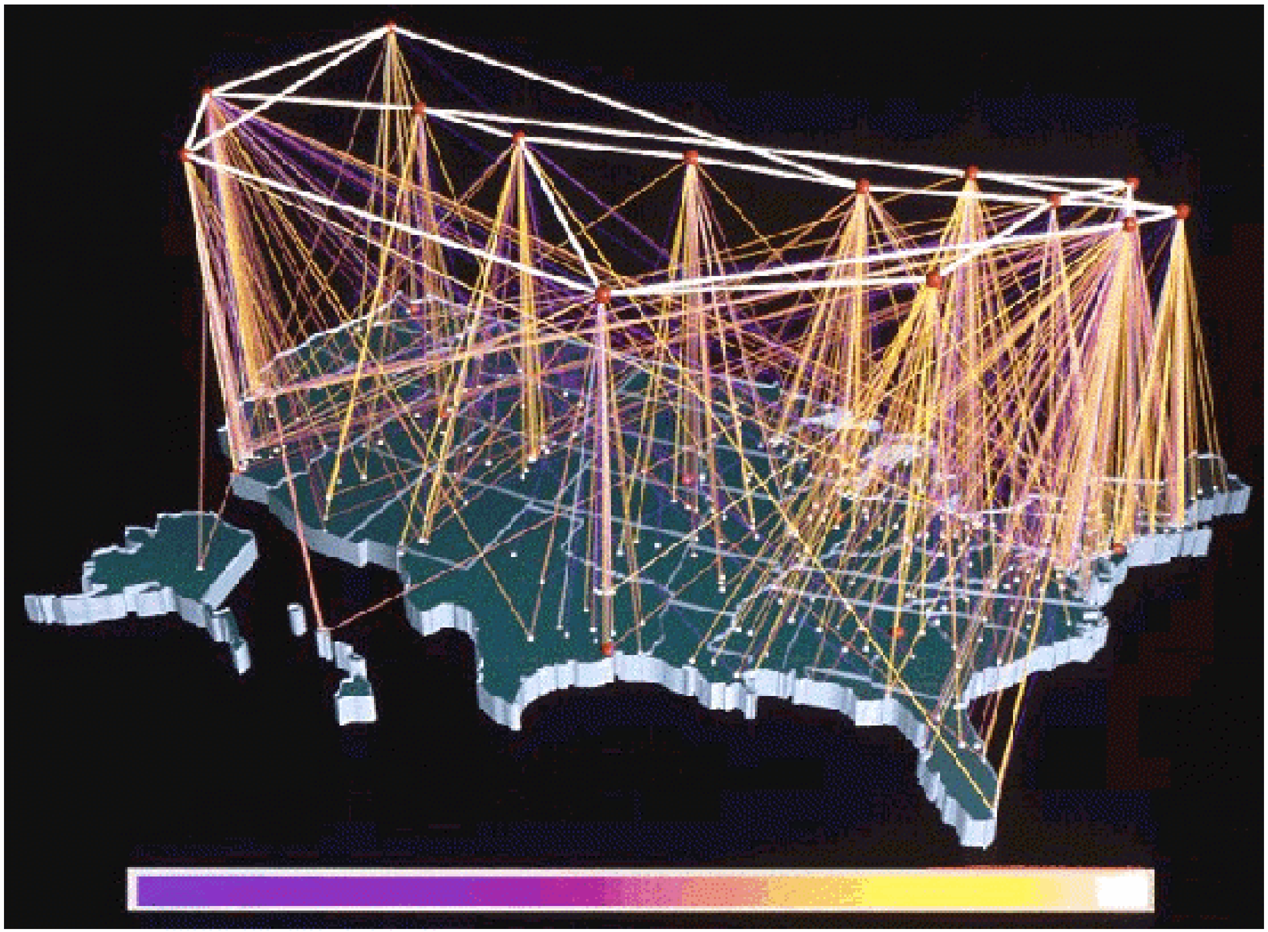
- Rensch (1960) and Bonner (1988) argued that more parts will allow a greater division of labor among parts
- Waddington (1969) and Arthur (1994) suggested that due to increasing diversity niches become more complex, and are then filled with more complex organisms
- Saunders and Ho (1976) claim component additions are more likely than deletions, because additions are less likely to disrupt normal function

Sources of Complexity Growth



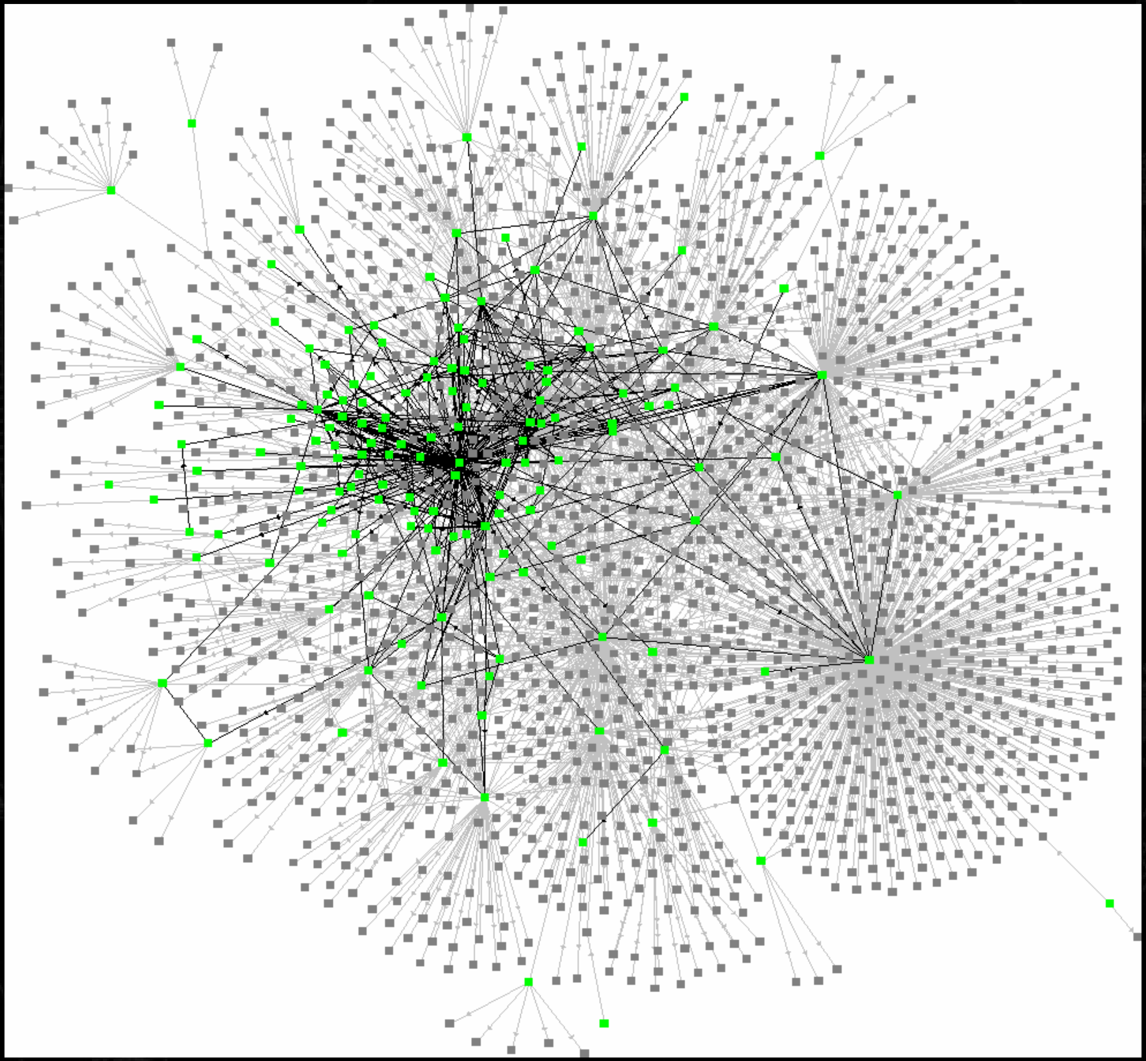
Liver metabolism network

Sources of Complexity Growth



The Fractal Properties of Growing Networks

Sources of Complexity Growth



Friendships of Blog Network members (http://radio.weblogs.com)

Complexity

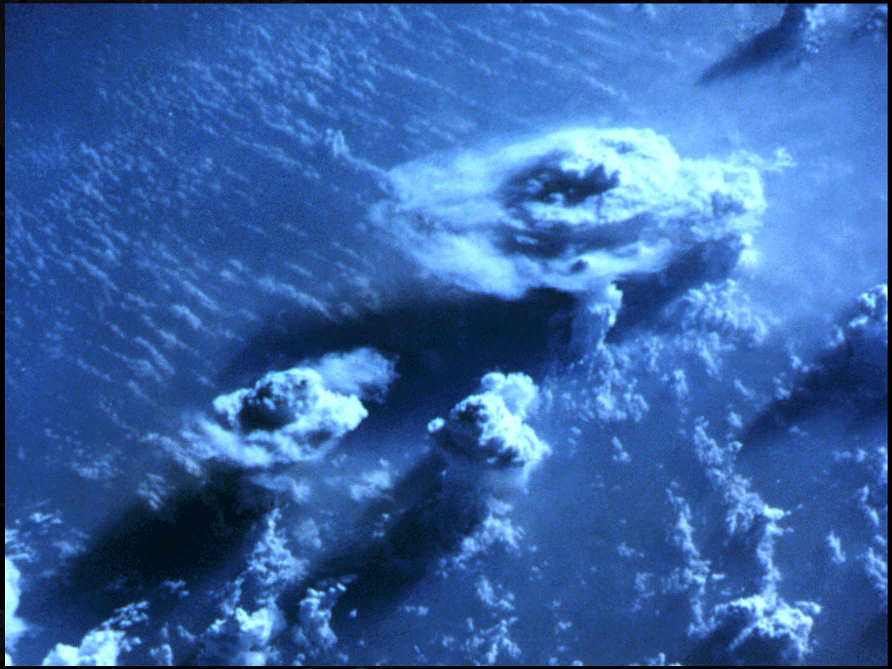
- Complexity seems a likely candidate for providing our quantitative assessment of life (and intelligence)
- But how do we actually measure it?
- So far no single measure of complexity has emerged as an uncontested winner

But What Does It Mean To Be ‘Complex’?

What is complexity?

Is there a way to quantify ‘complexity’?

What Is Complexity?



Most complicated wrist watch in the world made in 1983 by Gerald Genta (Patrick DesBiolles). It was sold in 1983 for 1.4 million dollars! Leave out one part and the whole thing doesn't work anymore

Complexity - Difficulty of Description

The complexity of a system is often associated with the degree of difficulty involved in <u>completely describing the system</u>. Prominent examples of these measures are:

- Information
- Entropy
- Algorithmic complexity
- Minimum description length
- Fisher information
- Renyi entropy
- Code length (prefix-free, Huffman, Shannon-Fano, error-correcting, Hamming)
- Chernoff information
- Dimension
- Fractal dimension
- Lempel-Ziv complexity

List compiled by Seth Lloyd

Complexity - Difficulty of Creation

Degree of difficulty involved in constructing or duplicating a system can form the basis of a complexity measure:

- Computational complexity
- Time computational complexity
- Space computational complexity
- Information-based complexity
- Logical depth
- Thermodynamic depth
- Cost
- Crypticity

List compiled by Seth Lloyd

Complexity - Degree of Organization

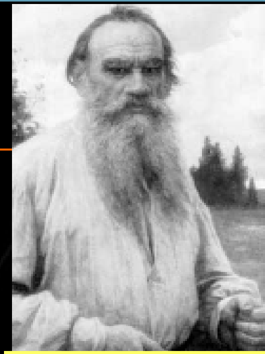
Difficulty of describing organizational structure/amount of information shared between the parts, whether corporate, chemical, cellular, etc

- Effective complexity
- Metric entropy
- Fractal dimension
- Excess entropy
- Stochastic complexity
- Sophistication
- Effective measure complexity
- True measure complexity
- Topological epsilon-machine size
- Conditional information
- Conditional algorithmic information content
- Schema length
- Ideal complexity
- Hierarchical complexity
- Tree subgraph diversity
- Homogeneous complexity
- Grammatical complexity
- Mutual information
- Channel capacity
- Stored information

Complexity - Yet Another Variation

Although this is not the whole story, one could say that:

Complexity emerges from the co-existence of tendencies to generate randomness and regularities



"All work and no play makes Jack a dull boy. All work and no play makes Jack a dull boy. All work and no play makes Jack a dull boy."

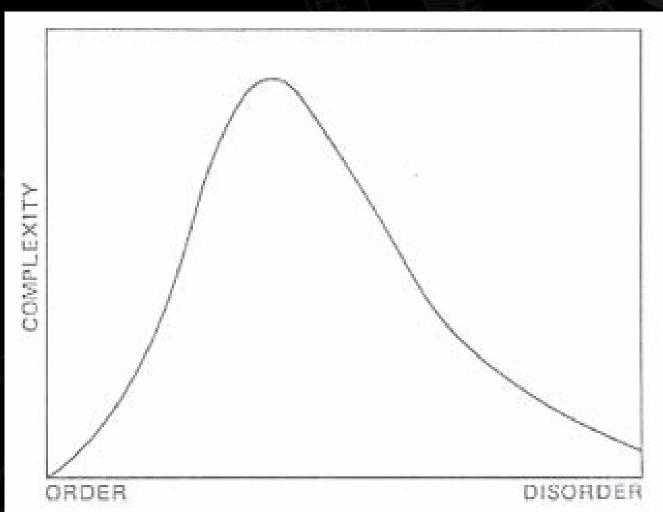
"Happy families are all alike; every unhappy family is unhappy in its own way."

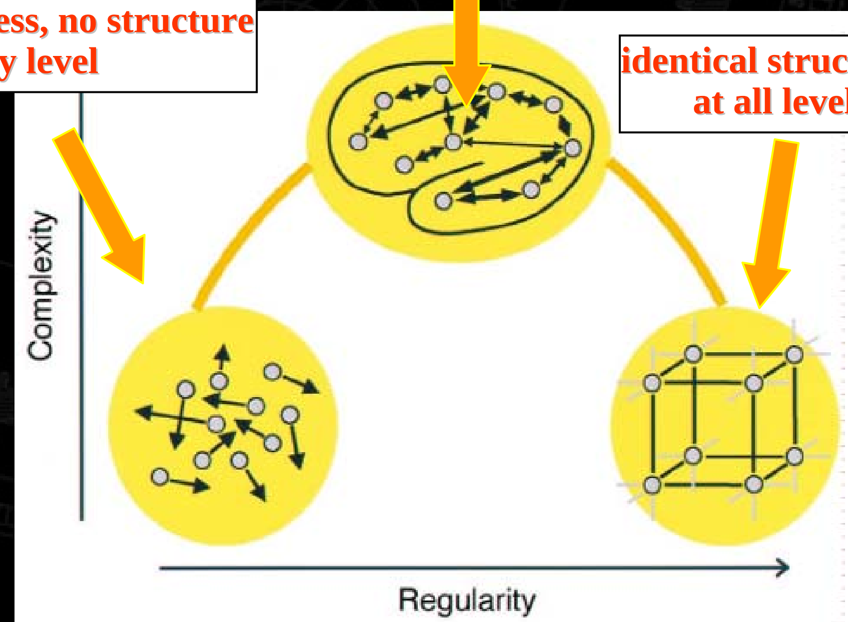
"What clashes here of wills gen wonts, oystrygods gaggin fishygods! Brekkek Kekkek Kekkek Kekkek! Koax Koax Koax! Ualu Ualu Ualu! Quaouahu!"

Non-repeating structure at multiple levels

randomness, no structure at any level

identical structure at all levels

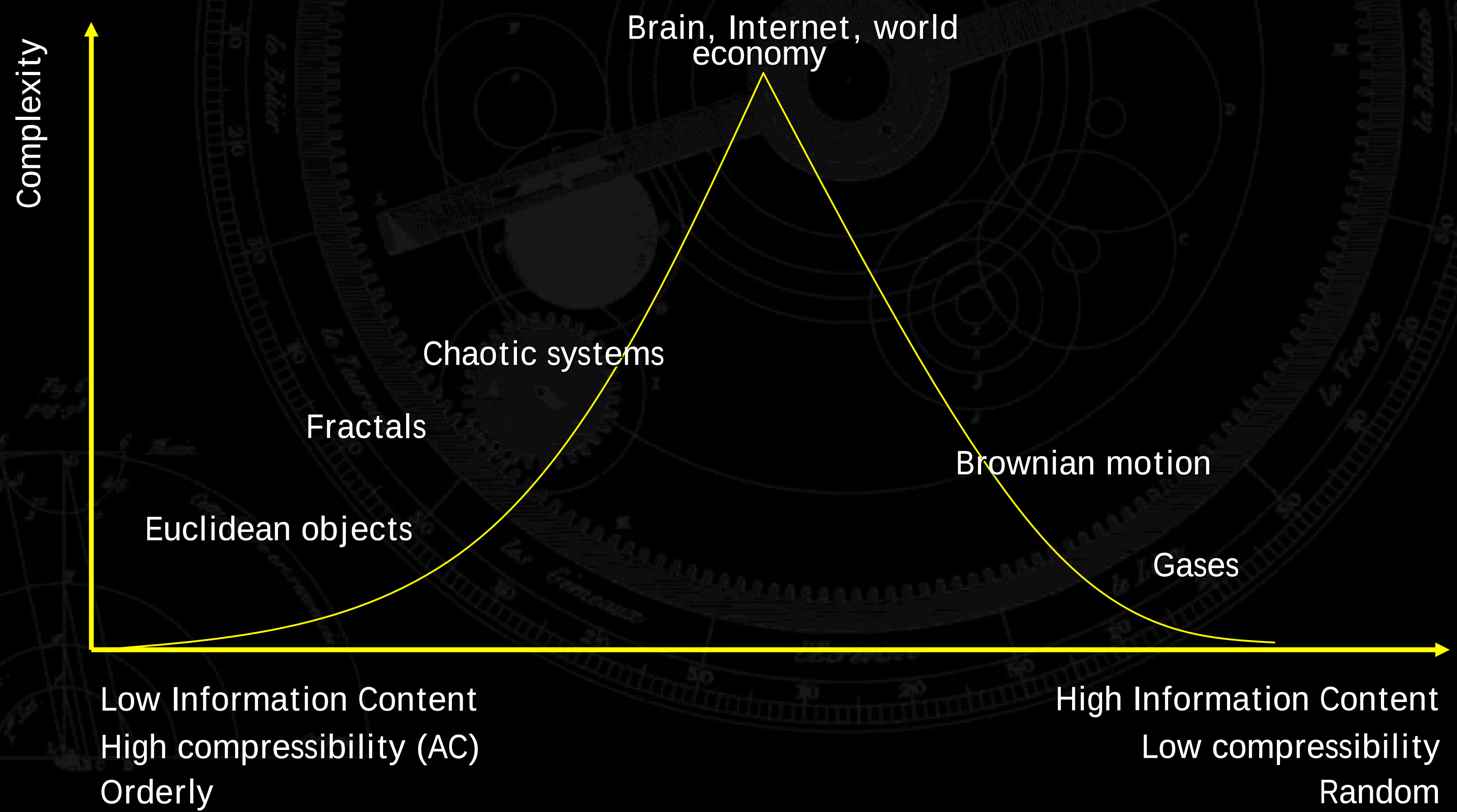




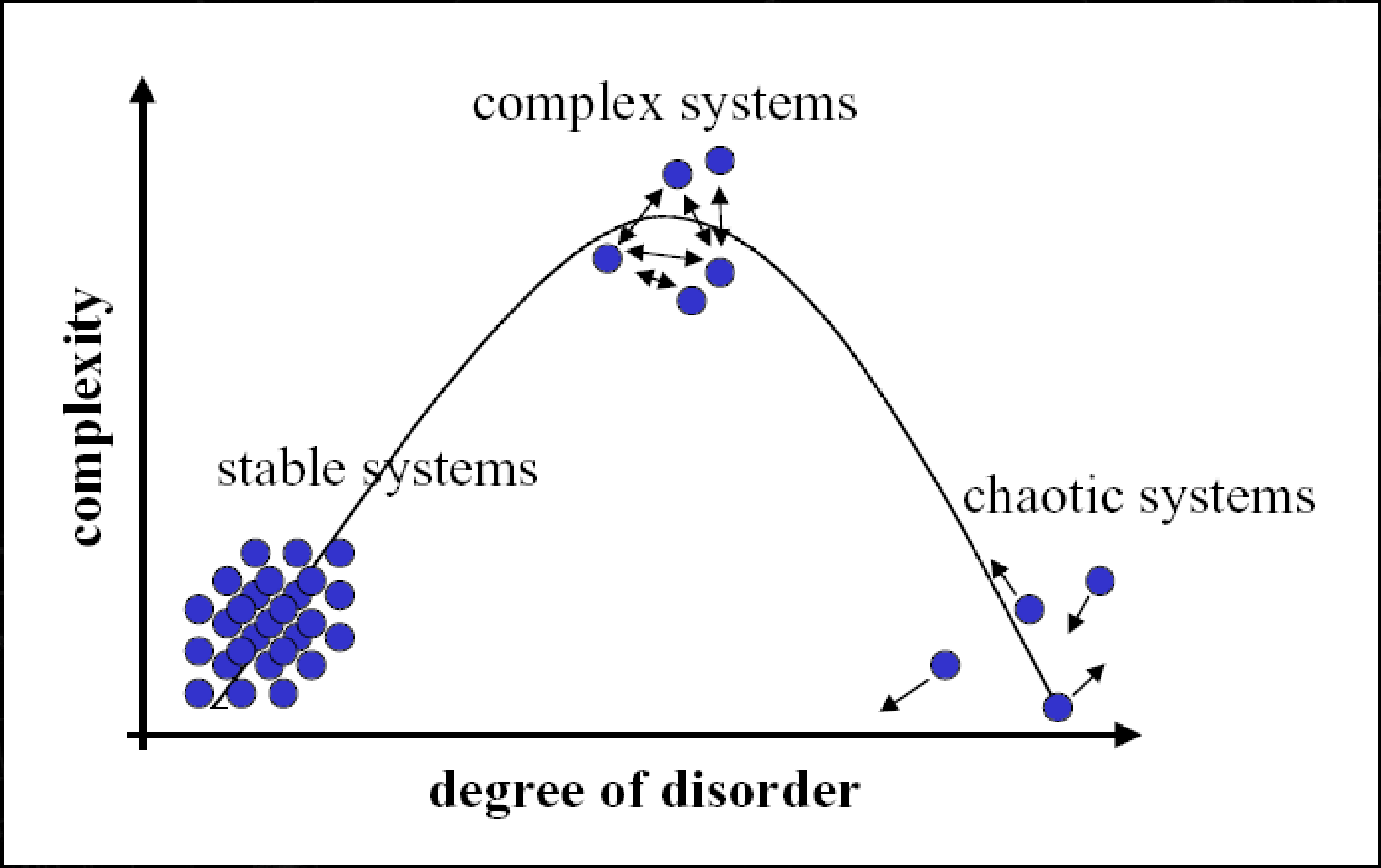
From Tononi, Sporns, Edelman, 1998

Effective Complexity

Complexity in terms of information, compressibility, and randomness



Where Do We Find Complexity?



Edge of Chaos (by Langton)

Dynamical systems must constantly balance a need for homeostasis with a need for dynamic variation

Life itself may have its origin in the kind of extended transients seen at the phase transition in dynamical systems, and that we may be "examples of the kind of 'computation' that can emerge in the vicinity of a phase transition given enough time"

Thus computation and life itself exist at, and because of, the edge of chaos

Characterization of Alife

From the script:

"Artificial life is based on ideas of <u>emergence</u> and <u>self-organization</u> in distributed system with <u>many elements</u> that interact with each by means of local rules."

How Do We Deal With Many Elements?

Statistical Physics

The object of statistical physics and thermodynamics is to describe the behavior of an aggregate, <u>knowing only the forces between the microscopic constituents</u>.

Statistical physics studies the laws that govern the behavior and properties of macroscopic bodies that are made up of <u>very many microscopic elements</u> (e.g. atoms, molecules, or any units)

Information Theory

Claude E. Shannon also called it "communication theory"

The theory was developed and published as "The Mathematical Theory of Communication" in the July and October 1948 issues of the *Bell System Technical Journal*

Shannon's concerns were clearly rooted in the communication of signals and symbols in a telephony system, but his formalization was so rigorous and general that it has since found many applications

He was aware of similarities and concerned about differences with thermodynamic entropy, but was encouraged to adopt the term by Von Neumann, who said, "Don't worry. No one knows what entropy is, so in a debate you will always have the advantage."

(from Larry Yaeger's lecture notes)

History

Samuel F.B. Morse worried about letter frequencies when designing (both versions of) the Morse code (1838)

- Made the most common letters use the shortest codes
- Obtained his estimate of letter frequency by counting the pieces of type in a printer's type box
- Observed transmission problems with buried cables

William Thompson, aka Lord Kelvin, Henri Poincaré, Oliver Heaviside, Michael Pupin, and G.A. Campbell all helped formalize the mathematics of signal transmission, based on the methods of Joseph Fourier (mid to late 1800's)

Harry Nyquist published the Nyquist Theorem in 1928

R.V.L. Hartley published "Transmission of Information" in 1928, containing a definition of information that is the same as Shannon's for equiprobable, independent symbols

More History

During WWII, A.N. Kolmogoroff, in Russia, and Norbert Weiner, in the U.S., devised formal analyses of the problem of extracting signals from noise (aircraft trajectories from noisy radar data)

In 1946 Dennis Gabor published “Theory of Communication”, which addressed related themes, but ignored noise

In 1948 Norbert Wiener published *Cybernetics*, dealing with communication and control

In 1948 Shannon published his work

In 1949 W.G. Tuller published “Theoretical Limits on the Rate of Transmission of Information” that parallels Shannon’s work on channel capacity

(from Larry Yaeger’s lecture notes)

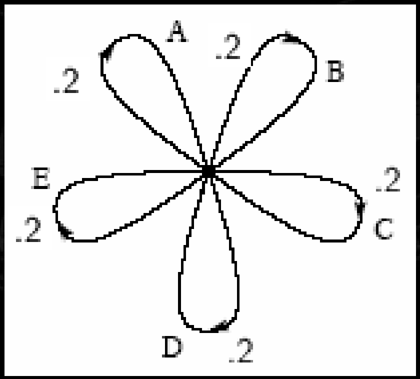
Stochastic Signal Sources

Suppose we have a set of 5 symbols—the English letters A, B, C, D, and E

If symbols from this set are chosen with equal (0.2) probability, you would get something like:

B D C B C E C C C A D C B D D A A E C E E A A B B D A E E C A C E E B A E E C B C E A D

This source may be represented as follows



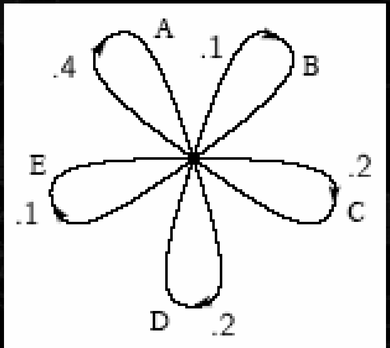
(from Larry Yaeger's lecture notes)

Stochastic Signal Sources

If the same symbols (A, B, C, D, E) are chosen with uneven probabilities 0.4, 0.1, 0.2, 0.2, 0.1, respectively, one obtains:

A A A C D C B D C E A A D A D A C E D A E A D C A B E D A D D C E C A A A A A D

This source may be represented as follows



(from Larry Yaeger's lecture notes)

Approximations in English

Assume we have a set of 27 symbols—the English alphabet plus a space

A zero-order model of the English language might then be an equiprobable, independent sequence of these symbols:

XFOML RXKHRJFFJUJ ZLPW CFW KCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZLHJQD

A first-order approximation, with independent symbols, but using letter frequencies of English text might yield:

OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL

(from Larry Yaeger's lecture notes)

Approximations in English

A second-order approximation using bigram probabilities from English text might yield:

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE

A third-order approximation using trigram probabilities from English text might yield:

IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE

(from Larry Yaeger's lecture notes)

Information in Markov Processes

The language models just discussed and many other symbol sources can be described as Markoff processes (stochastic processes in which future states depend solely on the current state, and not on how the current state was arrived at)

Can we define a quantity that measures the information produced by, or the information rate of, such a process?

Let's say that the information produced by a given symbol is exactly the amount by which we reduce our uncertainty about that symbol when we observe it

We therefore now seek a measure of uncertainty

(from Larry Yaeger's lecture notes)

Uncertainty

Suppose we have a set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$

Say these probabilities are known, but that is all we know concerning which event will occur next

What properties would a measure of our uncertainty, $H(p_1, p_2, \ldots, p_n)$, about the next symbol require:

1) H should be continuous in the p_i
2) If all the p_i are equal ($p_i = 1/n$), then H should be a monotonic increasing function of n
    - With equally likely events, there is more choice, or uncertainty, when there are more possible events
3) If a choice is broken down into two successive choices, the original H should be the weighted sum of the individual values of H

(from Larry Yaeger's lecture notes)

Entropy

In a proof that explicitly depends on this decomposibility and on monotonicity, Shannon establishes

Theorem 2: The only H satisfying the three above assumptions is of the form:

$$H = - K \Sigma pi \log pi$$

where K is a positive constant

Observing the similarity in form to entropy as defined in statistical mechanics, Shannon dubbed H the entropy of the set of probabilities $p_1, p_2, \ldots, p_n$

Generally, the constant K is dropped; Shannon explains it merely amounts to a choice of unit of measure

(from Larry Yaeger's lecture notes)

Behavior of Entropy Function

In general, $H = 0$ if and only if all the p_i are zero, except one which has a value of one

For a given n, H is a maximum (and equal to log n) when all p_i are equal (1/n)

- Intuitively, this is the most uncertain situation

Any change toward equalization of the probabilities $p_1, p_2, \ldots, p_n$ increases H

- If $p_i \neq p_j$, adjusting p_i and p_j so they are more nearly equal increases H
- Any "averaging" operation on the p_i increases H

(from Larry Yaeger's lecture notes)

Joint Entropy

For two events, x and y, with m possible states for x and n possible states for y, the entropy of the joint event may be written in terms of the joint probabilities

while

$$H(x,y) = - \Sigma\, p(x_i,y_j) \log p(x_i,y_j)$$

$$H(x) = - \Sigma\, p(x_i,y_j) \log \Sigma\, p(x_i,y_j)$$

$$H(y) = - \Sigma\, p(x_i,y_j) \log \Sigma\, p(x_i,y_j)$$

It is "easily" shown that

$$H(x,y) \leq H(x) + H(y)$$

- Only equal if the events are independent
  - $p(x,y) = p(x)\, p(y)$
- Uncertainty of a joint event is less than or equal to the sum of the individual uncertainties

(from Larry Yaeger's lecture notes)

Maximum and Normalized Entropy

Maximum entropy, when all probabilities are equal is

$$H_{Max} = \log n$$

Normalized entropy is the ratio of entropy to maximum entropy

$$H_o(x) = H(x) \ / \ H_{Max}$$

Since entropy varies with the number of states, n, normalized entropy is a better way of comparing across systems

Shannon called this *relative entropy*

(Some cardiologists and physiologists call entropy divided by total signal power normalized entropy)

(from Larry Yaeger's lecture notes)

Mutual Information

Define *Mutual Information* (aka *Shannon Information Rate*) as

$$I(x,y) = - \Sigma\, p(x_i,y_j) \log [\, p(x_i,y_j) / p(x_i)p(y_j)\,]$$

When x and y are independent $p(x_i,y_j) = p(x_i)p(y_j)$, so $I(x,y)$ is zero

When x and y are the same, the mutual information of x,y is the same as the information conveyed by x (or y) alone, which is just $H(x)$

Mutual information can also be expressed as

$$I(x,y) = H(x) - H(x|y) = H(y) - H(y|x)$$

Mutual information is nonnegative

Mutual information is symmetric; i.e., $I(x,y) = I(y,x)$

(from Larry Yaeger's lecture notes)