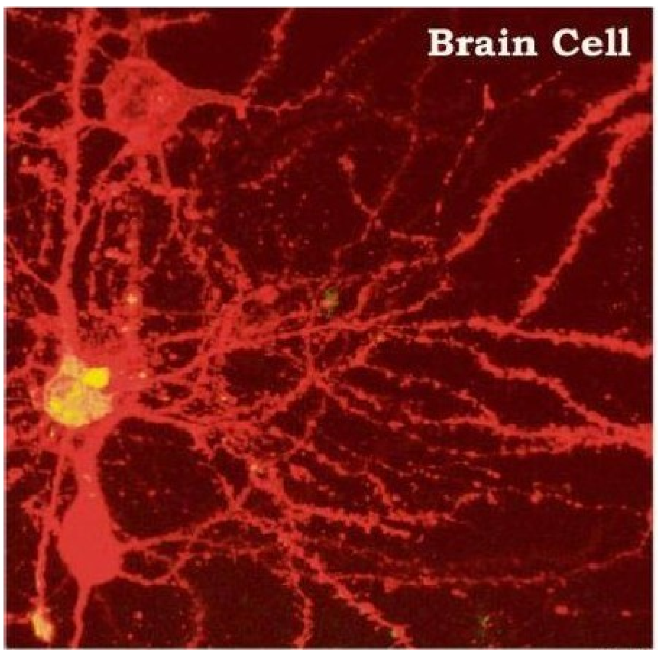
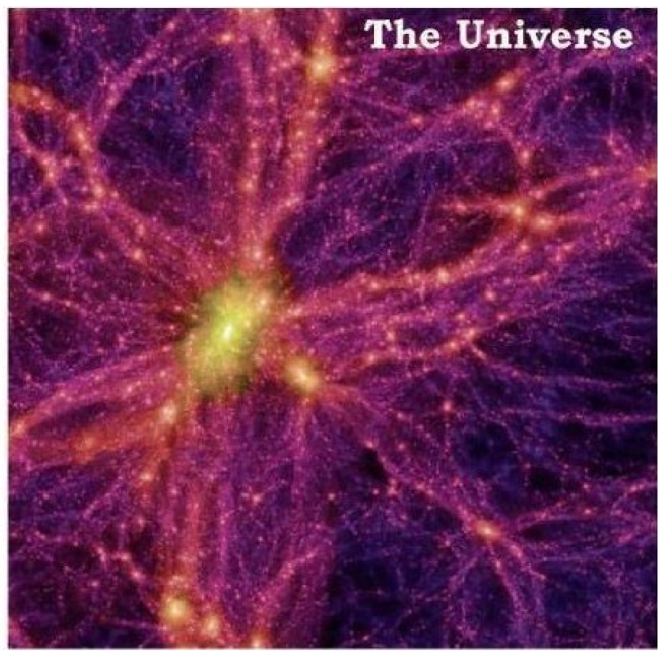
NEURAL NETWORK NATURE

Fractal Hierarchies of 'Perceptrons' from Clusters of galaxies to the World Wide Web

a small handbook

Peter Winiwarter





e-mail : winiwarter@bordalierinstitute.com

Table of Content

"*Things arise in Space as Thoughts arise in Mind"*

Parmenides

"The Universe is a vast system of systems which strikingly resemble one another in the details of their structures and processes. Among theses systems, or realms, are matter, life and mind"

George Perrigo Conger in *A World of Epitomizations*

"Life and mind have a common abstract pattern or set of basic organizational properties. The functional properties characteristic of mind are an enriched version of the functional properties that are fundamental to life in general. Mind is literally life-like. "

Godfrey-Smith, P. (1996). Complexity and the Function of Mind in Nature. Cambridge: Cambridge University Press.

Introduction

Every creation in the field of science or art is the realization of a child's or juvenile's dream.

I received my high school education at the "humanistic gymnasium" at Linz, Austria. Since the age of nine our mind was formed with Latin lessons six days a week during every school year. Daily lessons in ancient Greek were added at the age of eleven. The major goal of this education was to form our minds in the old tradition of Greek-roman culture without neglecting mathematics and philosophy.

At the age of fourteen we had to choose between two additional subjects: music or art. I played the violin and should have been attracted by music, but I had not the least ear – tuning my instrument was a daily nightmare. I liked to draw and to paint so I chose Art as my additional subject for the remaining years of my education.

At the final exam of graduation, which in German speaking countries is called Matura or Abitur, we had three compulsory subjects: Latin, Greek and Mathematics and one subject of our choice. I was always a fan of the "principle of least effort" so I choose Art as fourth subject. For the exam we were questioned on the history of Art but we also had to produce one work of art corresponding to a given topic.

The topic of the exam was "Big fish eat small fish" to be realized as a painting in four hours.

I liked the subject and at the end of the four hours I admired my creation. The biggest part of my painting was filled with a big monster fish, who was about to swallow a medium sized fish. The medium sized fish was about to swallow a small fish. On the left side of the painting two more medium sized fishes were busy swallowing small fishes and the rest of the painting was filled with all sorts of small fishes swimming around. For an animated version see http://www.funny-games.biz/fishtales.html

Fish Tales

Meet Sunny, a small fish in a vast ocean. Use YOUR MOUSE to help Sunny survive in these dangerous waters. To win you have to follow these rules. Eat the fish smaller than yourself, avoid the fish bigger than yourself and eat enough fish to grow up. Have fun!

This was just a juvenile's dream.

Today, waking up during my years of research I often had this image in mind “big fish eat small fish”.

This book is a scientific answer to the question. From a science point of view the painting is called an aquatic ecosystem showing the food web for the biggest organisms fish. There is a strict hierarchical order in the systems with the constraints “who swallows whom” which follows a power law. A few hubs, the biggest fishes, swallow almost everything, while a few medium sized fish modules are constrained on the feeding of a great number of small and very small fishes. There is also a fractal like feature in the image; independent of the scaling we see the same building block a big fish swallowing a small fish.

Todays studies of ecosystems go further down in the scaling hierarchy to plankton and bacteria over more than ten orders of magnitude in size. What is noteworthy that virtually all observed ecosystems reveal power law biomass size distributions.

Why do we observe these Pareto-Zipf-Mandelbrot (PZM) regularities not only in ecosystems but also for complex networks on virtually every level of the evolutionary hierarchy from stars to the World Wide Web?

This book has a simple aim: to get you to think “real world” complex networks in terms of Neural Nets, that have memory, are learning and could be considered as intelligent, since they strive to reach a goal.

The intelligence is not only located in brains, its located out there in the topology and weighted links of the numerous small world networks ranging from massive stars to the World Wide Web.

Hierarchy Theory

To grasp the key idea put forward in this book, that the universe can be understood as a self similar hierarchy of neural networks some basic concepts like hierarchy, self-similarity, fractal, network and neural network have to be understood by the reader.

When we found a suitable Wikipedia entry we have cited the article in extenso to avoid for the reader the necessity to be on line to the Internet while reading the book. When the reader desires to deepen his understanding he can follow the links in the text when connected on line.

Likewise several sections like the one on operator hierarchy and compositional vs. subsumption hierarchy have been written by the authors and been included in the book with their permission. Why rewrite when the authors or an encyclopedia can say it better?

What is a Hierarchy? Wikipedia

A **hierarchy** is an arrangement of objects, people, elements, values, grades, orders, classes, etc., in a ranked or graduated series. The word derives from the Greek **ἱεραρχία** (*hierarchia*), from *ἱεράρχης* (*hierarches*), "president of sacred rites, high-priest" and that from *ἱερός* (*hieros*), "sacred" + *ἄρχω* (*arkho*), "to lead, to rule"[1][2]. The word can also refer to a series of such items so arranged. Items in a hierarchy are typically thought of as being "above," "below," or "at the same level as" one another.[3][4]

This is as opposed to anarchy where there is no concept of higher or lower items (or people) -- everything is considered equal.

The first use of the word "hierarchy" cited by the Oxford English Dictionary was in 1880, when it was used in reference to the three orders of three angels as depicted by Pseudo-Dionysius the Areopagite. Pseudo-Dionysius used the word both in reference to the celestial hierarchy and the ecclesiastical hierarchy. [5] His term is derived from the Greek for 'Bishop' (hierarch), and Dionysius is credited with first use of it as an abstract noun. Since hierarchical churches, such as the Roman Catholic and Eastern Orthodox churches, had tables of organization that were "hierarchical" in the modern sense of the word (traditionally with God as the pinnacle of the hierarchy), the term came to refer to similar organizational methods in more general settings.

A hierarchy can link entities either directly or indirectly, and either vertically or horizontally. The only direct links in a hierarchy, insofar as they are hierarchical, are to one's immediate superior or to one of one's subordinates, although a system that is largely hierarchical can also incorporate other organizational patterns. Indirect hierarchical links can extend "vertically" upwards or downwards via multiple links in the same direction. All parts of the hierarchy which are not vertically linked to one another can nevertheless be "horizontally" linked by traveling up the hierarchy to find a common direct or indirect superior, and then down again. This is akin to two co-workers, neither of whom is the other's boss, but both of whose chains of command will eventually meet.

These relationships can be formalized mathematically; see hierarchy (mathematics).

Computation and electronics

Large electronic devices such as computers are usually composed of modules, which are themselves created out of smaller components (integrated circuits), which in turn are internally organized using hierarchical methods (e.g. using standard cells). The order of tasks in a computational algorithm is often managed hierarchically, with repeated loops nested within one another. Computer files in a file system are stored in an hierarchy of directories in most operating systems. In object-oriented programming, classes are organized hierarchically; the relationship between two related classes is called inheritance. In the Internet, IP addresses are increasingly organized in an hierarchy (so that the routing will continue to function as the Internet grows).

Computer graphic imaging (CGI)

Within most CGI and computer animation programs is the use of hierarchies. On a 3D model of a human, the chest is a parent of the upper left arm, which is a parent of the lower left arm, which is a parent of the hand. This is used in modeling and animation of almost everything built as a 3D digital model.

Biological taxonomy

In biology, the study of taxonomy is one of the most conventionally hierarchical kinds of knowledge, placing all living beings in a nested structure of divisions related to their probable evolutionary descent. Most evolutionary biologists assert a hierarchy extending from the level of the specimen (an individual living organism — say, a single newt), to the species of which it is a member (perhaps the Eastern Newt), outward to further successive levels of genus, family, order, class, phylum, and kingdom. (A newt is a kind of salamander (family), and all salamanders are types of amphibians (class), which are all types of vertebrates (phylum).) Essential to this kind of reasoning is the proof that members of a division on one level are more closely related to one another than to members of a different division on the same level; they must also share ancestry in the level above. Thus, the system is hierarchical because it forbids the possibility of overlapping categories. For example, it will not permit a 'family' of beings containing some examples that are amphibians and others that are reptiles — divisions on any level do not straddle the categories of structure that are hierarchically above it. (Such straddling would be an example of heterarchy.)

Organisms are also commonly described as assemblies of parts (organs) which are themselves assemblies of yet smaller parts. When we observe that the relationship of cell to organ is like that of the relationship of organ to body, we are invoking the hierarchical aspects of physiology. (The term "organic" is often used to describe a sense of the small imitating the large, which suggests hierarchy, but isn't necessarily hierarchical.) The analogy of organ to body also extends to the relationship of a living being as a system that might resemble an ecosystem consisting of several living beings; physiology is thus hierarchically nested in ecology.

Physics

In physics, the standard model of reasoning on the nature of the physical world decomposes large bodies down to their smallest particle components. Observations on the subatomic (particle) level are often seen as fundamental constituent axioms, on which conclusions about the atomic and molecular levels depend. The relationships of energy and gravity between celestial bodies are, in turn, dependent upon the atomic and molecular properties of smaller bodies. In energetics, energy quality is sometimes used to quantify energy hierarchy.

Language and semiotics

In linguistics, especially in the work of Noam Chomsky, and of later generative linguistics theories, such as Ray Jackendoff's, words or sentences are often broken down into hierarchies of parts and wholes. Hierarchical reasoning about the underlying structure of language expressions leads some linguists to the hypothesis that the world's languages are bound together in a broad array of variants subordinate to a single Universal Grammar.

Hierarchical verbal alignment

In some languages, such as Cree and Mapudungun, subject and object on verbs are distinguished not by different subject and object markers, but via a hierarchy of persons.

In this system, the three (or four with Algonquian languages) persons are placed in a hierarchy of salience. To distinguish which is subject and which object, *inverse markers* are used if the object outranks the subject.

In music, the structure of a composition is often understood hierarchically (for example by Heinrich Schenker (1768–1835, see Schenkerian analysis), and in the (1985) Generative Theory of Tonal Music, by composer Fred Lerdahl and linguist Ray Jackendoff). The sum of all notes in a piece is understood to be an all-inclusive surface, which can be reduced to successively more sparse and more fundamental types of motion. The levels of structure that operate in Schenker's theory are the foreground, which is seen in all the details of the musical score; the middle ground, which is roughly a summary of an essential contrapuntal progression and voice-leading; and the background or Ursatz, which is one of only a few basic "long-range counterpoint" structures that are shared in the gamut of tonal music literature.

The pitches and form of tonal music are organized hierarchically, all pitches deriving their importance from their relationship to a tonic key, and secondary themes in other keys are brought back to the tonic in a recapitulation of the primary theme. Susan McClary connects this specifically in the sonata-allegro form to the feminist hierarchy of gender (see above) in her book *Feminine Endings*, even pointing out that primary themes were often previously called "masculine" and secondary themes "feminine."

Hierarchies in programming

The concept of hierarchies plays a large part in object oriented programming. For more information see Hierarchy (object-oriented programming) and memory hierarchy.

Containment hierarchy

A containment hierarchy of the subsumption kind is a collection of strictly nested sets. Each entry in the hierarchy designates a set such that the previous entry is a strict superset, and the next entry is a strict subset. For example, all rectangles are quadrilaterals, but not all quadrilaterals are rectangles, and all squares are rectangles, but not all rectangles are squares. (See also: Taxonomy.) A containment hierarchy of the compositional kind refers to parts and wholes, as well as to rates of change. Generally the bigger changes more slowly. Parts are contained in wholes and change more rapidly than do wholes.

- In geometry: {shape {polygon {quadrilateral {rectangle {Square (geometry)|square }}}}}
- In biology:subsumption hierarchy {animal {bird {bird of prey|raptor {eagle {golden eagle}}}}}
    - compositional hierarchy: [population [organism [biological cell [macromolecule]]]]
- The Chomsky hierarchy in formal languages: recursively enumerable, context-sensitive, context-free, and regular
- In physics: subsumption hierarchy {elementary particle {fermion {lepton {electron }}}}
    - compositional hierarchy: [galaxy [star system [star]]]

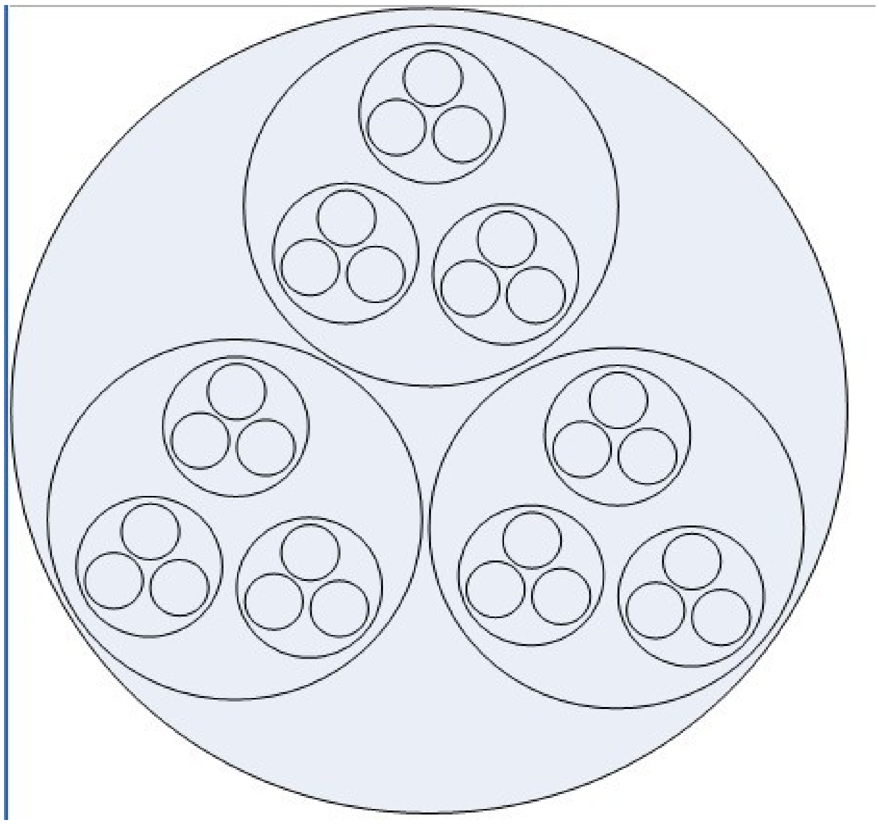
Social hierarchies

Many human organizations, such as governments, educational institutions, businesses, churches, armies and political movements are hierarchical organizations, at least officially; commonly seniors, called "bosses", have more power. Thus the relationship defining this hierarchy is "commands" or "has power over". Some analysts question whether power "actually" works in the way the traditional organizational chart indicates, however. This view tends to emphasize the significance of the informal organization. See also chain of command.

Retrieved from "http://en.wikipedia.org/wiki/Hierarchy"

Hierarchy of Holons (1968 Koestler)

Some 40 years ago, Arthur Koestler proposed the word "holon" [Koestler 1968]. It is a combination from the Greek 'holos' = whole, with the suffix 'on' which, as in proton or neutron, suggests a particle or part.



Selfsimilar hierarchy of holons

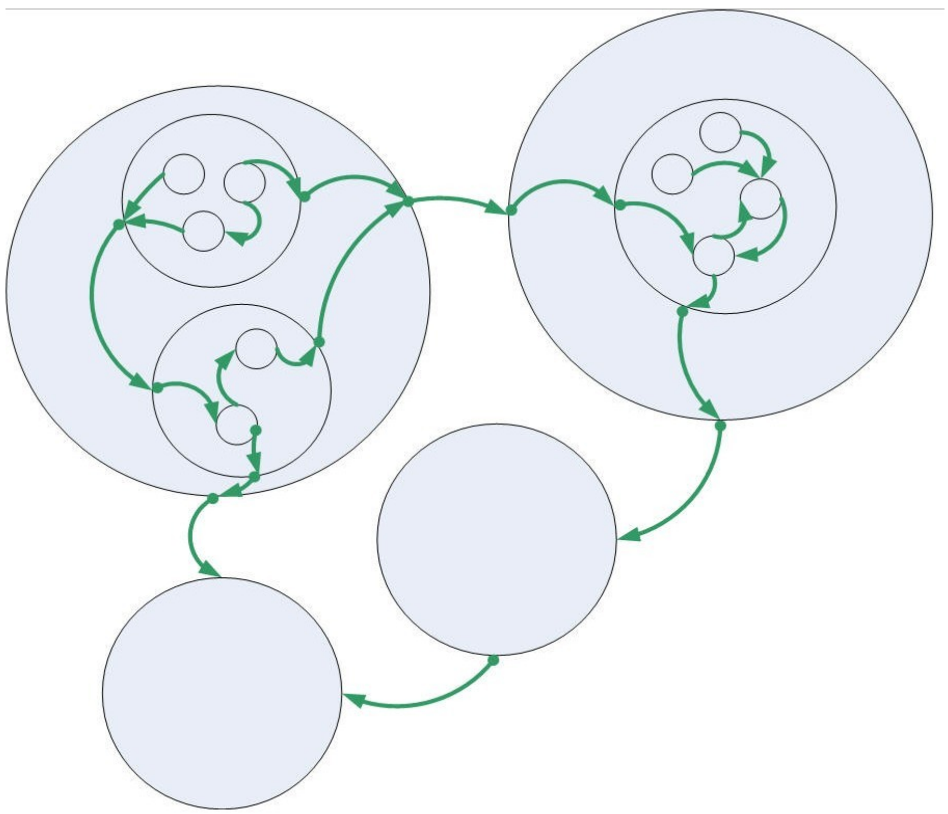
Two observations impelled Koestler to propose the word holon. The first comes from Herbert Simon, a Nobel prize winner, and is based on his 'parable of the two watchmakers'.

The Parable

There once were two watchmakers, named Hora and Tempus, who made very fine watches. The phones in their workshops rang frequently and new customers were constantly calling them. However, Hora *prospered while Tempus became poorer and poorer. In the end, Tempus lost his shop. What was the reason behind this?*

The watches consisted of about 1000 parts each. The watches that Tempus made were designed such that, when he had to put down a partly assembled watch, it immediately fell into pieces and had to be reassembled from the basic elements. Hora had designed his watches so that he could put together sub-assemblies of about ten components each, and each sub-assembly could be put down without falling apart. Ten of these subassemblies could be put together to make a larger sub-assembly, and ten of the larger sub-assemblies constituted the whole watch.

From this parable, Simon concludes that complex systems will evolve from simple systems much more rapidly if there are stable intermediate forms than if there are not; the resulting complex systems in the former case will be hierarchic.



Dynamics of of a holarchy

The second observation, made by Koestler while analyzing hierarchies and stable intermediate forms in living organisms and social organization, is that although it is easy to identify sub-wholes or parts 'wholes' and 'parts' in an absolute sense do not exist anywhere. This made Koestler propose the word holon to describe the hybrid nature of sub- wholes/parts in real-life systems; holons simultaneously are self-contained wholes to their subordinated parts, and dependent parts when seen from the inverse direction.

Koestler also establishes the link between holons and the watchmakers' parable from professor Simon. He points out that the sub-wholes/holons are autonomous self-reliant units, which have a degree of independence and handle contingencies without asking higher authorities for instructions. Simultaneously, holons are subject to control from (multiple) higher authorities. The first property ensures that holons are stable forms, which survive disturbances. The latter property signifies that they are intermediate forms, which provide the proper functionality for the bigger whole.

Finally, Koestler defines a holarchy as a hierarchy of self-regulating holons which function (a) as autonomous wholes in supra-ordination to their parts, (b) as dependent parts in sub- ordination to controls on higher levels, (c) in co-ordination with their local environment

What is a Holon? Wikipedia

General definition

A holon is a system (or phenomenon) that is a whole in itself as well as a part of a larger system. It can be conceived as systems nested within each other. Every system can be considered a holon, from a subatomic particle to the universe as a whole. On a non-physical level, words, ideas, sounds, emotions—everything that can be identified—is simultaneously part of something, and can be viewed as having parts of its own, similar to sign in regard of semiotics.

Since a holon is embedded in larger wholes, it is influenced by and influences these larger wholes. And since a holon also contains subsystems, or parts, it is similarly influenced by and influences these parts. Information flows bidirectionally between smaller and larger systems as well as rhizomatic contagion. When this bidirectionality of information flow and understanding of role is compromised, for whatever reason, the system begins to break down: wholes no longer recognize their dependence on their subsidiary parts, and parts no longer recognize the organizing authority of the wholes. Cancer may be understood as such a breakdown in the biological realm.

A hierarchy of holons is called a holarchy. The holarchic model can be seen as an attempt to modify and modernise perceptions of natural hierarchy.

Ken Wilber comments that the test of holon hierarchy (e.g. holarchy) is that if a type of holon is removed from existence, then all other holons of which it formed a part must necessarily cease to exist too. Thus an atom is of a lower standing in the hierarchy than a molecule, because if you removed all molecules, atoms could still exist, whereas if you removed all atoms, molecules, in a strict sense would cease to exist. Wilber's concept is known as the doctrine of the **fundamental** and the **significant**. A hydrogen atom is more fundamental than an ant, but an ant is more significant.

The doctrine of the fundamental and the significant are contrasted by the radical rhizome oriented pragmatics of Deleuze and Guattari, and other continental philosophy.

Types of holons

Individual holon

An individual holon possesses a dominant monad; that is, it possesses a definable "I-ness". An individual holon is discrete, self-contained, and also demonstrates the quality of agency, or self-directed behavior. [3] The individual holon, although a discrete and self-contained is made up of parts; in the case of a human, examples of these parts would include the heart, lungs, liver, brain, spleen, etc. When a human exercises agency, taking a step to the left, for example, the entire holon, including the constituent parts, moves together as one unit.

Social holon

A social holon does not possess a dominant monad; it possesses only a definable "we-ness", as it is a collective made up of individual holons. [4] In addition, rather than possessing discrete agency, a social holon possesses what is defined as nexus agency. An illustration of nexus agency is best described by a flock of geese. Each goose is an individual holon, the flock makes up a social holon. Although the flock moves as one unit when flying, and it is "directed" by the choices of the lead goose, the flock itself is not mandated to follow that lead goose. Another way to consider this would be collective activity that has the potential for independent internal activity at any given moment.

Applications

Ecology

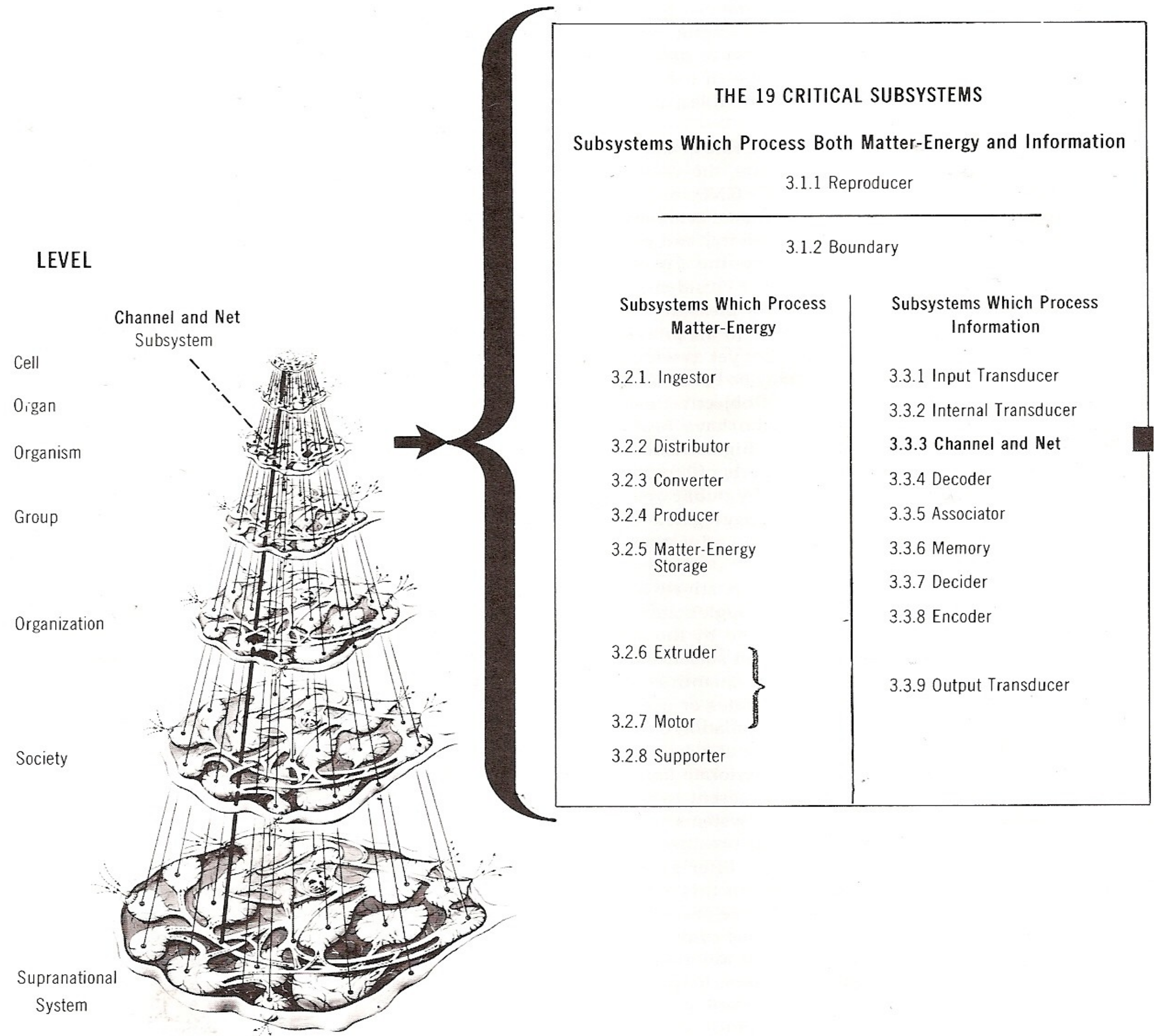
The concept of the holon is used in environmental philosophy, ecology and human ecology. Ecosystems are often seen as holons within one or many holarchies. Holons are seen as open subsystems of systems of higher order, with a continuum from the cell to the ecosphere.

Philosophy of history

In the philosophy of history, a holon is a historical event that makes other historical events inevitable. A holon is a controversial concept, in that some reject the inevitability of any historical event. A special category of holon is technology, which implies a perspective on how technologies have the potential to dictate history.

Living Systems (1978 Miller)

In 1978, together with his wife and collaborator Jessie, Miller made the case for a unified approach to the biological, psychological and social sciences in the book "Living Systems" a compilation and synthesis that he regarded as the capstone of his career, 25 years in the making[2] which founded the field of Living systems theory.



The self-similar nested hierarchy of living systems from the cell to the supranational system: on each level we identify the same 8 subsystems processing matter energy and the 9 subsystems processing information.

Living systems Wikipedia

Miller considers living systems as a subset of all systems. Below the level of living systems, he defines space and time, matter and energy, information and entropy, levels of organization, and physical and conceptual factors, and above living systems ecological, planetary and solar systems, galaxies, and so forth.[1].

Living systems are by definition open self-organizing systems that have the special characteristics of life and interact with their environment. This takes place by means of information and material-energy exchanges. Living systems can be as simple as a single cell or as complex as a supranational organization such as the European Economic Community. Regardless of their complexity, they each depend upon the same essential twenty subsystems (or processes) in order to survive and to continue the propagation of their species or types beyond a single generation.[2].

Miller said that systems exist at eight "nested" hierarchical levels: cell, organ, organism, group, organization, community, society, and supranational system. At each level, a system invariably comprises 20 critical subsystems, which process matter/ energy or information except for the first two, which process both matter/energy and information: reproducer & boundary.

The processors of matter/energy are:

- Ingestor, Distributor, Converter, Producer, Storage, Extruder, Motor, Supporter

The processors of information are

- Input transducer, Internal transducer, Channel and net, Timer (added later), Decoder, Associator, Memory, Decider, Encoder, Output transducer.

Miller's Living systems theory

James Grier Miller in 1978 wrote a 1,102-page volume to present his living systems theory. He constructed a general theory of living systems by focusing on concrete systems—nonrandom accumulations of matter-energy in physical space-time organized into interacting, interrelated subsystems or components. Slightly revising the original model a dozen years later, he distinguished eight "nested" hierarchical levels in such complex structures. Each level is "nested" in the sense that each higher level contains the next lower level in a nested fashion.

His central thesis is that the systems in existence at all eight levels are open systems composed of 20 critical subsystems that process inputs, throughputs, and outputs of various forms of matter/energy and information. Two of these subsystems—reproducer and boundary—process both matter/energy and information. Eight of them process only matter/energy. The other 10 process information only.

> All nature is a continuum. The endless complexity of life is organized into patterns which repeat themselves—theme and variations—at each level of system. These similarities and differences are proper concerns for science. From the ceaseless streaming of protoplasm to the many-vectored

activities of supranational systems, there are continuous flows through living systems as they maintain their highly organized steady states.*[3]*

Seppänen (1998) says that Miller applied general systems theory on a broad scale to describe all aspects of living systems" [4]

Topics in living systems theory

Miller's theory posits that the mutual interrelationship of the components of a system extends across the hierarchical levels. Examples: Cells and organs of a living system thrive on the food the organism obtains from its suprasystem; the member countries of a supranational system reap the benefits accrued from the communal activities to which each one contributes. Miller says that his eclectic theory "ties together past discoveries from many disciplines and provides an outline into which new findings can be fitted".[5]

Miller says the concepts of space, time, matter, energy, and information are essential to his theory because the living systems exist in space and are made of matter and energy organized by information. Miller's theory of living systems employs two sorts of spaces: physical or geographical space, and conceptual or abstracted spaces. Time is the fundamental "fourth dimension" of the physical space-time continuum/spiral. Matter is anything that has mass and occupies physical space. Mass and energy are equivalent as one can be converted into the other. Information refers to the degrees of freedom that exist in a given situation to choose among signals, symbols, messages, or patterns to be transmitted.

Other relevant concepts are system, structure, process, type, level, echelon, suprasystem, subsystem, transmissions, and steady state. A system can be conceptual, concrete or abstracted. The structure of a system is the arrangement of the subsystems and their components in three-dimensional space at any point of time. Process, which can be reversible or irreversible, refers to change over time of matter/energy or information in a system. Type defines living systems with similar characteristics. Level is the position in a hierarchy of systems. Many complex living systems, at various levels, are organized into two or more echelons. The suprasystem of any living system is the next higher system in which it is a subsystem or component. The totality of all the structures in a system which carry out a particular process is a subsystem. Transmissions are inputs and outputs in concrete systems. Because living systems are open systems, with continually altering fluxes of matter/energy and information, many of their equilibria are dynamic—situations identified as steady states or flux equilibria.

Miller identifies the comparable matter-energy and information processing critical subsystems. Elaborating on the eight hierarchical levels, he defines society, which constitutes the seventh hierarchy, as "a large, living, concrete system with [community] and lower levels of living systems as subsystems and components". [6] Society may include small, primitive, totipotential communities; ancient city-states, and kingdoms; as well as modern nation-states and empires that are not supranational systems. Miller provides general descriptions of each of the subsystems that fit all eight levels.

A supranational system, in Miller's view, "is composed of two or more societies, some or all of whose processes are under the control of a decider that is superordinate to their highest echelons" [7]. However, he contends that no supranational system with all its 20 subsystems under control of its decider exists today. The absence of a supranational decider precludes the existence of a concrete supranational system. Miller says that studying a supranational system is problematical because its subsystems

> ...tend to consist of few components besides the decoder. These systems do little matter-energy processing. The power of component societies [nations] today is almost always greater than the power of supranational deciders. Traditionally, theory at this level has been based upon intuition and study of history rather than data collection. Some quantitative research is now being done, and construction of global-system models and simulations is currently burgeoning.[8]

At the supranational system level, Miller's emphasis is on international organizations, associations, and groups comprising representatives of societies (nation-states). Miller identifies the subsystems at this level to suit this emphasis. Thus, for example, the reproducer is "any multipurpose supranational system which creates a single purpose supranational organization" (p. 914); and the boundary is the "supranational forces, usually located on or near supranational borders, which defend, guard, or police them" (p. 914).

Strengths of Miller's theory

Not just those specialized in international communication, but all communication science scholars could pay particular attention to the major contributions of LST to social systems approaches that Bailey[9] has pointed out:

- The specification of the 20 critical subsystems in any living system.
- The specification of the eight hierarchical levels of living systems.
- The emphasis on cross-level analysis and the production of numerous cross-level hypotheses.
- Cross-subsystem research (e.g., formulation and testing of hypotheses in two or more subsystems at a time).
- Cross-level, cross-subsystem research.

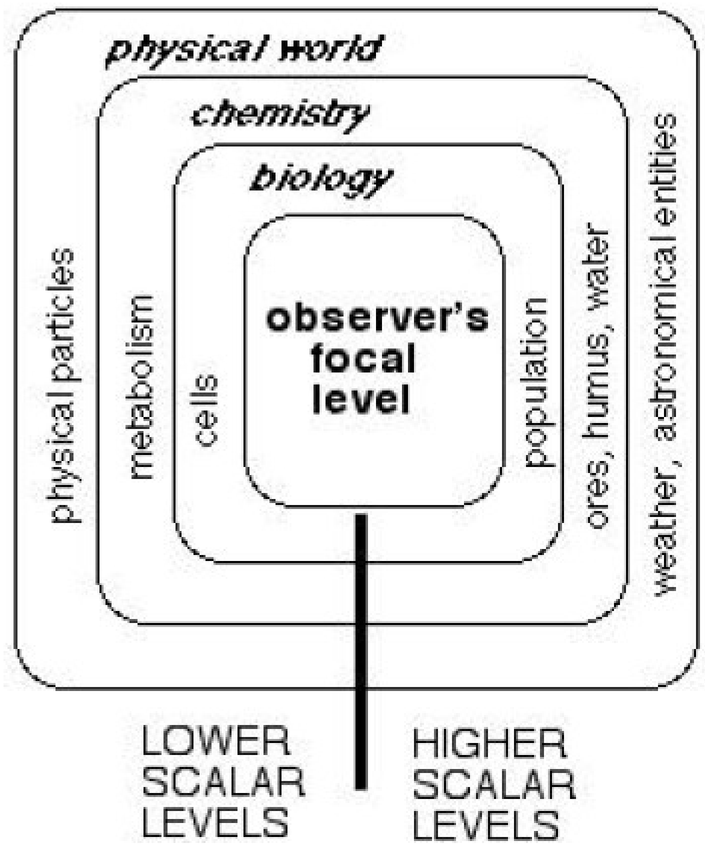
Bailey says that LST, perhaps the "most integrative" social systems theory, has made many more contributions that may be easily overlooked, such as: providing a detailed analysis of types of systems; making a distinction between concrete and abstracted systems; discussion of physical space and time; placing emphasis on information processing; providing an analysis of entropy; recognition of totipotential systems, and partipotential systems; providing an innovative approach to the structure-process issue; and introducing the concept of joint subsystem—a subsystem that belongs to two systems simultaneously; of dispersal—lateral, outward, upward, and downward; of inclusion—inclusion of something from the environment that is not part of the system; of artifact—an animal-made or human-made inclusion; of adjustment process, which combats stress in a system; and of critical subsystems, which carry out processes that all living systems need to survive.[10]

LST's analysis of the 20 interacting subsystems, Bailey adds, clearly distinguishing between matter/energy processing and information-processing, as well as LST's analysis of the eight interrelated system levels, enables us to understand how social systems are linked to biological systems. LST also analyzes the irregularities or "organizational pathologies" of systems functioning (e.g., system stress and strain, feedback irregularities, information-input overload). It explicates the role of entropy in social research while it equates negentropy with information and order. It emphasizes both structure and process, as well as their interrelations [11]

Limitations

It omits the analysis of subjective phenomena, and it overemphasizes concrete Q-analysis (correlation of objects) to the virtual exclusion of R-analysis (correlation of variables). By asserting that societies (ranging from totipotential communities to nation-states and non-supranational systems) have greater control over their subsystem components than supranational systems have, it dodges the issue of transnational power over the contemporary social systems. Miller's supranational system bears no resemblance to the modern world-system that Wallerstein (1974) described although both of them were looking at the same living (dissipative) structure.

Compositional hierarchy vs. Subsumption hierarchy (2002 Salthe)



This figure from Salthe [Salthe, 2005] can be taken as a mandala, suggesting the relationship between the scalar levels of extensional complexity and the integrative levels of intensional complexity. The observer arises out of the physical – chemical and biological realms as the peak of a pyramid rising from the left, but at the same time is embedded in these containing realms as a thought from the right.

In order to underline the crucial difference between compositional hierarchies (extensional complexity) and subsumption hierarchies (intensional complexity) we extensively cite Salthe [Salthe 2002 revised 2008]:

Hierarchies have two known logical forms:

1. the **compositional hierarchy** (including a synchronic map of the command hierarchy), which in applications I have called the 'scale hierarchy'. The picture of macromolecules inside of a living cell inside of an organism is a familiar image of one important application. This form is suited to synchronic modeling of systems as they are at any given moment.

- the **subsumption hierarchy** (including a diachronic model of the trajectory of a given command), which I have called the 'specification hierarchy'. The Linnaean hierarchy in biological systematics has this form. This form is suitable to diachronic modeling of emergent forms.

- Cliff Joslyn has provided the following comparative table of logical properties:

Meronomy	Taxonomy
--------	--------
Whole/part	General/specific
is-a-part-of	is-a-kind-of
Composition	Subsumption
Containment	Inheritance
Modularity	Specification

General properties:

Hierarchies are examples of 'partial ordering' in logic. That is, the items being ordered could be ordered in other ways as well. Hierarchies order entities, processes or realms into a system of levels. The ordering principle ('is-a-part-of' or 'is-a-kind-of') is transitive across levels. In both of these hierarchies, when used to model systems, higher levels control (regulate, interpret, harness) lower levels, whose behaviors are made possible by properties generated at still lower levels. So higher levels provide boundary conditions on the behaviors of lower levels -- behaviors initiated by still lower level configurations (see below for the usage of 'higher' and 'lower'). It is important to realize that only some users of hierarchical forms would insist that particular levels exist in actuality. Levels are discerned from hierarchical analysis, aimed at constructing / discovering Nature's 'joints' with respect to given projects. Hierarchies thus provide models of systems that are susceptible to analysis into different levels.

(a) To use the compositional hierarchy we need to stipulate a focal level, as well as a lower and a higher, making up a 'basic triadic system' -- as, e.g., when the behavior of living cells is initiated by chemical events, and controlled by organismic events. The three level form insures stability because with it in place (a third level always anchoring relations between the other two), the focal level cannot be reduced either upward or downward by assimilation into a contiguous level. Here we should note that this hierarchy has been invoked to explain how the world manages to be as stable as it is. The triadic form reflects the putative way in which levels would have evolved, by interpolation between primal highest and lowest ones, as when biology would have emerged as organizational forms between chemical activities in an environmental energy dissipative configuration.

(b) In the subsumption hierarchy the highest relevant level is always the one in focus, with all the lower levels of the hierarchy providing cumulative initiating conditions simultaneously upon it. This reflects the fact that this hierarchy is implicitly evolutionary, with the levels being viewed as having emerged consecutively from the lowest, or most general (or generally present), up -- as with, e.g., biology emerging from chemistry, both historically and at any given moment. The two-level form is unstable, allowing new levels to emerge at the top of the hierarchy. Use of this form provides us with a model allowing for emergent changes in the world.

Hierarchical analysis is always driven by a given problem or project.

Formal relations between levels:

(a) The compositional hierarchy is one of parts nested within wholes, as, e.g., [... [species [population [organism [gene [...]]]]]], where [higher level [focal level [lower level]]]. The logic reflects Russell's logical types. In principle the levels just keep going, receding at both ends from the focal level. (It may be noted that this structure probably is rooted in our visual experiences.)
If the parts are functional in some given analysis, they are referred to as components, if not they are constituents. As one goes down the hierarchy, the relative number of constituents per level increases, giving a measure of the 'span' of the hierarchy.

(b) The subsumption hierarchy is one of classes and subclasses, as e.g., {material world {biological world {social world }}}, where {lower level(s) { highest level}}. The focus of analysis is always the highest level, which is the innermost level of the hierarchy. The logic reflects Ryle's categories. Higher levels inherit all the properties of the lower levels.

(c) A note on levels terminology: The levels in a subsumption hierarchy have been referred to as 'integrative levels' inasmuch as the higher levels integrate the lower levels' properties and dynamics under their own rules. 'Levels of reality' and 'ontological levels' have been used in subsumption as well. One sees other labels, such as 'levels of organization' or 'levels of observation' used for either kind of hierarchy. I have used 'scalar levels' or 'levels of scale' for application of the compositional hierarchy to material systems for dynamical reasons (see below under 'Criteria').

Style of growth of the hierarchy:

(a) A compositional hierarchy adds levels by interpolation between existing levels. In this way the system must be an expanding one. Therefore, an assumption required for application of this hierarchy would be the Big Bang (or other expanding system). The actual process of formation of a level would involve the cohesion of entities out of lower level units guided by higher level boundary conditions. This process is little understood since this hierarchy has largely been used for synchronic analyses.

(b) In the subsumption hierarchy new levels would emerge from the current highest one. So this system too can grow -- but not in space. Growth here is by the accumulation of informational constraints, modeled as a process of refinement by way of adding specification. New levels, marked by subclasses reflect thresholds of system structural reorganization.

Criteria:

(a) In application of the compositional hierarchy to actual natural systems, components at different levels must differ in size roughly by orders of magnitude. Otherwise components at different levels would interact dynamically, in which case there would not be different levels functionally.

(b) Levels in a subsumption hierarchy mark the qualitative differences of different realms of being, as in 'physical realm' versus 'biological realm'. This hierarchy is open at the top; the innermost level is unbounded above, and so free to give rise to ever higher levels.

Complexity:

(a) A compositional hierarchy provides a model of ‘extensional complexity’, the sign of which is nonlinear and chaotic dynamics, allowed by the fact that at any locale at any level in this hierarchy there could be a mixture of different kinds of information (relations, variables, constants of different kinds, attractors) which are not governed by a single overall structure. It is useful here to contrast complexity with complication. A flat hierarchy with few levels could tend to show more complicated behavior than a hierarchy with more levels, which would have more constraints imposed top-down.

(b) A subsumption hierarchy embodies intensional complexity, which characterizes a system to the degree that it is susceptible to many different kinds of analyses.

Dynamical relations:

(a) A compositional hierarchy represents a single moment in space, so its dynamics represent homeostasis, not change. Large scale moments "contain" many small scale moments. It is often suggested that scalar levels fundamentally signal rate differences rather than component size differences. We may note that the two most often go together. The problem appears in cases that are said to be non-nested, where, e.g., a much slower rate in a component of a cycle would regulate the rate of the entire cycle. It would be rare, however, for such rates to differ by orders of magnitude, and so many of these examples are likely not hierarchical at all. If we allowed mere size differences rather than scale differences to be the criterion, then the constraint of nestedness would be lifted. In any case:

Because of the order of magnitude differences between levels in the compositional hierarchy, dynamics at different levels do not directly interact or exchange energy, but transact by way of mutual constraint (i.e., via informational connections). The levels are screened off from each other dynamically. Because of this dynamical separation of levels, informational exchanges between levels are non-transitive, requiring interpretation at the boundaries between levels.

So, if focal level dynamics are represented by variables in an equation, then the results of dynamics at contiguous levels would be represented by (nonrecursive) constants. Larger scale dynamics are so slow with respect to those at the focal level, that the current value of their momentary result appears relatively unchanging at the focal level. Cumulated results of lower scale dynamics also appear relatively or statistically unchanging at the focal level, as it takes a very long time in lower scale moments to effect a change detectable at the focal level -- these points are the essence of dynamical 'screening off' in compositional hierarchy models.

Note that, because of these relations, thermodynamic equilibria would be more rapidly achieved per unit volume at a lower scalar level, delivering an adiabatic principle relating to screening off. While change of any kind (development, acceleration, diffusion) is relatively more rapid at lower levels, absolute translational motion is more rapid at higher levels. Thus, higher levels provide modes of convection for the dissipation of energy gradients, which would otherwise proceed by slow conduction instead. Related to these matters, we should note that metabolic rates and development are absolutely much faster in smaller dissipative structures (organisms, fluid vortices, etc.), and their natural life spans are shorter than in larger scale ones.

One sometimes sees the term ‘heterarchy’, posed in opposition to the scale hierarchy because of supposed failures of actual systems to conform to hierarchical constraints. One needs to recall here again that hierarchy is a conceptual construction, an analytical tool, and use of it does not imply that the world itself is actually hierarchically organized. It does seem to be so in many ways, but to suppose that this is the sole principle needed in understanding the world would be naive. It is one tool among many. But often this ‘hetero’ opposition to hierarchy is based merely on faulty understanding. For example, the tides are affected (partially controlled) by gravitational effects associated with the moon; yet the oceans are not nested inside the moon. As in classical thermodynamics, it is important to see the whole system correctly. The oceans are nested, along with the earth itself, within the solar system, and from the hierarchical point of view, these effects on the tides emanate from the solar system, not merely from the moon. (Demurrer: As we descend in applications through the realm of fundamental particles, it may be that some of these rules would break down [via nonlocality, etc.]. Hierarchical constructs model events and informational transactions in the material world, defined as the realm of friction and lag in the affairs of chemical elements and their compositions.)

(b) Dynamics in a subsumption hierarchy are entrained by development, which is modeled as a process of refinement of a class, or increased specification of a category. It is important to note that this process is open-ended in the sense that there could be many coordinate subclasses of a given class. That is, the potentials arising within any class form a tree. So, in {physical realm { material realm { biological realm }}}, or {mammal { primate { human }}} each hierarchy follows just one branch of a tree. Rylean categories can branch into new distinctions (and this forms a link with the scalar hierarchy because this would give rise as well to new logical types). Evolution (unpredictable change) is one -> many, and thus we have been able to picture organic evolution using the Linnaean hierarchy.

The fact that functionally this is a two-level hierarchy makes it susceptible to change, because, without the anchoring provided by a third level, it could be reduced to a single level. How is its direction into new subclasses insured (giving rise to the hierarchy)? In models of the material world this is afforded by the fact that information, once in place (or once having had an effect), marks a system irrevocably. Marks in material systems are permanent. If a system continues to exist, it must march forward if it changes; there can be no reversal of evolution. Since change in the material world is entrained by the Second Law of thermodynamics, we have here a link between the two hierarchy models because the Second Law can be seen to be a result of Universal expansion being too fast to allow the global equilibration of matter. As noted above, this expansion is also what affords the interpolation of new levels in a compositional hierarchy.

So, development of a subsumptive hierarchy model requires a two-level basic form. Yet these hierarchies involve more than just two levels. Why do not the more general levels prevent change, as by the weight of their accumulated information? Here we are led to note another aspect of development, which is perfectly general. The amount of change required to launch a new level is ever smaller as a hierarchy develops -- refinements are just that. The more general levels do continue to exert their influence; e.g., biology is a kind of chemistry, and humans are a kind of mammal. The key to understanding this situation is that in the subsumption hierarchy informational relations between levels are transitive. Thus, physical dynamics are fully active players in a biological system. This means that we can fully understand development in this hierarchical model using only two contiguous levels. New levels may branch off anywhere in the hierarchy, potentially giving rise to collections of coordinate subclasses.

Informational relations and semiotics:

(a) As noted above, informational relations between levels in a compositional hierarchy are non-transitive. The levels are screened off from each other dynamically, and influence each other only indirectly, via transformed informational constraints. Signals moving from one level to another are transformed at boundaries between the levels. When this is not the case, as when a signal from a higher level occasionally transits to a much lower level, that level suffers damage (as when an organism is hit by lightning, or, going the other way, if a given cell affects the whole organism, this could only be if its effect is promoted by the likes of cancer). Here we can note again the idea that levels different in scale dynamics deliver stability to a system, via the screening-off effect.

The interpolation of a new level between two others can be viewed as involving the appearance of a capability at the uppermost level (via fluctuation, self-organization and/or selection) for making a significant (to it) interpretation of events at what then becomes the lowermost level of the three. The upper level effectively disposes -- facilitates cohesion among -- some of what the lower level proposes. This requires energetic screening off between levels. As the arena of the upper level's interpretants, the new level acts as a filter or buffer between upper and lower. This allows us to see levels succeeding each other by a classification procedure whereby topological difference information is converted to (or coheres as) typological distinction information in an essentially top-down procedure.

(b) In a subsumption hierarchy the lower levels also make possible the emergence of a new realm, in an epigenetic process. And here too the process is top-down, but in a different sense, involving finality. Thus, e.g., we can see that organism sociality implies biology in the sense of material implication or conceptual subordination. Then, as organism sociality implies biology, biology implies chemistry, and so, because this is a process of refinement, only a very narrow set of possibilities could imply organism sociality. That is, chemistry could give rise to many kinds of supersystems, biology to fewer, and sociality to even fewer as the epigenetic system develops. Developments (in distinction from evolution) are always entrained by final causes, and approach them asymptotically with each emergence of a new realm. Involved here, as in all developments, is the process of senescence, a condition of information overload (recall that information in this hierarchy is transitive across levels), leading to overconnectivity, leading in turn to functional underconnectivity, leading in its turn to inflexibility and habit driven responses (loss of requisite variety), leading ultimately to loss of adaptability (inability to produce interpretants of novel situations).

Operator hierarchy (1999 Jagers op Akkerhuis)

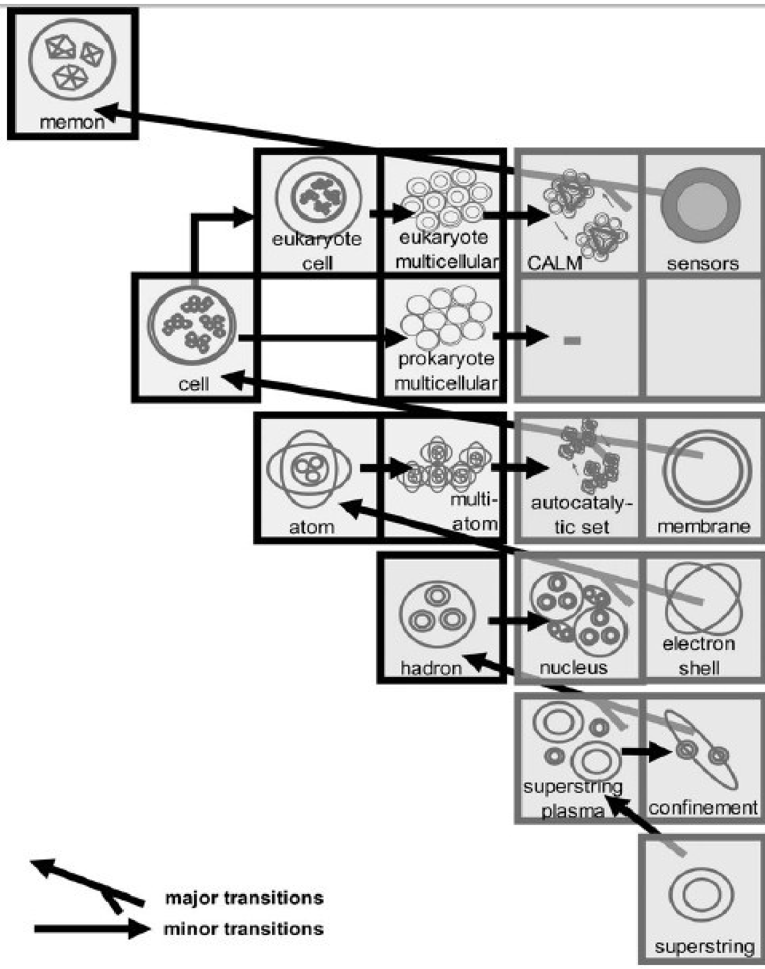


Fig. 1. The ranking of system types according to the operator hierarchy (Jagers op Akkerhuis & van Straalen, 1999; Jagers op Akkerhuis, 2001). Grey boxes indicate non-operator systems that play an important role in the operator hierarchy as intermediate closure states. Black upward arrows represent major transitions creating a new operator that shows a completely new type of closure. Black right-pointing arrows represent minor transitions. Empty cells and dashes indicate stages that have not yet evolved, but according to the logic of the hierarchy may potentially exist. Systems in the same vertical column share a common closure type. Titles above the columns indicate closure types. 'Interface' represents an emergent boundary. 'Hypercycle' represents an emergent second-order interaction cycle. 'Multi-operator' represents an emergent recurrent interaction between operators of the preceding type. 'Hypercycle mediating interface' (HMI) represents an interface that mediates the interactions of the hypercycle of the system involved with the world. 'Structural copying of information' (SCI) represents the property of systems to autonomously copy their structure and in this way reproduce their information. 'Structural auto-evolution' (SAE) represents the property of systems to improve, while living, the neural structures that contain their information. CALM stands for a Categorizing And Learning Module, representing a hypercyclic neural interaction pattern.

Network hierarchy (2002 Barabási)

"To build a modular network we started with a single node (see Figure 16.1 A) and created three copies of it, connecting them to the old node and to each other, obtaining a little four-node module (B). We next generated three copies of this module, linking the peripheral nodes of each new copy to the central node of the old module, obtaining a sixteen node network (C). Another "copy and link" step again quadrupled the number of nodes, resulting in a sixty-four-node network (D).

While we could have continued this process indefinitely, we stopped here and inspected the intricate structure of the network.

First it was modular by construction (self-similar fractal). At the lowest organizational level it was made of many highly connected four-node modules. These modules were the building blocks of the larger sixteen-node modules, which in turn were the major components of the sixty-four-node network.

Second, a highly connected central hub with thirty-nine links held the network together. The central nodes of the sixteen-node modules served as somewhat smaller local hubs, with fourteen links. Numerous nodes with a few links only accompanied these hubs, resulting in the familiar hierarchy of many small nodes held together by a few large hubs, a signature of scale-free networks. Indeed, the number of nodes with exactly k links followed a power law, confirming the model's scale-free nature. For the construction described above, the degree distribution follows a power law $P(k) = k^{\alpha}$ with alpha ~2.26." Source: Barabási 2003.

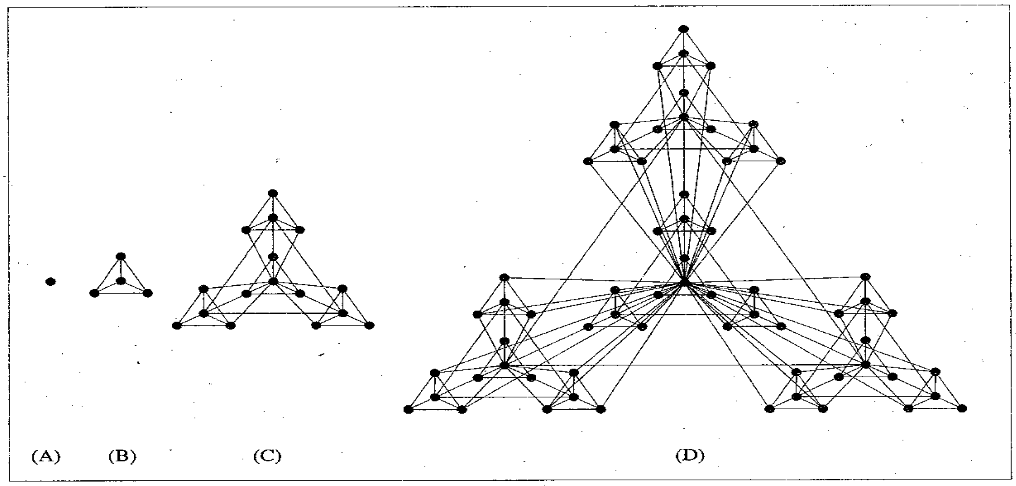


Figure 16.1 *We can generate a hierarchical network by starting from a single node (A) and making three copies of it, connecting the new nodes to the old node and to each other, obtaining the four-node structure shown in (B). In the next step we make three copies of our four-node structure and place them around the old module, connecting the peripheral nodes of the new modules to the central node of the old module and linking the central nodes of the new modules to each other. The obtained network will have sixteen nodes, as shown in (C). We can repeat the same process, creating now three copies of the C module and placing them around the old one, connecting the peripheral nodes to the center of the old module and the central in (D). The process can be continued indefinitely, each time generating a four-times-larger network. The obtained network is scale free: One can clearly distinguish a hierarchy of many small nodes, held together by a few large hubs. It is also modular, made of a hierarchy of larger and larger modules. Indeed, one can easily deconstruct the network shown in (C) into sixteen four-node modules, or four sixteen-node modules. An interesting property of the network is that it displays hierarchical clustering: It is made of many highly interlined four-node modules, which in turn form less interlinked sixteen-node modules, which are the building blocks of an even looser sixty-four-node module. Recently we learned that such hierarchical clustering is a generic property of a large number of real networks, from the cell to the World Wide Web.*

Note that the modular construction of the network follows a self-similar fractal like algorithm and suggests the fractal nature of scale-free networks. Hierarchical modularity is a generic property of most real networks accompanying scale-free architecture from cells over language to the Internet.

The Figure below shows an example of modular clustering in social networks. Small clusters of nodes interlinked with strong ties are interconnected with weak ties in a larger network.

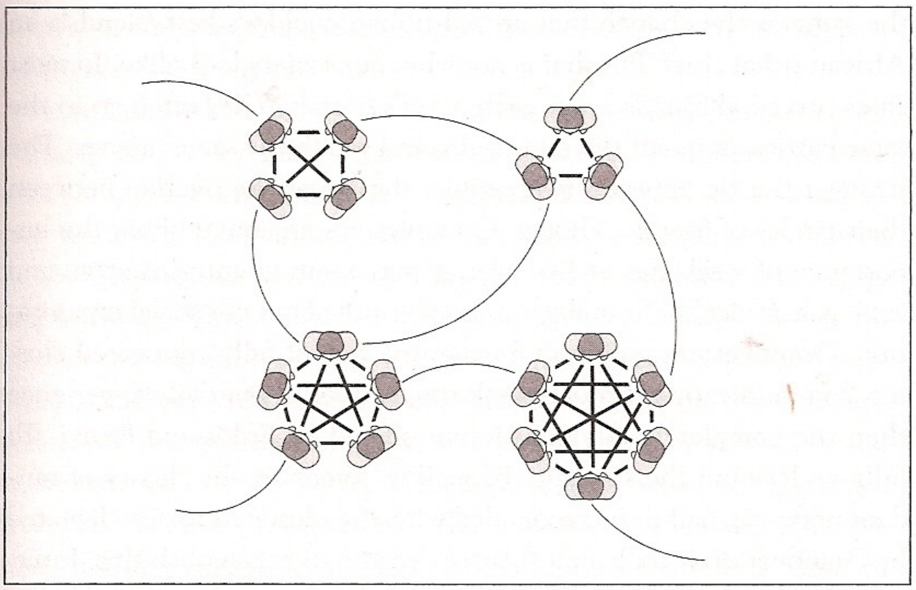


Figure 4.1 Strong and Weak Ties. *In Mark Granovetter's social world, our close friends are often friends with each other as well. The network behind such a clustered society consists of small, fully connected circles of friends connected by strong ties, shown as bold lines. Weak ties, shown as thin lines, connect the members of these friendship circles to their acquaintances, who have strong ties to their own friends. Weak ties play an important role in any number of social activities, from spreading rumors to getting a job.*

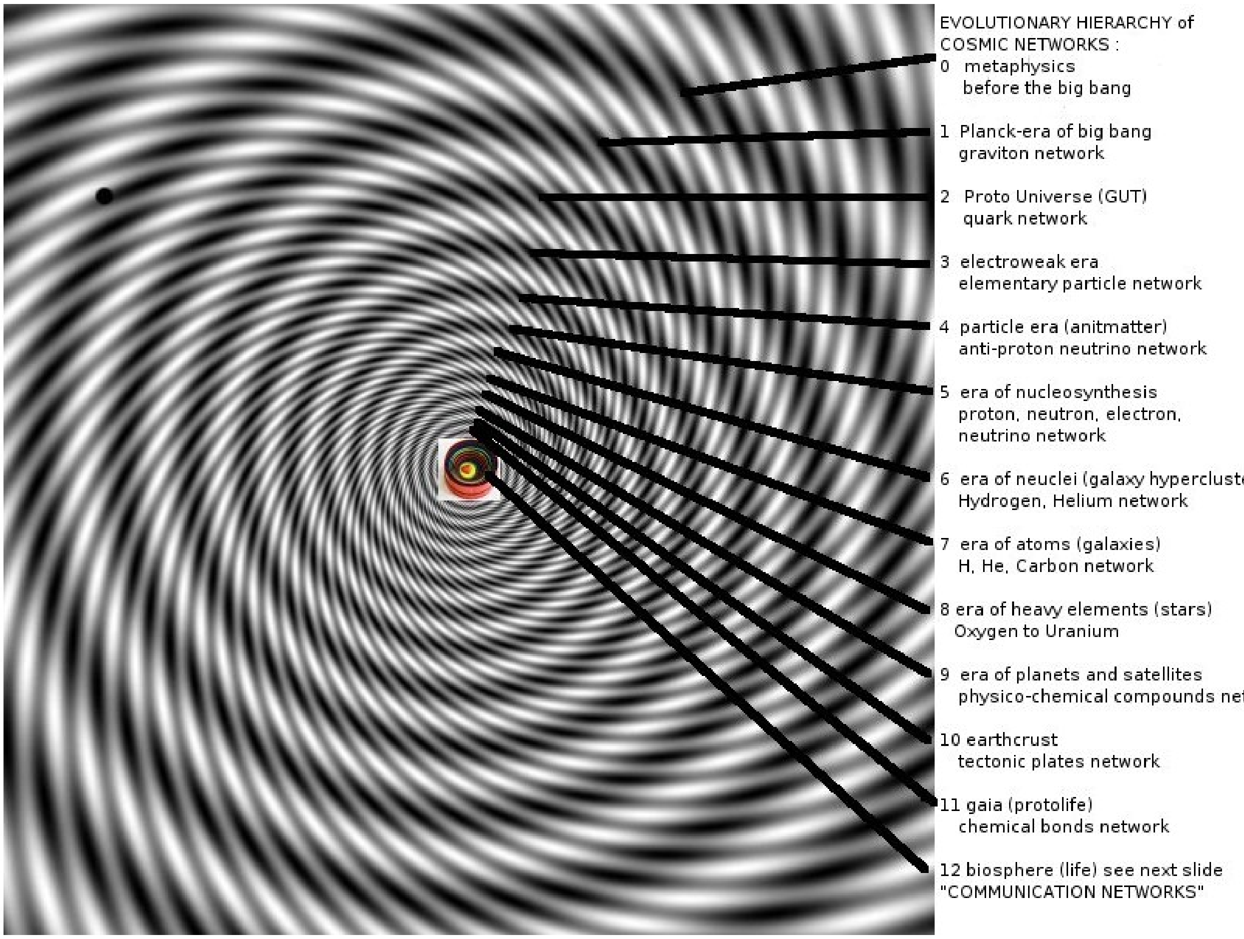
“Thanks to the high interest in clustering generated by Watts and Strogatz's unexpected discovery, the scientific community has subsequently scrutinized many networks. We now know that clustering is present on the Web, we have spotted it in physical lines that connect computers on the Internet; economists have detected it in the network describing how companies are linked by joint ownership, ecologists see it in food webs that quantify how species feed on each other in ecosystems; and cell biologists have learned that it characterizes the fragile network of molecules packed within a cell".

This citation of Barabási (Barabási 2003)shows that clustering is ubiquitous and a generic property of empirically observed complex networks.

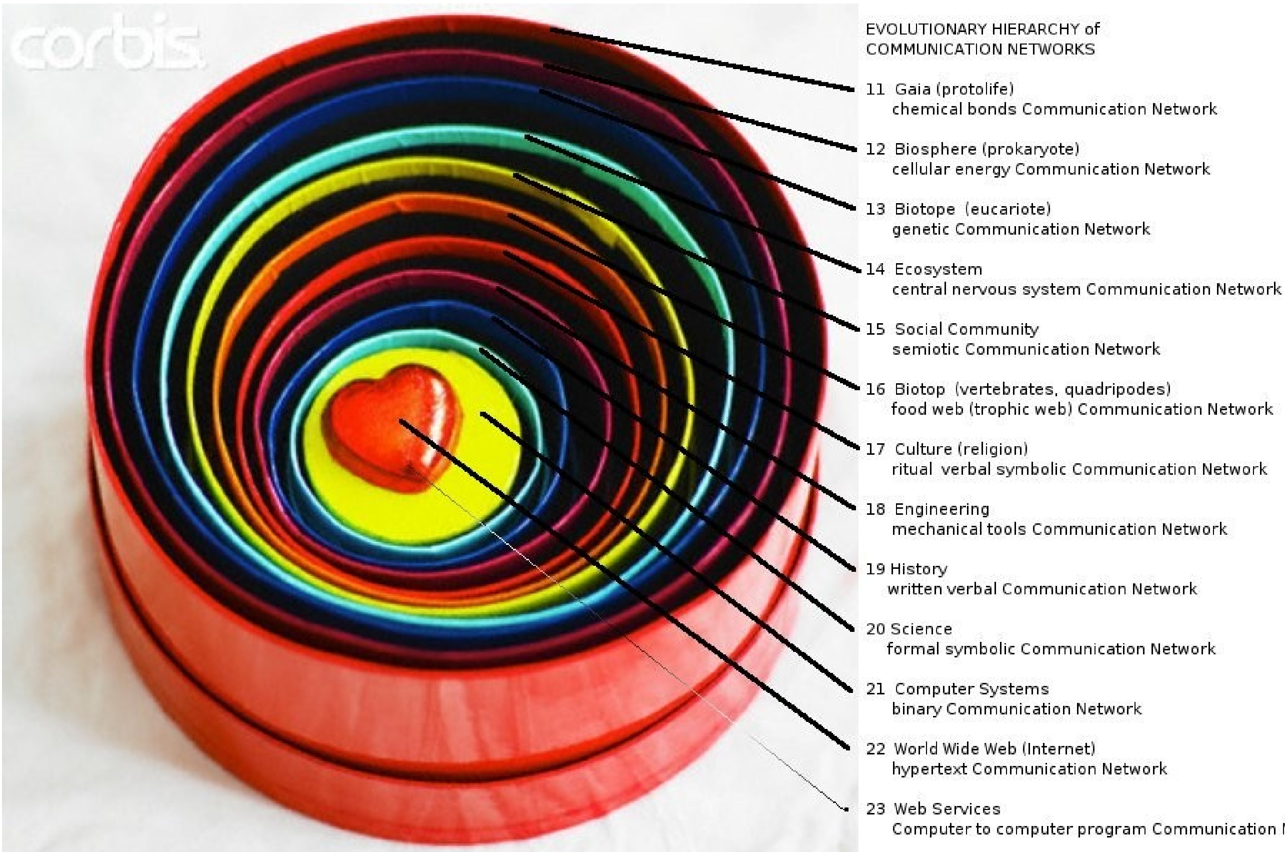
As we will show in the chapter on Neural Networks a modular network of the above type can be mapped on an Artificial Neural Network of a multilayer feed-forward network with back-propagation called also multilayer perceptron. While the above network limits itself to the description of the network topology the ANN model comprises the internal dynamics and information flow within the network: bottom up integration of inputs and top down differentiation through error back-propagation.

Levels of evolutionary hierarchy (2008 Winiwarter)

Hierarchies are ubiquitous. You find them in any science and in any field of research.
In fact the hierarchical "vision" of a system is a way to put a *static* order into the view of a complex system.
Networks are everywhere. You find them from galaxies to the World Wide Web. Again the networks don't exist, they are only a mental framework to put a *dynamic* order into the view of a complex system.
The Universe is a hierarchy – most people agree that it is not a flatland – but it can also be seen as a hierarchy of networks. How to put an order into this complex mess of viewpoints, points of view and world views?
We attempt to establish an evolutionary hierarchy based on clearly stated criteria.
A hierarchy is an ordered set - ordered according to an order criterion.
As order criterion for the universal evolutionary hierarchy we propose the time of emergence during evolution as observed by todays science.
By time of emergence we understand the first observation during the process of evolution of a given hierarchical level. Such the nested hierarchy of levels corresponds to the temporal sequence of their emergence.



The number of levels is arbitrary. For simplicity we choose 24 levels : 12 levels for the astrophysical evolution (deceleration and expansion of the universe from the big bang to the origins of biological life) and 12 levels from the early biosphere to the present of the Internet and Web Services (acceleration of evolution).



Evolutionary hierarchies are imbricated or embedded like "Russian dolls"

— *Humans have a natural tendency to find order in sets of information, a skill that has proven difficult to replicate in computers. Faced with a large set of data, computers don't know where to begin -- unless they're programmed to look for a specific structure, such as a hierarchy, linear order, or a set of clusters*. ScienceDaily (Aug. 28, 2008)

We introduced a zero level (background zero) for the metaphysical foundations of the model based on the standard big bang hypothesis.
Below an overview of astrophysical and biological hierarchical levels and the corresponding networks emerging at this level:

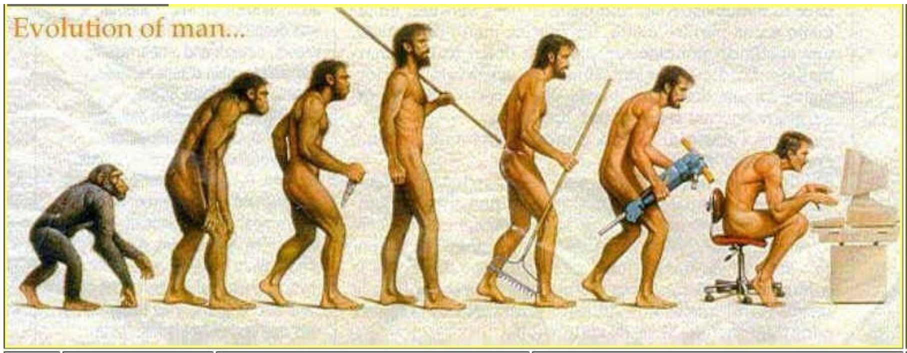
Such the 24 levels are imbricated like Russian dolls. Any other partition into a greater or smaller number of hierarchical levels would be equivalent as long as the partition respects the single order criterion, which is first time of emergence.

With this view, the higher levels in the hierarchy of complexity have autonomous causal powers that are functionally independent of lower-level processes. Topdown causation takes place as well as bottom-up action, with higher-level contexts determining the outcome of lower level functioning, and even modifying the nature of lower-level constituents.

Each of the hierarchical levels can be described as a complex interactive network. Each level having its characteristic "interaction units" or processors emerging from the prior level in the process of evolution.

Table of Evolutionary Hierarchies

	phase of evolution	Microscopic energy / information processor type	emerging energy / information field bonds linking the network	time scale since big bang	scientific discipline, link to empirical data	comments
0	metaphysics	man	consciousness	?	metaphysics	Short before the big bang?
1	Planck-era of big bang	graviton	super-grand-unified field	???	astrophysics theory	
2	proto-universe GUT era	quark	GUT, gravitation	10^{-43} sec	astrophysics theory	
3	electroweak era	elementary particles	strong force	10^{-35} sec	astrophysics theory	strong forces become distinct, perhaps causing inflation of the universe
4	particle era	antiproton, antineutrino antimatter	electromagnetic and weak force	10^{-10} sec	astrophysics theory	electromagnetic and weak forces become distinct
5	era of nucleosynthesis	protons, neutrons, electrons neutrinos	electromagnetic and weak force	10^{-3} sec	astrophysics theory	matter annihilates antimatter
6	era of nuclei	Hydrogen, Helium, electrons	strong nuclear	180 sec	astrophysics theory	Fusion ceases; normal matter is 75% hydrogen
7	era of atoms	H, He, C	carbon hypercycle	300.000 years	astrophysics	Atoms form, photons fly free and become background radiation
8	era of stars	H, He, C O, ... U	nuclear binding energy in atom (nuclide)	1 billion years	astrophysics	first galaxies form
	today			13.73 billion years	astronomy	humans observe the cosmos

				**time scale from today**		
	big bang			- 13.73×10^9 years *billion*		age of the universe
9	era of planets and moons	physico-chemical elements	trajectory bonds of planetary network	-4.54×10^9 years *billions*	astrophysics	
10	earthcrust	land-water agglomerates	geophysical bonds network	$-3.8x10^9$ years *billions*	geophysics	islands
11	gaia	macro-molecules hypercycles ATP	chemical bonds, trophic bonds in network (foodweb)	$-3.5x10^9$ years` *billions*	physics, phys-chem geology, geoscience, meteorology	mountains, lakes, rivers, floods
12	biosphere	proto-cells, selfreplicating unicellular metabolism (prokaryote) RNA	genetic network	$-3.3x10^9$ years *billions*	bio-chemistry	hypercycle, cell
13	biotope	cell (eucariote) DNA	energy transportation network, genetic network	$-2.5x10^9$ years *billions*	bio-chemistry, biology	genetic tradition, first photosynthesis by blue-green algae
14	ecosystem	multicellular organism	central nervous system network	$-2.0x10^9$ years *billions*	biology, plant, animal, ecosystems	oxygen from photosynthesis
15	social community	communication symbols, rituals	social communication network	$-1.0x10^9$ years *billions*	social systems; ethology,	first nucleated cells with organelles/ bacteria colonized single cells

16	Land plant animals	vertebrates	trophic web network	-500×10^6 years *millions*	plant, animal,	-200 to -100 million years Age of Dinosaurs
17	culture, religion	*homo economicus verbalis :* words	formal symbolic (verbal) communication network,	-2×10^6 years *millions*	economics linguistics, religion	oral traditiion, homo economicus
18	engineering & design	*homo mecanicus*: tools, buildings, machines	oral tradition and tools, mechanical systems network	-10.000 years	mechanical systems	it's in the cities that new forms of social interaction and mechanical systems emerge, (water, road)
19	history	*homo letteris*: icons, ideograms, letters	written communication network	-5.000 years	paleontology, history, communication systems	written tradition, communication networks, energy transportation networks (water, roads ... railroads, electricity, airports)
20	science	*homo scientificus*: formulas, laws	formal written symbolic community / network	-500 years	history of science,	formal written tradition; humans observe the cosmos
21	computers	*homo cyberneticus*: systems theories, systems laws	binary systems network	-50 years	cybernetics, science of science	transdisciplinary tradition
22	world-wide-web	*homo networked:* URL universal resource locator on computer networks	man computer internet network, artificial neural networks, scientometrics	-18 years	computer tech, neural networks scientometrics metaphilosophy	putting it all together on the web, this site
23	Semantic WEB	*computer communicans*	WEB services network	-2 years future	Amazon(Amazon Web Servicies), Google Semantic Web	computers communicating directly with computers

The network of processors create a field specific to the level, which is the interaction of all processors specific to the hierarchical level.
For a level to exist, all prior levels are necessary, since they constitute the environment of the new emerging level. There is no science without language, there is no language without semiotic communication, there is no semiotic communication without central nervous systems ...
There are no chemical compounds without atoms, there are no atoms without nucleons, there are no nucleons without quarks ...

A general process of emergence is described in the paper Autognosis, the theory of Hierarchical self-

image building systems (Winiwarter 1986). In this paper we advance the hypothesis of an underlying isomorphic self-organizational core process by which learning and evolutionary processes in general take place.
The core idea is that evolution is a simultanous process of global top down differention of the environment and local bottom up integration of elements or processors. As a generic example we describe the evolution of nucleosynthesis in a massive star with the emergence of nested cores. In each core there is synthesis of nucleons from protons to helium, from helium to carbon ...

A Summary of Principles of Hierarchy Theory

The Hierarchy theory is a dialect of general systems theory. It has emerged as part of a movement toward a general science of complexity. Rooted in the work of economist, Herbert Simon, chemist, Ilya Prigogine, and psychologist, Jean Piaget, hierarchy theory focuses upon levels of organization and issues of scale. There is significant emphasis upon the observer in the system.

Hierarchies occur in social systems, biological structures, and in the biological taxonomies. Since scholars and laypersons use hierarchy and hierarchical concepts commonly, it would seem reasonable to have a theory of hierarchies. Hierarchy theory uses a relatively small set of principles to keep track of the complex structure and a behavior of systems with multiple levels. A set of definitions and principles follows immediately:

Hierarchy: in mathematical terms, it is a partially ordered set. In less austere terms, a hierarchy is a collection of parts with ordered asymmetric relationships inside a whole. That is to say, upper levels are above lower levels, and the relationship upwards is asymmetric with the relationships downwards.

Hierarchical levels: levels are populated by entities whose properties characterize the level in question. A given entity may belong to any number of levels, depending on the criteria used to link levels above and below. For example, an individual human being may be a member of the level i) human, ii) primate, iii) organism or iv) host of a parasite, depending on the relationship of the level in question to those above and below.

Level of organization: this type of level fits into its hierarchy by virtue of set of definitions that lock the level in question to those above and below. For example, a biological population level is an aggregate of entities from the organism level of organization, but it is only so by definition. There is no particular scale involved in the population level of organization, in that some organisms are larger than some populations, as in the case of skin parasites.

Level of observation: this type of level fits into its hierarchy by virtue of relative scaling considerations. For example, the host of a skin parasite represents the context for the population of parasites; it is a landscape, even though the host may be seen as belonging to a level of organization, organism, that is lower than the collection of parasites, a population.

The criterion for observation: when a system is observed, there are two separate considerations. One is the spatiotemporal scale at which the observations are made. The other is the criterion for observation, which defines the system in the foreground away from all the rest in the background. The criterion for observation uses the types of parts and their relationships to each other to characterize the system in the foreground. If criteria for observation are linked together in an asymmetric fashion, then the criteria lead to levels of organization. Otherwise, criteria for observation merely generate isolated classes.

The ordering of levels: there are several criteria whereby other levels reside above lower levels. These criteria often run in parallel, but sometimes only one or a few of them apply. Upper levels are above lower levels by virtue of: 1) being the context of, 2) offering constraint to, 3) behaving more slowly at a lower frequency than, 4) being populated by entities with greater integrity and higher bond strength than, and 5), containing and being made of - lower levels.

Nested and non-nested hierarchies: nested hierarchies involve levels which consist of, and contain, lower levels. Non-nested hierarchies are more general in that the requirement of containment of lower levels is relaxed. For example, an army consists of a collection of soldiers and is made up of them. Thus an army is a nested hierarchy. On the other hand, the general at the top of a military command does not consist of his soldiers and so the military command is a non-nested hierarchy with regard to the soldiers in the army. Pecking orders and a food chains are also non-nested hierarchies.

Duality in hierarchies: the dualism in hierarchies appears to come from a set of complementarities that line up with: observer-observed, process-structure, rate-dependent versus rate-independent, and part-whole. Arthur Koestler in his "Ghost in The Machine" referred to the notion of holon, which means an entity in a hierarchy that is at once a whole and at the same time a part. Thus a holon at once operates as a quasi-autonomous whole that integrates its parts, while working to integrate itself into an upper level purpose or role. The lower level answers the question "How?" and the upper level answers the question, "So what?"

Constraint versus possibilities: when one looks at a system there are two separate reasons behind what one sees. First, it is not possible to see something if the parts of the system cannot do what is required of them to achieve the arrangement in the whole. These are the limits of physical possibility. The limits of possibility come from lower levels in the hierarchy. The second entirely separate reason for what one sees is to do with what is allowed by the upper level constraints. An example here would be that mammals have five digits. There is no physical reason for mammals having five digits on their hands and feet, because it comes not from physical limits, but from the constraints of having a mammal heritage. Any number of the digits is possible within the physical limits, but in mammals only five digits are allowed by the biological constraints. Constraints come from above, while the limits as to what is possible come from below. The concept of hierarchy becomes confused unless one makes the distinction between limits from below and limits from above. The distinction between mechanisms below and purposes above turn on the issue of constraint versus possibility. Forget the distinction, and biology becomes pointlessly confused, impossibly complicated chemistry, while chemistry becomes unwieldy physics.

Complexity and self-simplification: Howard Pattee has identified that as a system becomes more elaborately hierarchical its behavior becomes simple. The reason is that, with the emergence of intermediate levels, the lowest level entities become constrained to be far from equilibrium. As a result, the lowest level entities lose degrees of freedom and are held against the upper level constraint to give constant behavior. Deep hierarchical structure indicates elaborate organization, and deep hierarchies are often considered as complex systems by virtue of hierarchical depth.

Complexity versus complicatedness: a hierarchical structure with a large number of lowest level entities, but with simple organization, offers a low flat hierarchy that is complicated rather than complex. The behavior of structurally complicated systems is behaviorally elaborate and so complicated, whereas the behavior of deep hierarchically complex systems is simple.

Hierarchy theory is as much as anything a theory of observation. It has been significantly operationalized in ecology, but has been applied relatively infrequently outside that science. There is a negative reaction to hierarchy theory in the social sciences, by virtue of implications of rigid autocratic systems or authority. When applied in a more general fashion, even liberal and non-authoritarian systems can be described effectively in hierarchical terms. There is a politically correct set of labels that avoid the word hierarchy, but they unnecessarily introduce jargon into a field that has enough special vocabulary as it is.

Power laws and the laws of Power

> “Power laws are emergent general features of complex systems. Despite the complex and idiosyncratic features of organisms and the ecosystems where they occur, there are aspects of the structure and function of these systems that remain self-similar or nearly so over a wide range of spatial and temporal scales. Empirical power laws describe mathematically the hierarchical, fractal-like organization of these systems. Presumably these power laws reflect the outcome of simple rules or mechanisms. On the one hand, simple mechanisms that determine the structure and function of the fundamental components at the smallest scales constrain how these parts function when they are assembled in progressively larger subsets or hierarchies.
> On the other hand, simple mechanisms constrain the structure, and dynamics at the largest scales also place powerful limits on how the components interact and assemble in the large, complex system. Together, these bottom–up and top–down mechanisms give rise to power laws and other emergent features.”

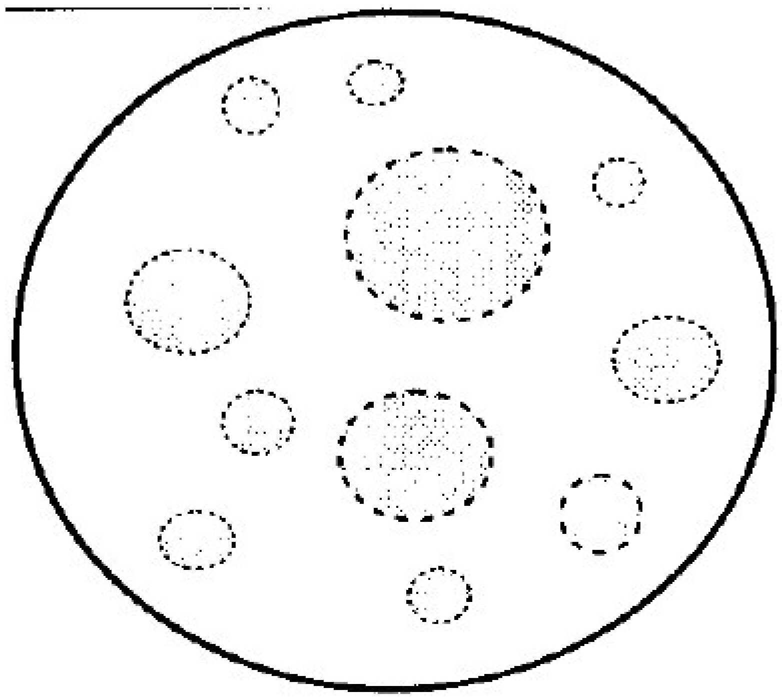
The fractal nature of nature: power laws, ecological complexity and biodiversity
James H. Brown, Vijay K. Gupta, Bai-Lian Li, Bruce T. Milne, Carla Restrepo and Geoffrey B.West
http://www.fractal.org/Bewustzijns-Besturings-Model/Fractal-Nature.pdf

"It is an interesting possibility that the power laws followed by so many different kinds of systems might be the result of downward constraintes exerted by encompassing supersystems."

Stanley N. Salthe, ***Entropy*** 2004, **6**, 335

Common 3-level hierarchical structure

Power laws of the Pareto-Zipf-Mandelbrot (hyperbolic fractal) type are observed for class-size distributions of virtually all evolutionary hierarchical levels ranging from the field of astrophysics to the Internet.



All observed regularities are based on a 3-level hierarchical description, see figure below:

Figure. The three-level hierarchy of a Pareto-Zipf-Mandelbrot PZM distribution: local processing units (small dots), processing unit classes (dotted circles) and global interaction system (fat circle)

Let us have a closer look at this hierarchy at hand of and example.

City-size distribution show PZM regularities for any country of the world.

Interaction units

Interaction units – small dots in the figure – are the third and basic level of the 3-level hierarchy: interaction system, equivalence classes, interaction units. In our example the basic local interaction unit is an inhabitant, which is assigned to a class (city) during the snapshot of the system.

The class size distribution of the system changes only due to three possible interactions:

- birth of an interaction unit (new inhabitant)
- death of an interaction unit (disappearance of an inhabitant) and
- migration of an interaction unit from one class (city) to another class (city) within the network during two consecutive snapshots (US census)

Interaction units may be closed energy information processors or operators as defined in the operator hierarchy approach.

In our example above the basic interaction units are human inhabitants, better households or oikos in our terminology. Basic households are the building blocks for aggregates on a town or city level.

Equivalence classes of interaction units

Equivalence classes – dotted circles in the figure – are aggregates of interaction units, cities in our example. The interaction units (inhabitants) belonging to the same class (inhabitants of the same city) are equivalent for the statistical analysis. The class sizes, number of operators per class, show the characteristic PZM distribution at a census measurement, that is a count of all individual inhabitants during a snapshot of the system.

There are few very big agglomerations like New York and Los Angeles with millions of inhabitants, few big agglomerations of hundred thousand inhabitants and very many small agglomerations in the range of 10.000 inhabitants. In quantitative geography this regularity is called rank-size rule

Interaction system, closed network of interaction units

The global system – fat circle in the figure - for which we observe a PZM regularity we call interaction system. This system is delimited within a boundary, frontier of the US in our example. This boundary or frontier is more or less impermeable to the interaction units of the network, while movements of interaction units (inhabitants) between equivalence classes (cities) within the system are frequent and relatively free.

Note that PZM regularities are observed only within a closed boundary of an interaction system. We observe PZM regularities for the entire United States but also for each individual state with the exception of Texas. An explanation for this exception may be the fact, that the frontiers of Texas are arbitrary straight lines on a map not corresponding to a quasi impermeable membrane.

The same approach of description of a 3-level hierarchy can be applied in astrophysics to massive stars for which we observe PZM regularities.

The **interaction system** level is the entire massive star (e.g. the sun or) with its surface as boundary. Within this system we have interactions between local **interaction units** called atoms (nuclei), which can be classified into **equivalence classes** called chemical elements. The sizes of the equivalence classes (frequencies of chemical elements) follow a PZM regularity. See figure later in this chapter.

Likewise we can analyze **any** interaction system revealing PZM regularity.

Let us take another example, a national economy.

The **interaction system** is the entire economy (e.g. a country or the entire world). Within this system we have interactions between local **interaction units** called monetary units (Dollars or Euros), which can be classified into **equivalence classes** called firms (turnover of a firm or assets of a firm). The sizes of the equivalence classes (firm sizes) follow a PZM regularity.

A short history of discovery across the disciplines

1897 Wilfredo Pareto, income distribution

The first extensive discussion of the problem how income is distributed among the citizens of a state was made by Vilfredo Pareto in 1897 [Pareto, 1987]. On the basis of data collected from numerous sources Pareto arrived at the following law:

In all places and at all times the distribution of income in a stable economy, when the origin of measurement is at a sufficiently high income level, will be given approximately by the empirical formula

(1) $n = a S^{Y}$

where n is the number of people having the income S or greater, a and .y are constants.

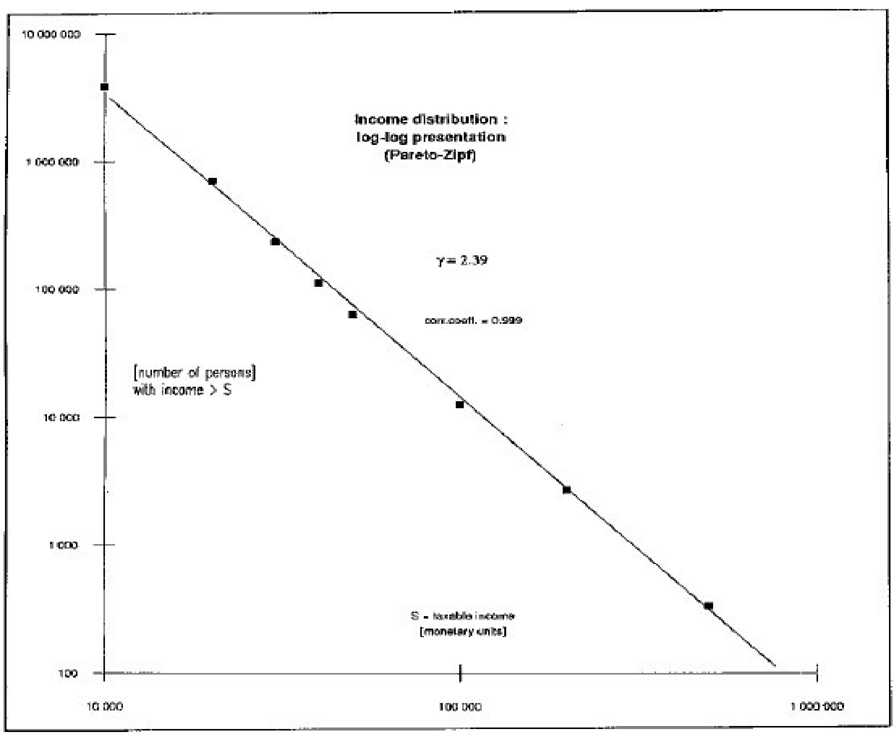


Figure. It is difficult to represent the data graphically within ordinary arithmetic scales . The data are taxable incomes of 1937 in France, but any other country and year yields distributions of the above type. Note the almost perfect correlation coefficient. Data Source [Winiwarter, 1992]

It is extremely interesting to note, that empirical observations of Pareto distributions are:

i) not markedly influenced by the socio-economic structure of the community under study

ii) not markedly influenced by the definition of "income" .

The Pareto law holds for a few hundred burghers of a city-state of the Renaissance up to the more than l00 million taxpayers in the USA. Essentially the same law continues to be followed by the distribution of "income", despite the changes in the definition of this term.

Note: this empirical evidence is a contradiction to any ideology striving for equal distribution of incomes. As we shall see below, this goal is just as unrealistic and unnatural as the goal to make all cities of a country of the same size i.e. the same number of inhabitants. Likewise it is 'unnatural' to make all business firms of equal size or to use in a text all words with equal frequency.

Pareto was intrigued by the generality of his discovery: "These results are very. remarkable . It is absolutely impossible to admit that they are due only to chance . “There is most certainly a cause, which produces the tendency of incomes to arrange themselves according to a certain curve.”

1913 Auerbach, the distribution of city sizes in countries

Looking for a new measure for population concentration, Auerbach [Auerbach, 1913] analyzed the distribution of cities within a country. He ranked the cities in decreasing order of inhabitants and discovered a relationship between rank and size of the type

(2) $S(j) = a\ j^{\beta}$

with $S(j)$ the size of the city ranked j, a and β are constants.

As an example let us consider the city-size distribution of France.

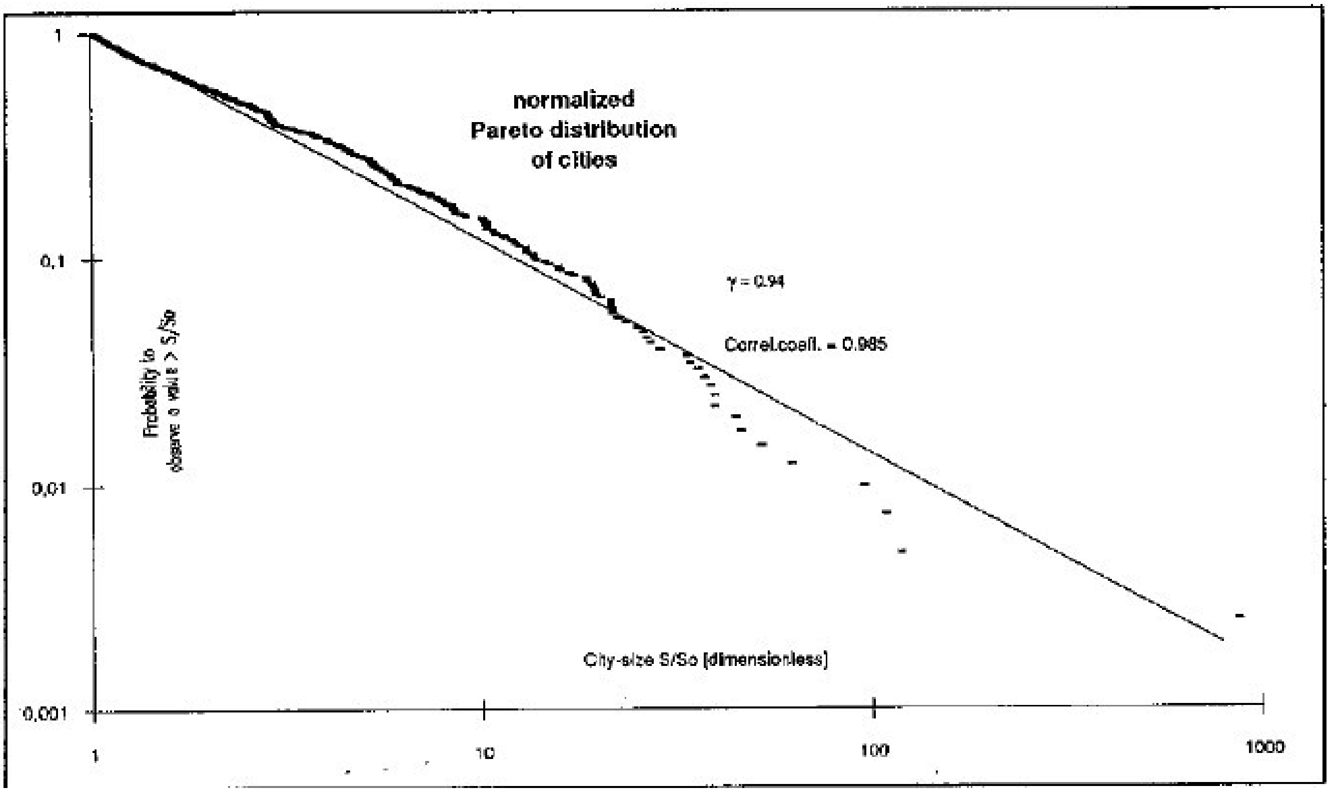


Figure. The cumulative probability as a function of the dimensionless symptom S/So. S is the city size and S_o the smallest or threshold size for of the observed set of cities (here 10.000). Data Source [Winiwarter, 1992]

1912 Willis-Yule, the distribution of species, genera and families in biological systems

Based on field observation in Ceylon in Willis [Willis, 1912] first noticed, that the distribution of species within the genera of an ecosystem follows a regularity, which is of the Pareto-Zipf type.

“ this type of curve holds not only for all the genera of the world, but also for all the individual families both of plants and animals, for endemic and non-endemic genera, for local floras and faunas ... it obtains too, for all the deposits of Tertiary fossils examined ."

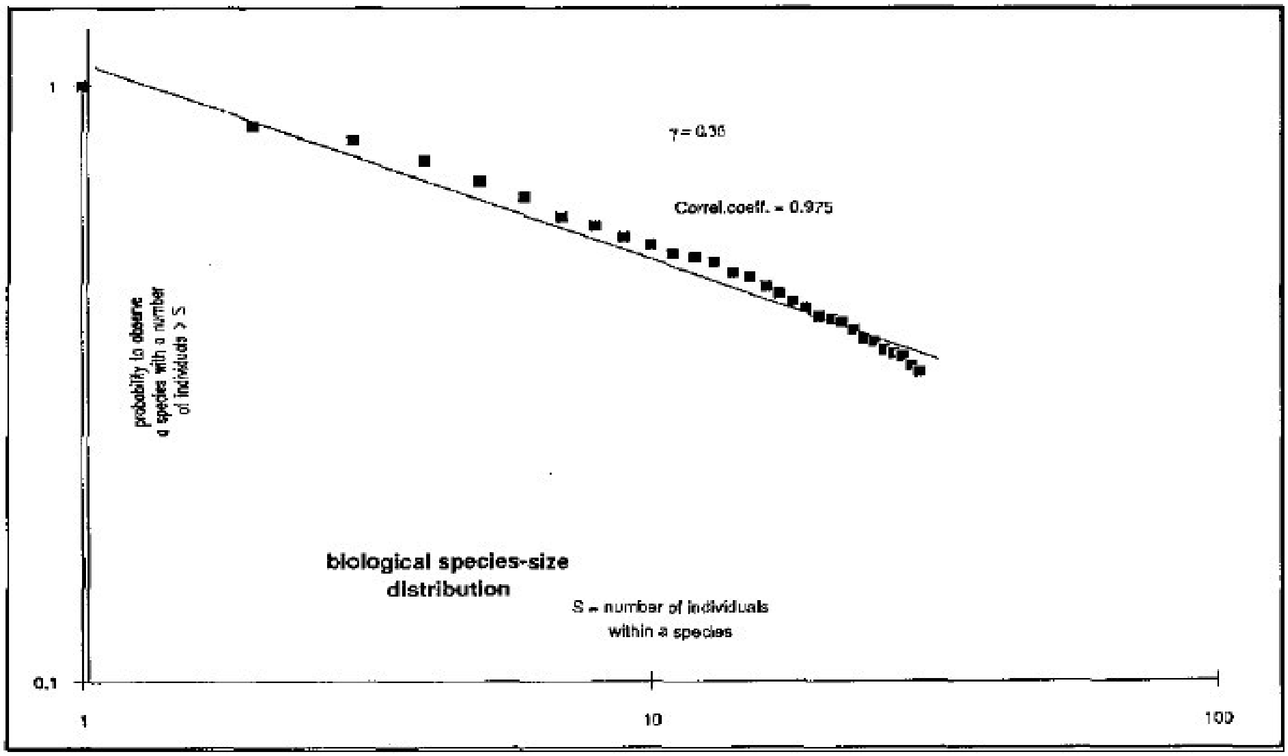


Figure. Species-size distribution of Macrolepidoptera. 15 609 individuals were captured belonging to 240 species. Data Source [Winiwarter, 1992]

Further analysis of data have shown, that similar regularities hold also for the distribution of parasites on hosts, the distribution of individuals within species and the distribution of genera within families of any observed ecosystem at any time.

1948 George Kingsley Zipf the linguist, word frequencies

In his magnum opus Zipf [Zipf, 1948] reports regularities of the above type for a wide variety of fields, but his main interest is human language for which he analyzed word-frequency distributions.
James Joyce's Ulysses is the "richest" known text with almost 30 000 words and word occurrences ranging from 1 to 2 653. The empirical data can be approximated almost too perfectly by a Pareto-Zipf distribution.
Zipf found regularities of similar type for all types of English text, for all types of languages and for all times, even for Chinese text and also for spoken language of children of different ages. The exponent is in all cases close to 1.
The only exceptions reported by Zipf are texts written by schizophrenics and scientific English.

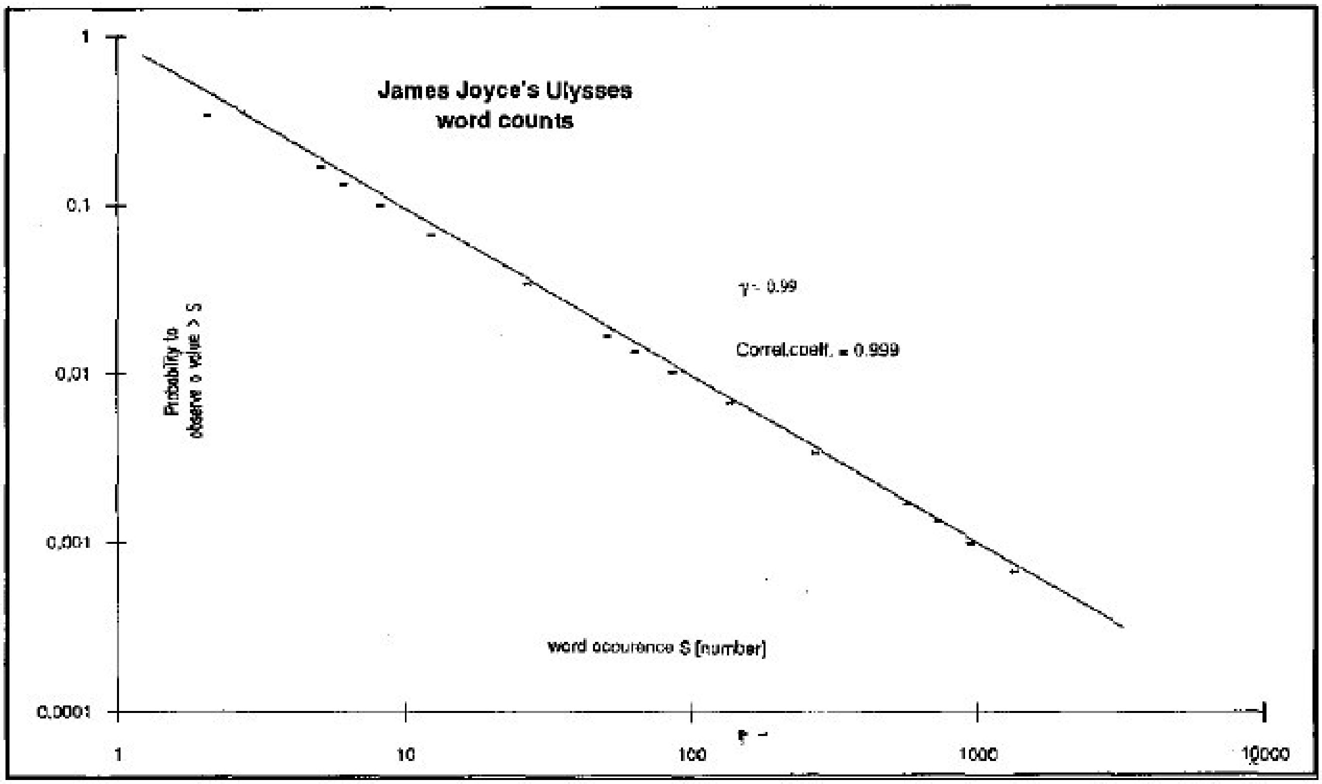


Figure. Word counts for texts in any language yield Pareto-Zipf distributions. In normalized form the graph shows the probability of a word to occur more the S times in the text. Data Source [Winiwarter, 1992]

Zipf also reports, that the distribution of scientists within a research discipline is of Pareto-Zipf type. The observed "symptom" of a scientist is measured as the number of citations in the physical or chemical abstracts.
As a side remark we note that the author [Winiwarter, 1992] has discovered that the size-distribution of programs on the hard disk of a computer are of Pareto-Zipf type.

1955 Herbert Simon, firm size distributions

Herbert Simon, who won the Nobel prize for economics in 1978, has intensively studied firm-sizes: Whether sales, assets, number of employees, value added, profits, or capitalization are used as a size measure, the observed distribution always are of the Pareto-Zipf type. This is true for the data for individual industries (economic sectors) and for all industries taken together. It holds for sizes of plants as well as of firms.
Take any annual number of the Fortune 500 magazine and you can verify this assertion, which also

holds for any national economy and also for multinational companies on a world level.
We have analyzed the Fortune data over a period of 30 years [Roehner, Winiwarter, 1984] and found, that the parameter γ of the size-distributions remains almost constant over the entire period of observation. This self-similarity of the distribution curves holds in periods of overall economic growth as well as in periods of economic recession and despite the fact, that firms appear and disappear. From the 50 largest industrial firms in 1954 only 20 can be found among the 50 largest 3 decades later, the other 30 have declined in size, been absorbed in mergers and acquisitions or simply have gone out of business . On the other hand, 12 of the 50 largest firms were not even ranked among the 500 largest in 1954 or did not even exist at that time.
To observe a constant size-distribution despite this intensive shuffling around within the system is quite remarkable .
As Herbert Simon stated in the conclusion of his paper: "We need to know more about the relations between the distributions and the **generating processes"** .
Since the graphs oft the empirical data are monotonously similar, we will not burden the reader with examples.
Over time, the Pareto-Zipf line seems to act as an ***attractor*** for "deviating points" . For example in the computer industry we had a similar situation as in the case of the largest French cities. IBM, the number one, was "too big" and the next ten following companies were "too small" deviating from the attracting straight line. The evolution of the following 10 years has brought the "deviations" almost back in line again due to:
i) a relative decline of the growth rate of IBM reducing its "deviation"
ii) an above average growth rate of DEC, the number two follower bringing it closer to the attractor
iii) several mergers and acquisitions among the top computer companies reducing the overall deviations.

1956 Gutenberg-Richter the distribution of earthquakes

In 1956, the geologists Beno Gutenberg and Charles Richter (the father of the seismological scale of the same name) discovered [Richter, 1958], that the number of important earthquakes is linked to the number of small earthquakes :
the law of Gutenberg-Richter states, that the number of annual earthquakes as a function of the liberated Energy, is a Pareto-Zipf-distribution . The exponent $\gamma = 1.5$ is universal and does not depend on the geographical region!

1983 Winiwarter, the distribution of chemical elements in cosmic systems

The analysis of chemical element distributions within stars or within the entire cosmos is traditionally presented as relative abundance versus the mass number of the elements.

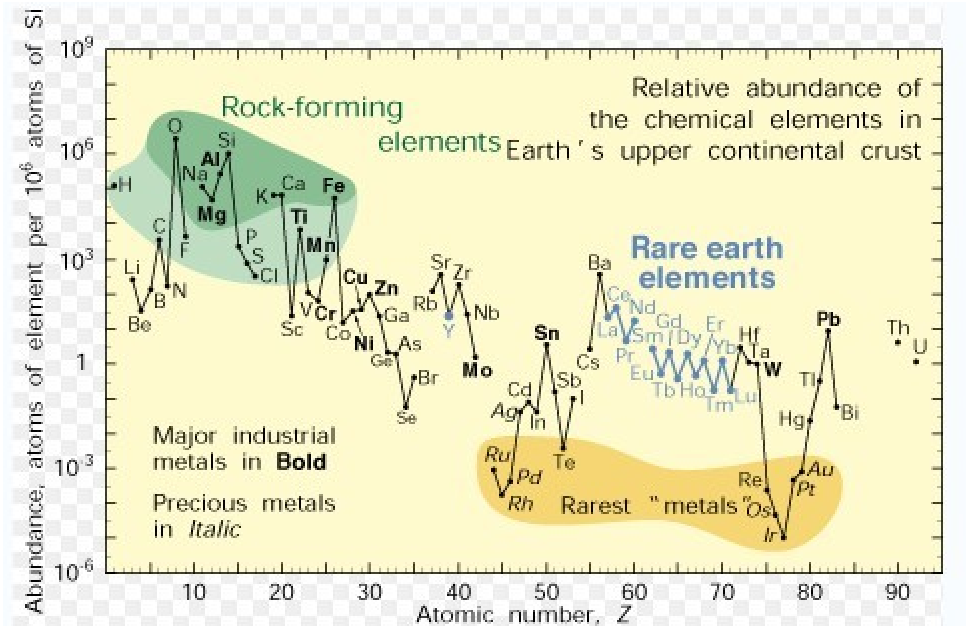


Figure. Relative abundance of chemical elements in the universe as a function of atomic mass. This graph does not allow to deduce any quantitative regularity except a decrease of abundance with mass number with peaks around the "magic numbers". Data Source [Wikipedia]

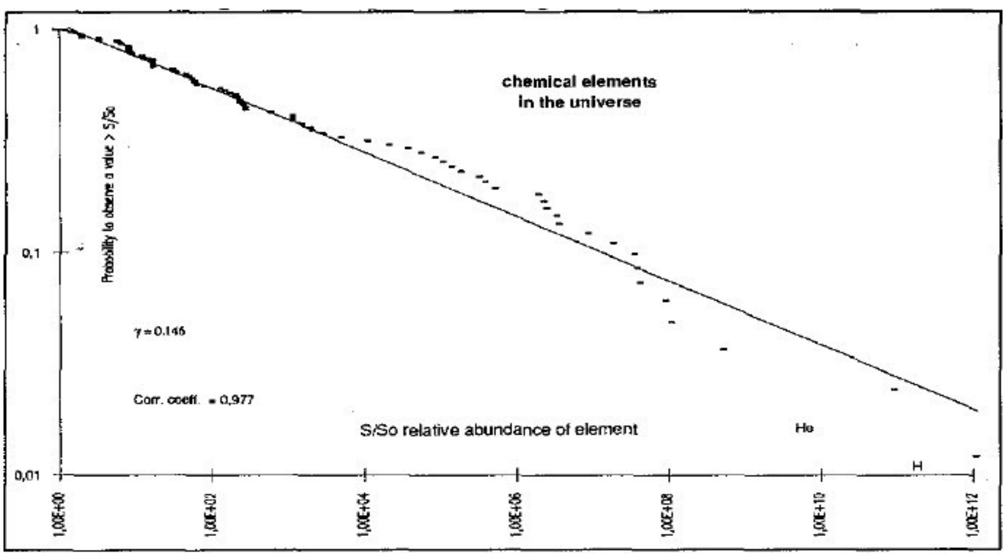


Figure. The same data as in the figure above presented as a normalized Pareto-Zipf distribution revealing a distinct quantitative regularity. Data Source [Winiwarter, 1992]

This type of regularity for the abundance of chemical elements can be observed for the universe, for single stars, for meteorites, for the lithosphere ...
Similar regularities can be observed for star-size distributions in galaxies, for the planet-size distribution in our solar system, for the moon-size distributions of the Jupiter system ...

1991 Cempel, the distribution of vibration amplitudes in mechanical machine systems

Research in the field of vibration diagnostics [Cempel, 1991] has revealed, that long-tailed Pareto-like distributions are a good approximation for the data yielded by empirical measurements of vibration symptoms for a set of "running" machines :
the regularities are observed independently of the machine type (electro motors, diesel engines ...)

Do we live in a Pareto-Zipf world?

This short historical overview showed the discovery of similar regularities for incomes, cities, species, words, earthquakes, chemical elements, machine vibrations. How can this possibly make sense without postulating similar underlying structures and processes of the observed systems?
With the rapid development of complex network theory in the 1990ties Power laws have been observed in almost all systems of research ranging from protein networks to the World Wide Web.
In the following chapter we will give some illustrated examples of the systems, for which we observe Pareto-Zipf-Mandelbrot Power laws.

Pareto-Zipf-Mandelbrot (PZM) and parabolic fractal distributions

There exists a great variety of names for the same type of empirically observed distributions in self-organized systems:

long tail , "*longtailed* "/ "*heavy tailed* "/ "*skewed*" distributions, Pareto law, Zipf's law, Zipf-Mandelbrot law, lognormal distribution , Yule-Simon distribution, Frechet Weibull distribution, rank-size rule, parabolic fractal distribution, 80/20 rule, the law of the vital few ***and the principle of factor sparsity*** ...law of Gutenberg-Richter, Lotka's law, Bradford's law , Benford's law ... selforganized critically ***power laws***, scaling laws, scalefree networks, ...

all are synonyms of the same statistical **power law structur**e called **PZM** (***Pareto-Zipf-Mandelbrot***).

The common statistical feature of all the distribution types cited above are simple, they yield more of less straight lines in log log coordinates.
The mathematical forms of the distributions are more or less complicated. Statisticians have done extensive studies http://arxiv.org/PS_cache/arxiv/pdf/0706/0706.1062v1.pdf, trying to find out which distribution yields the best fit to a given data set. But they show, that if one distribution yields a good

fit, then all the other distributions yield good fits also. (Error type three in the inquiry question, which is not what distribution is best, but why do we observe always similar distributions).
Let's apply Occam's razor and say that the most simple distribution will do it (the simple Pareto power law, which is equivalent to Zipf's law or the rank size rule by simple inversion of coordinates. For discrete distributions (which are the case in most real examples) the Zipf-Mandelbrot or parabolic fractal distribution is the most simple form to prefer to complex constructs like Yule-Simon or Frechet Weibull.
In the following we therefore speak of PZM (**Pareto-Zipf-Mandelbrot or parabolic fractal distribution**).

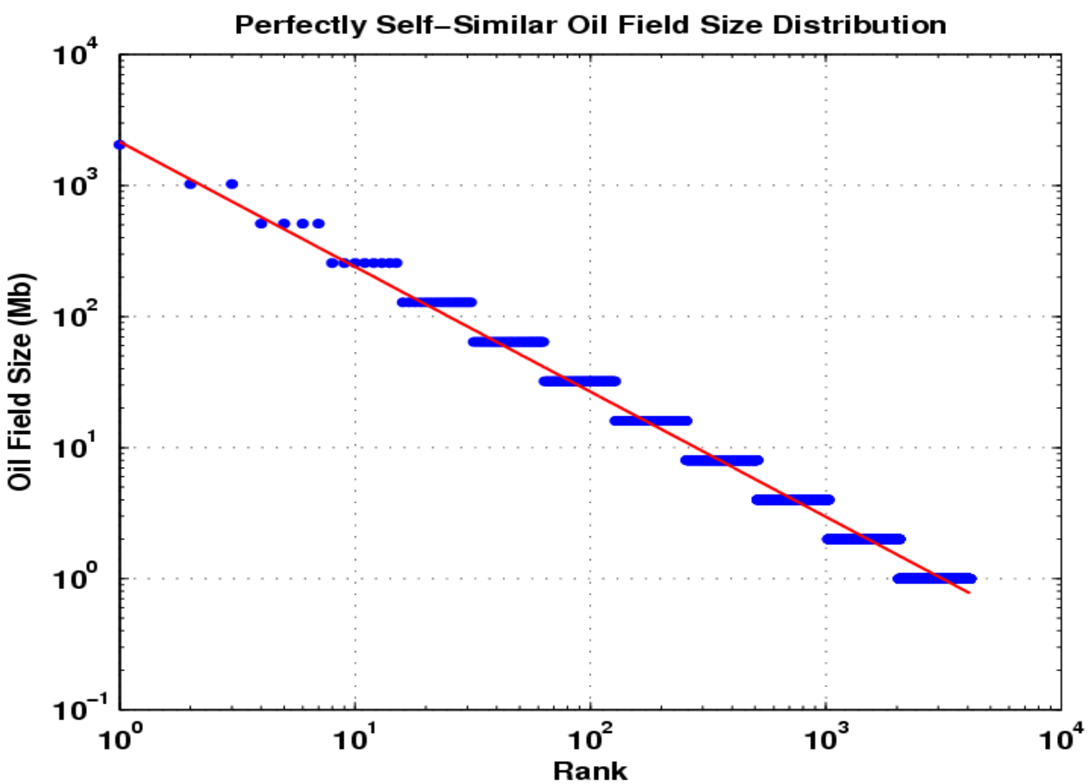


Figure. Parabolic fractal distribution, the logarithm of the frequency or size of entities in a population is a quadratic polynomial of the logarithm of the rank.

How to explain this to a non-mathematician?
Very simple:

Let's take the income size distribution of any country. There are very few very rich billionaires, there are few rich millionaires, there are many middle class people and the remaining vast majority are just plain poor. This inequality can be described by a mathematical distribution, which yields a straight line in log-log coordinates. (Pareto law).

Another example for PZM Pareto-Zipf-Mandelbrot regularity. Let's take the city size distribution of a country and rank the cities in decreasing order of number of inhabitants. There are a few very big metropolis, there are a few big cities, there are many cities of medium size and the vast number of agglomerations are small towns. This inequality can be described by a mathematical distribution, which yields a straight line in log-log coordinates. (rank size rule). etc...

This asymmetric distribution is very different from the well known Gaussian bell shaped symmetric distribution.

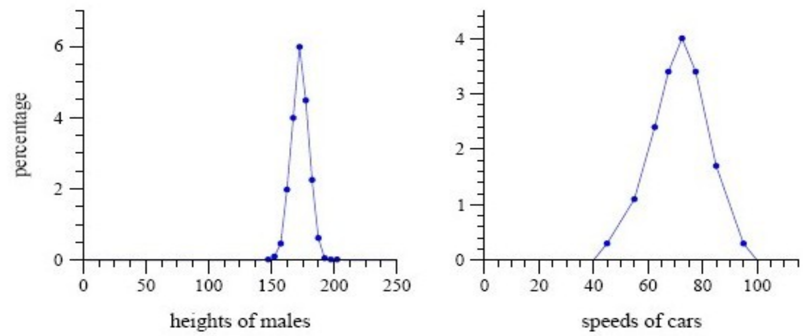


FIG. 1 Left: histogram of heights in centimetres of American males. Data from the National Health Examination Survey, 1959–1962 (US Department of Health and Human Services). Right: histogram of speeds in miles per hour of cars on UK motorways. Data from Transport Statistics 2003 (UK Department for Transport).

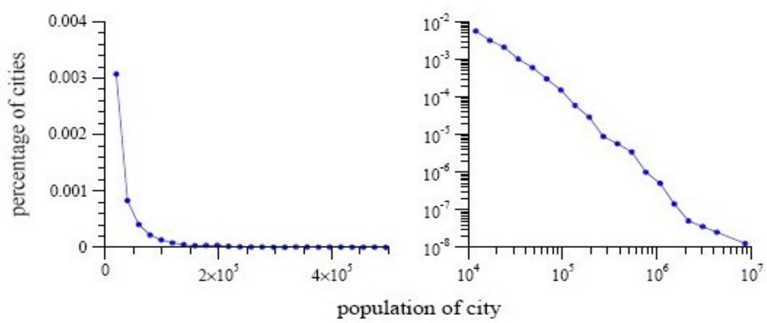


FIG. 2 Left: histogram of the populations of all US cities with population of 10 000 or more. Right: another histogram of the same data, but plotted on logarithmic scales. The approximate straight-line form of the histogram in the right panel implies that the distribution follows a power law. Data from the 2000 US Census.

normal Gaussian distribution versus Pareto-Zipf distribution

For those who persist to say they are the same mathematical structures, there is a major difference even in the second and third degree of a Taylor development.

Pareto-Zipf-Mandelbrot (parabolic fractal) distributions are scalefree.

In the PZM or parabolic fractal distribution the right tail of the poor is very much longer than the short left tail of the rich. Therefore the term longtailed. For these distributions there is no such thing calculable like a mean or average income, since there is no symmetry and a value would be different for every arbitrary cutoff point in the ranking.

Note that there are only two families of mathematical distributions which do not change their form after the merger or split of system distributions :

1) the Gaussian *(Gauss folded with Gauss yields Gauss).*

2) the Pareto-Zipf-Mandelbrot distribution *(Pareto folded with Pareto yields Pareto). see the chapter on stability under addition.*

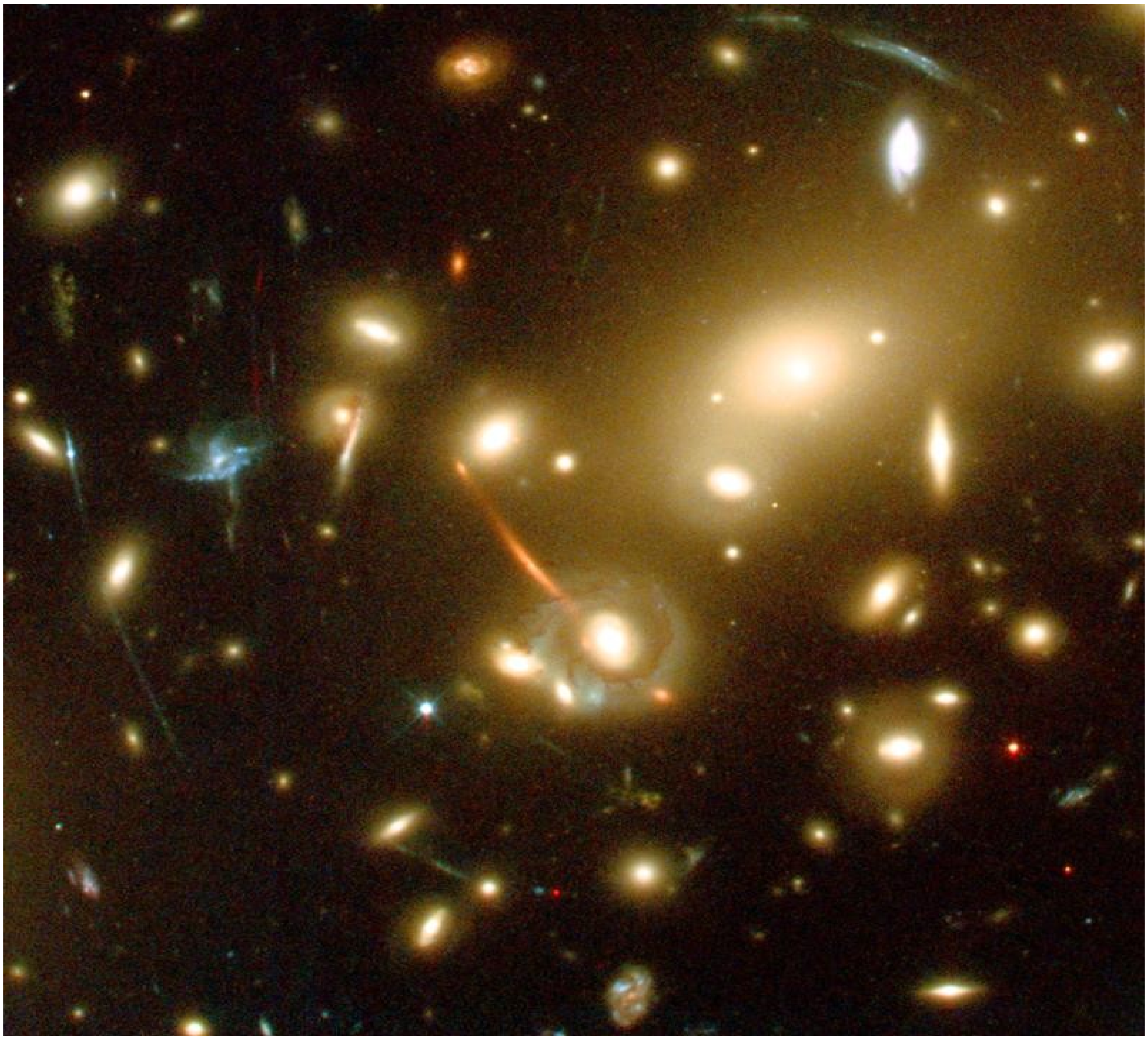
For all the following examples in the next chapter we observe similar regularities of the Pareto-Zipf-Mandelbrot (parabolic fractal) type in the fields of:

astrophysics, geophysics, geology, geoscience, physics, meteorology, physico-chemistry, biochemistry, biology, plant, animal, ecosystems, environment, social systems, transportation systems, economics, sociology, religion, linguistics, mechanical systems, computer technology, world wide web, scientometrics, brain, neural networks ...

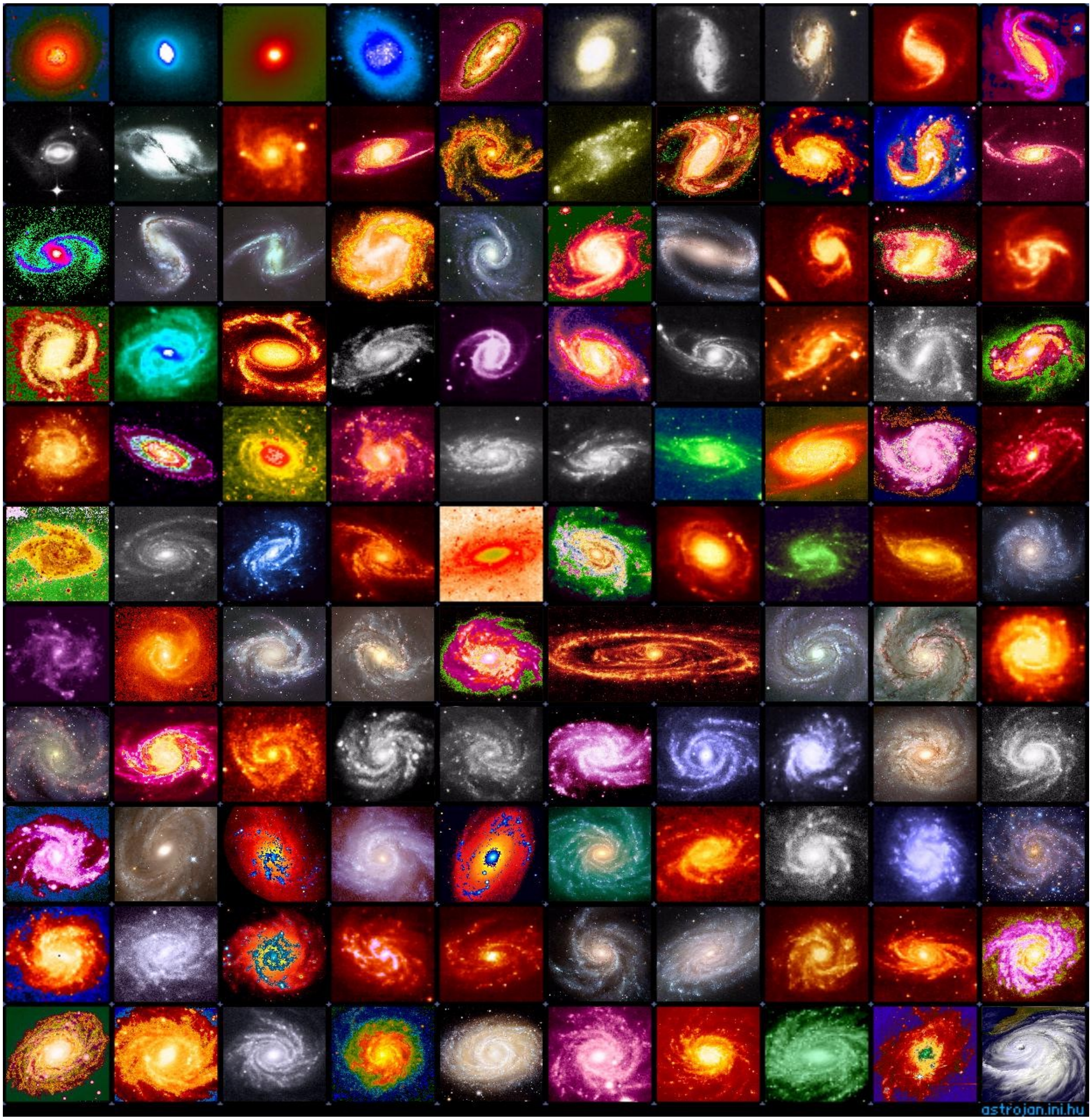
Illustrated regularities of the Pareto-Zipf-Mandelbrot type

Data Source: Google Images

Astrophysics, Nuclear networks

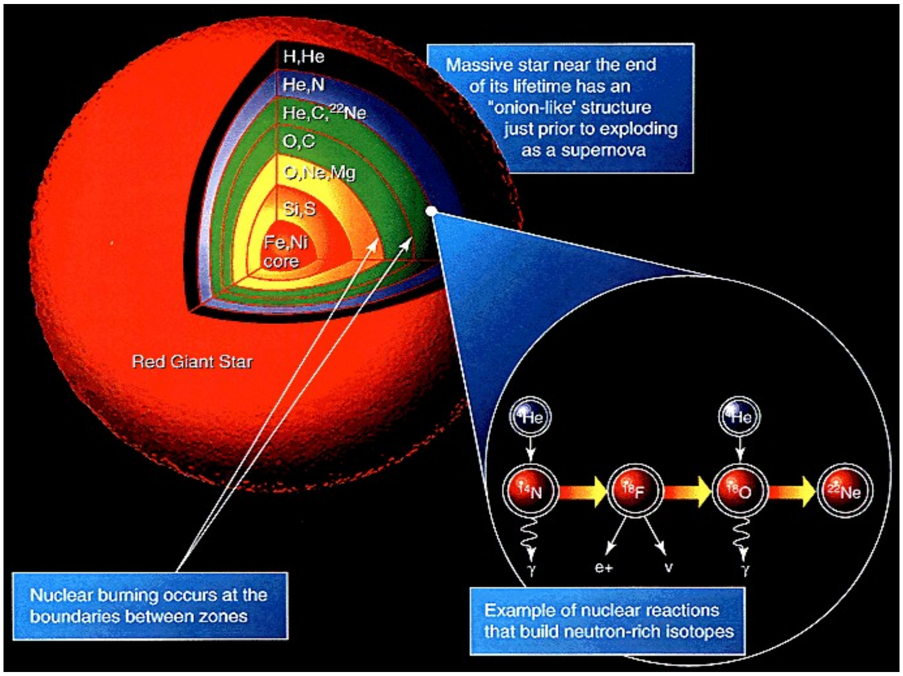


Universe network: PZM distribution of galaxy cluster size



Universe network: PZM distribution of galaxy sizes

Galaxy network: PZM distribution of star size



Massive star nuclear network : PZM distribution of chemical element frequencies from Hydrogen to Uranium

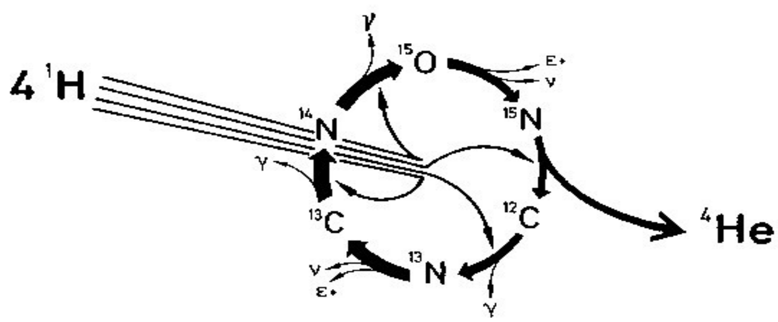
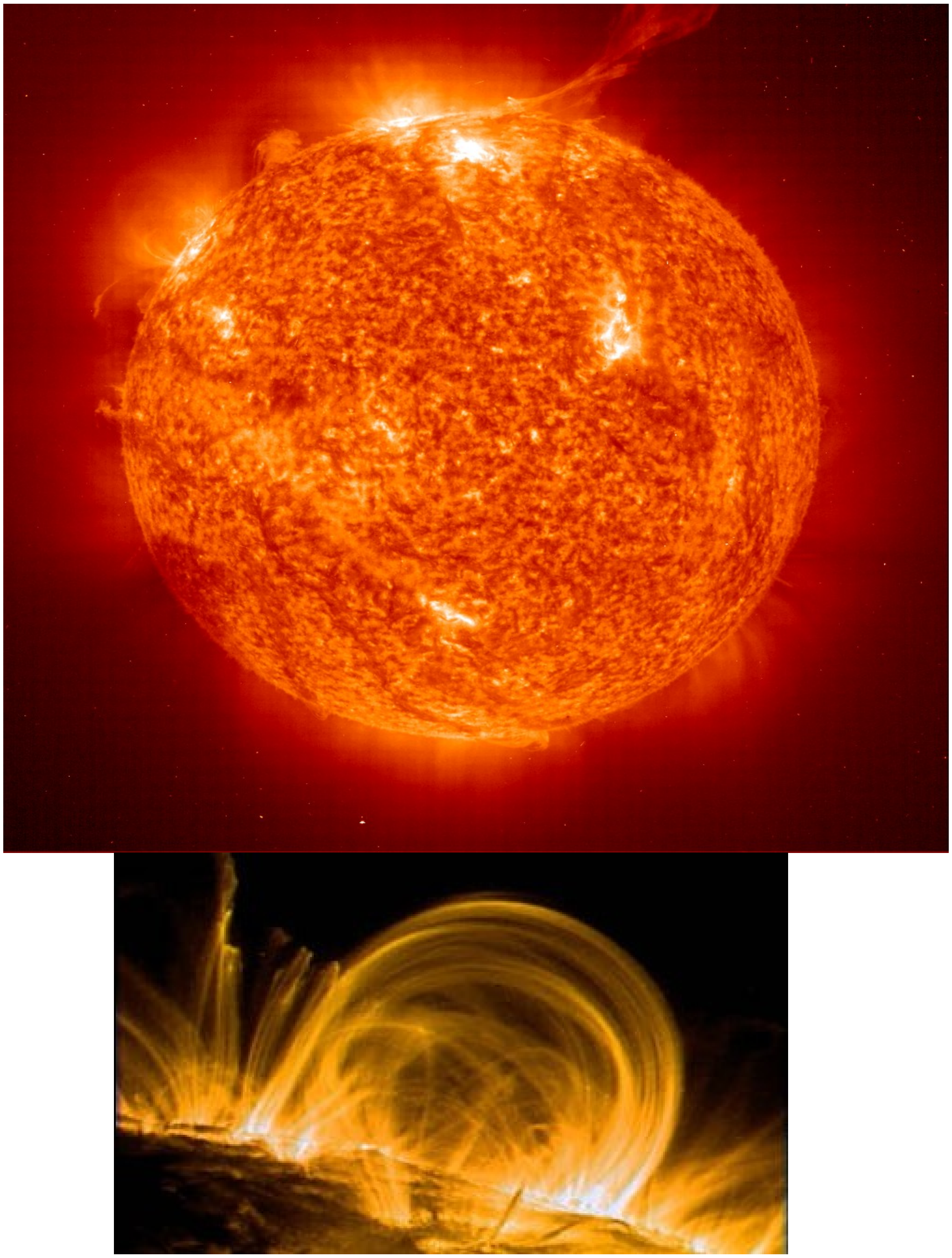
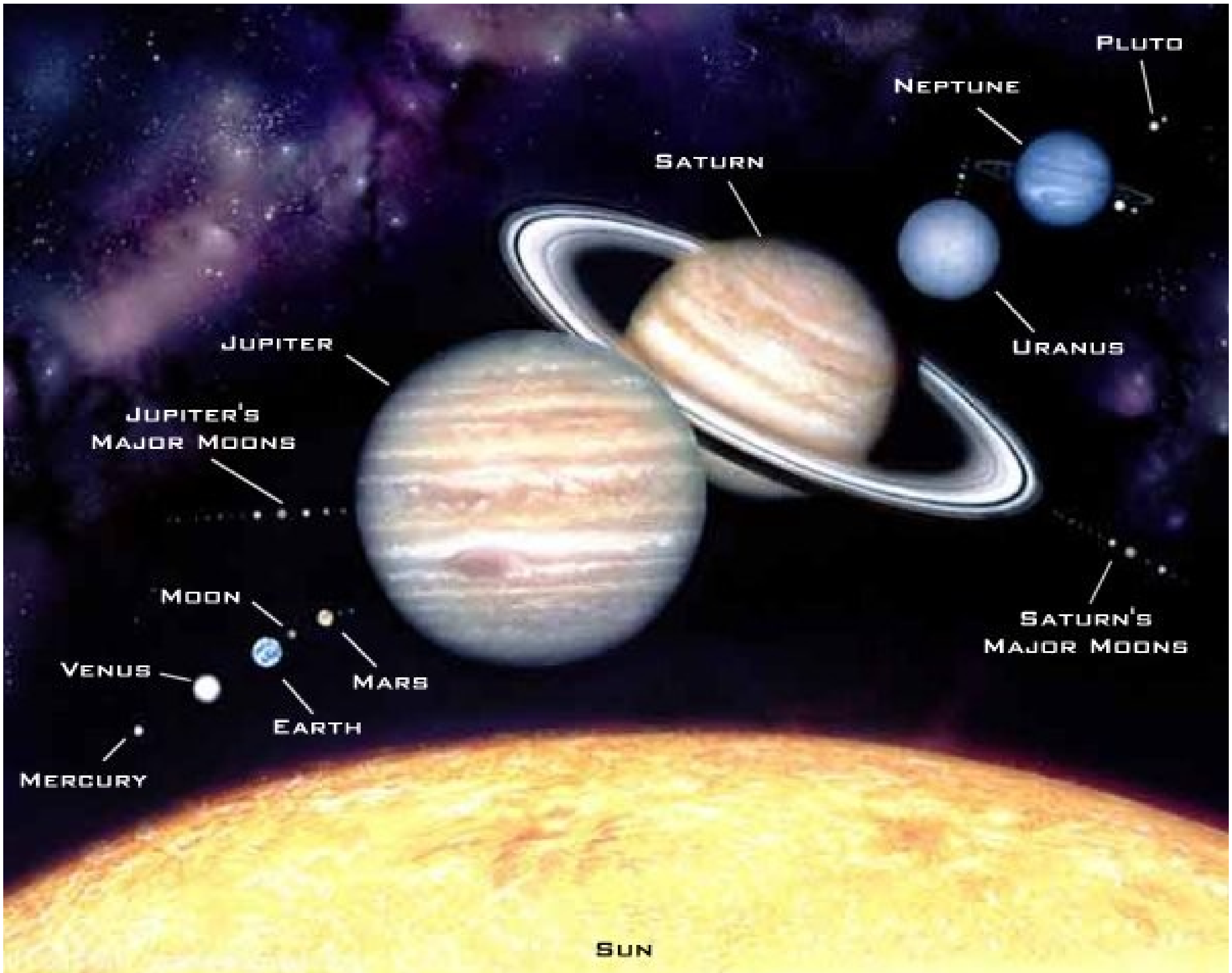


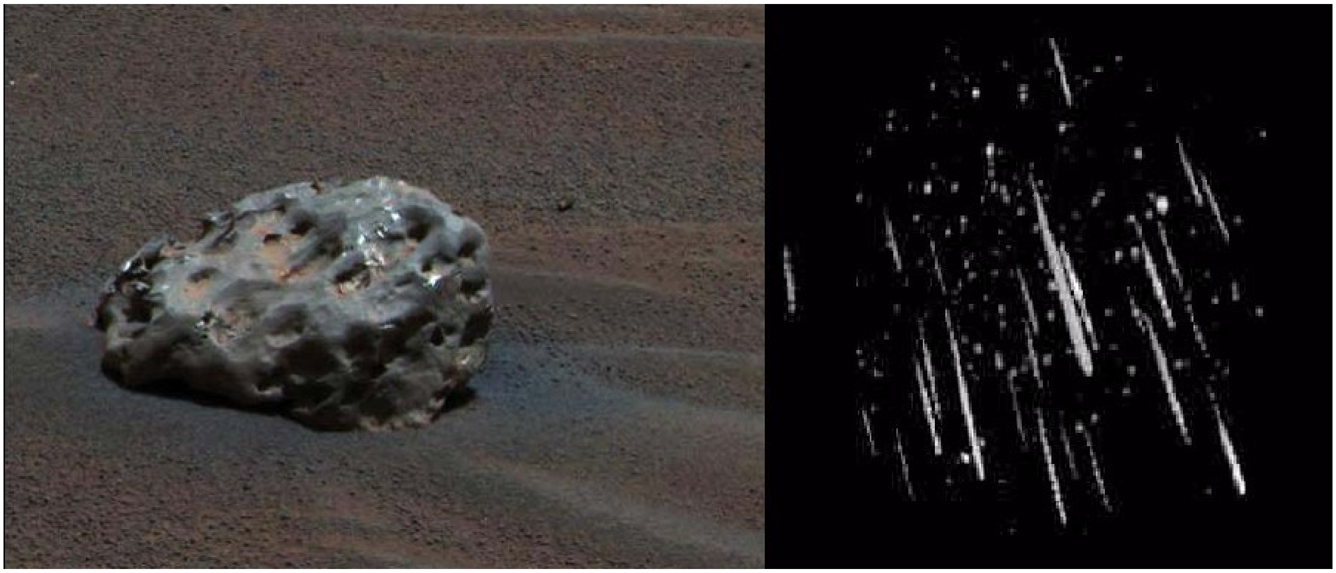
Fig. 2. *The carbon cycle,* proposed by Bethe and v. Weizsäcker, is responsible — at least in part — for the energy production of massive stars. The constituents: ^{12}C, ^{13}N, ^{13}C, ^{14}N, ^{15}O, and ^{15}N are steadily reconstituted by the cyclic reaction. The cyclic scheme as a whole represents a catalyst which converts four ^{1}H atoms to one ^{4}He atom, with the release of energy in the form of γ-quanta, positrons (ε^{+}) and neutrinos (ν).



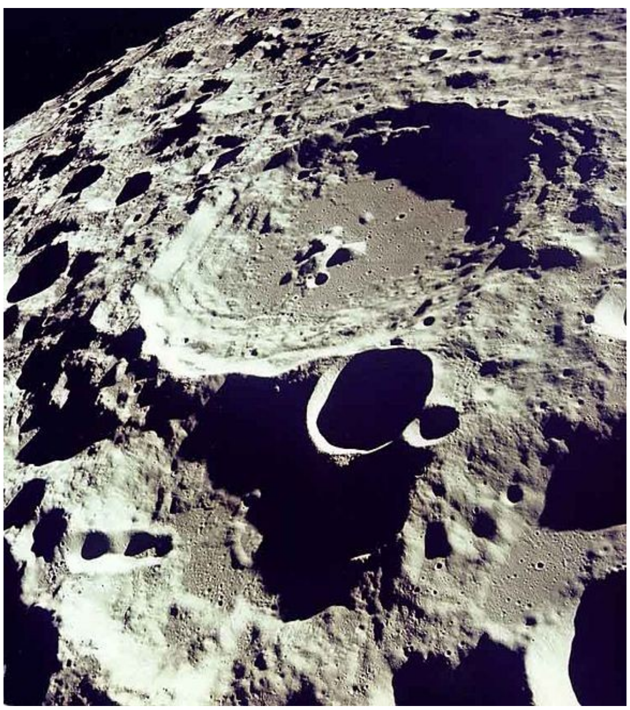
Sun network: solar flares reveal PZM distribution



Solar Planetary network: planet size, satellite of planet (moon) size distributions are of the PZM Pareto-Zipf-Mandelbrot type (parabolic fractal)

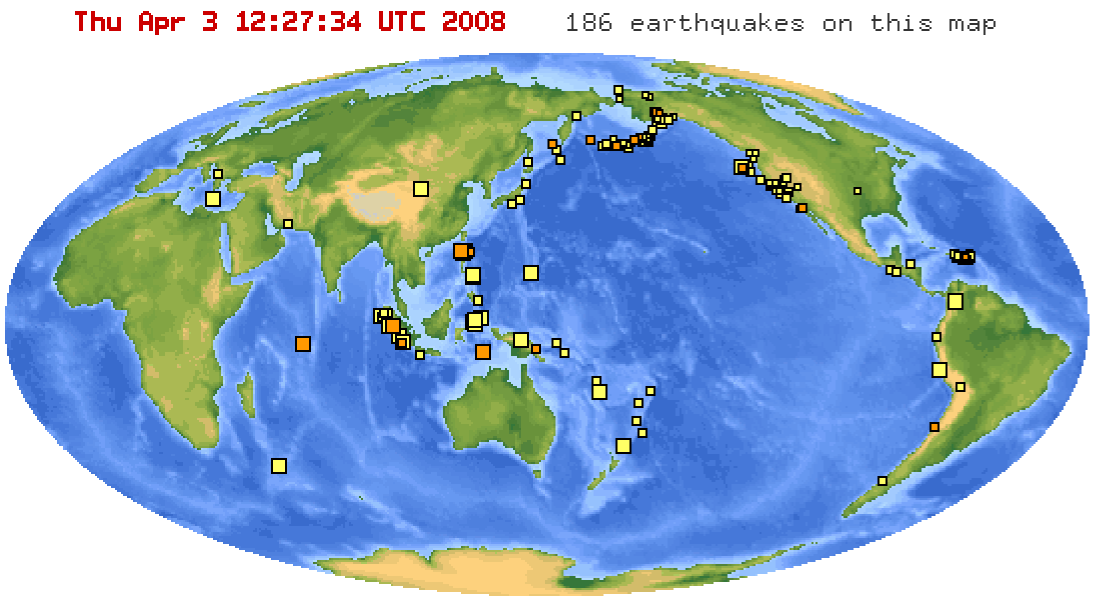


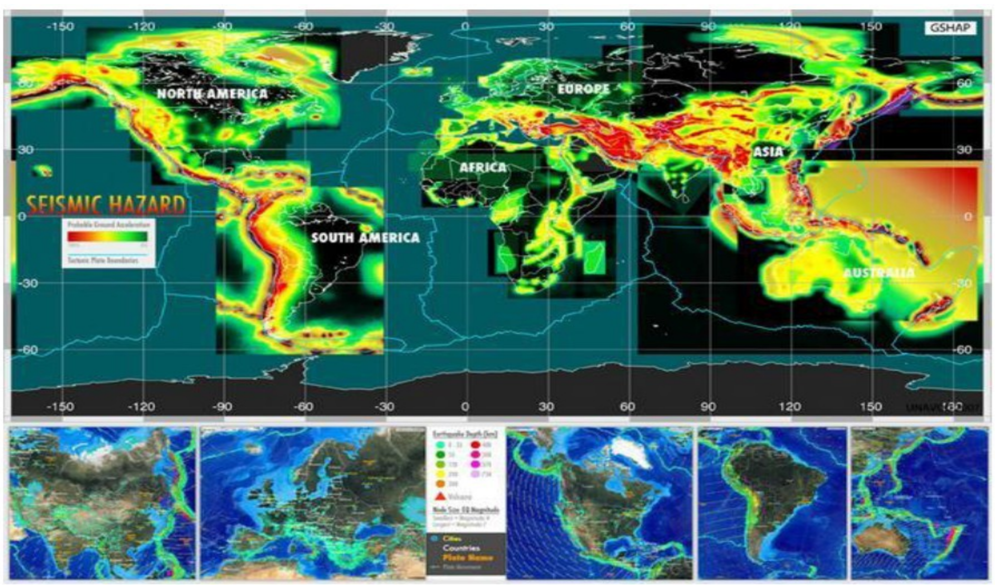
Satellite network: Meteorites show PZM distribution



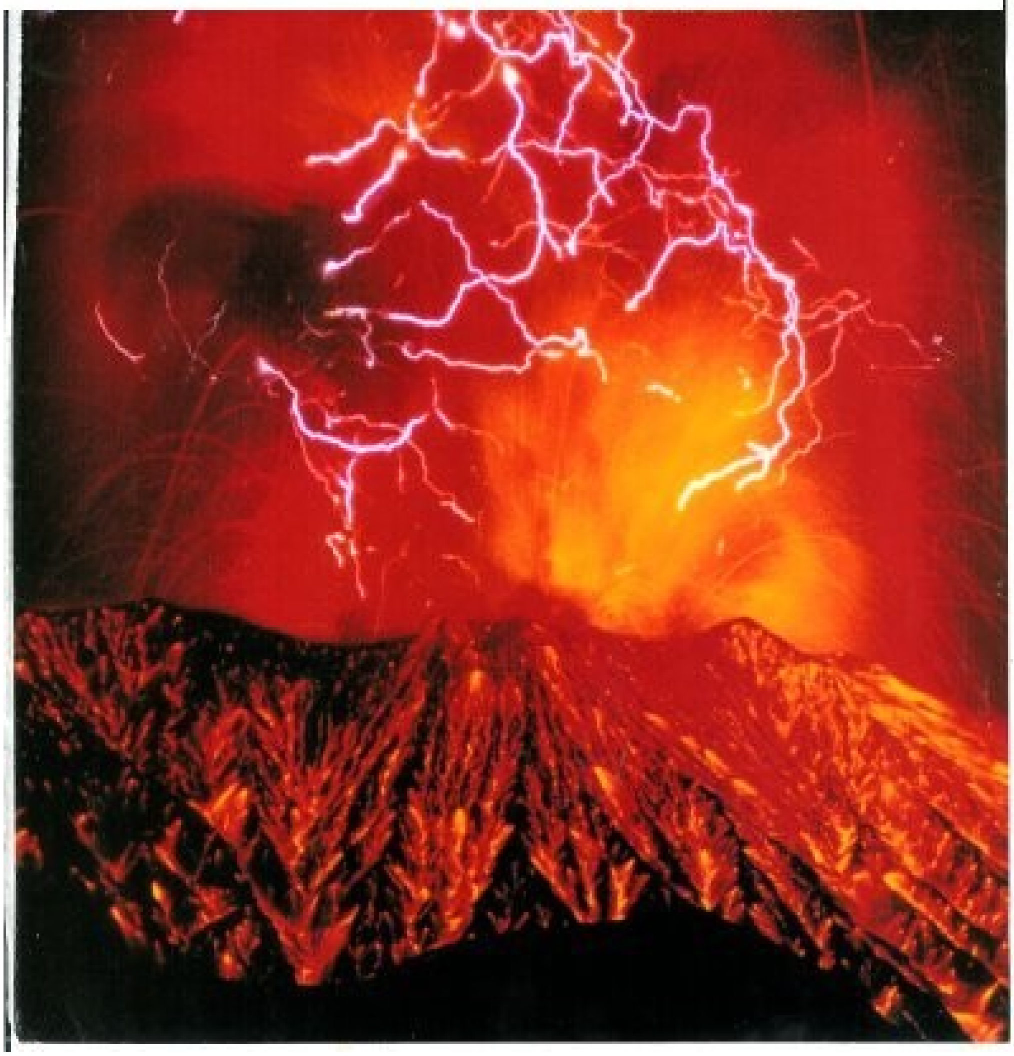
Moon surface network : craters reveal PZM distribution

Geophysics (Gaia), Tectonic networks

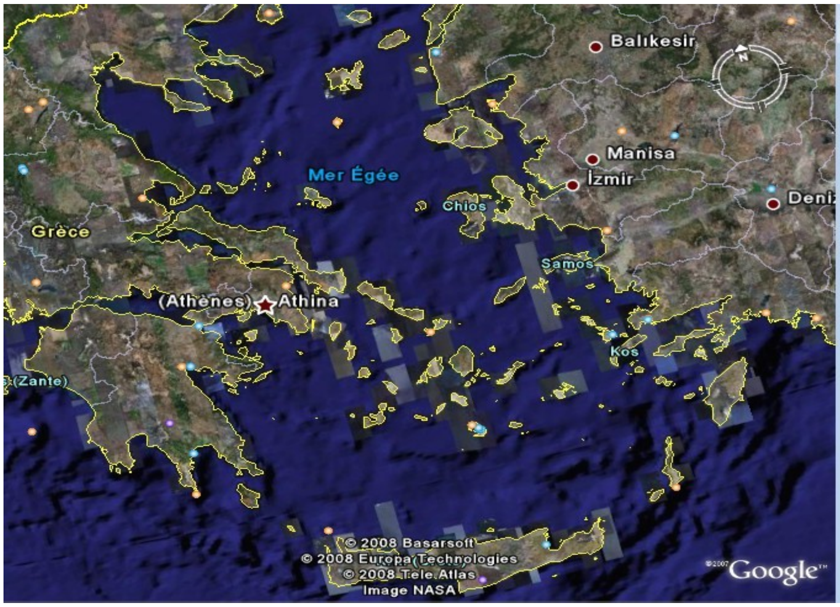




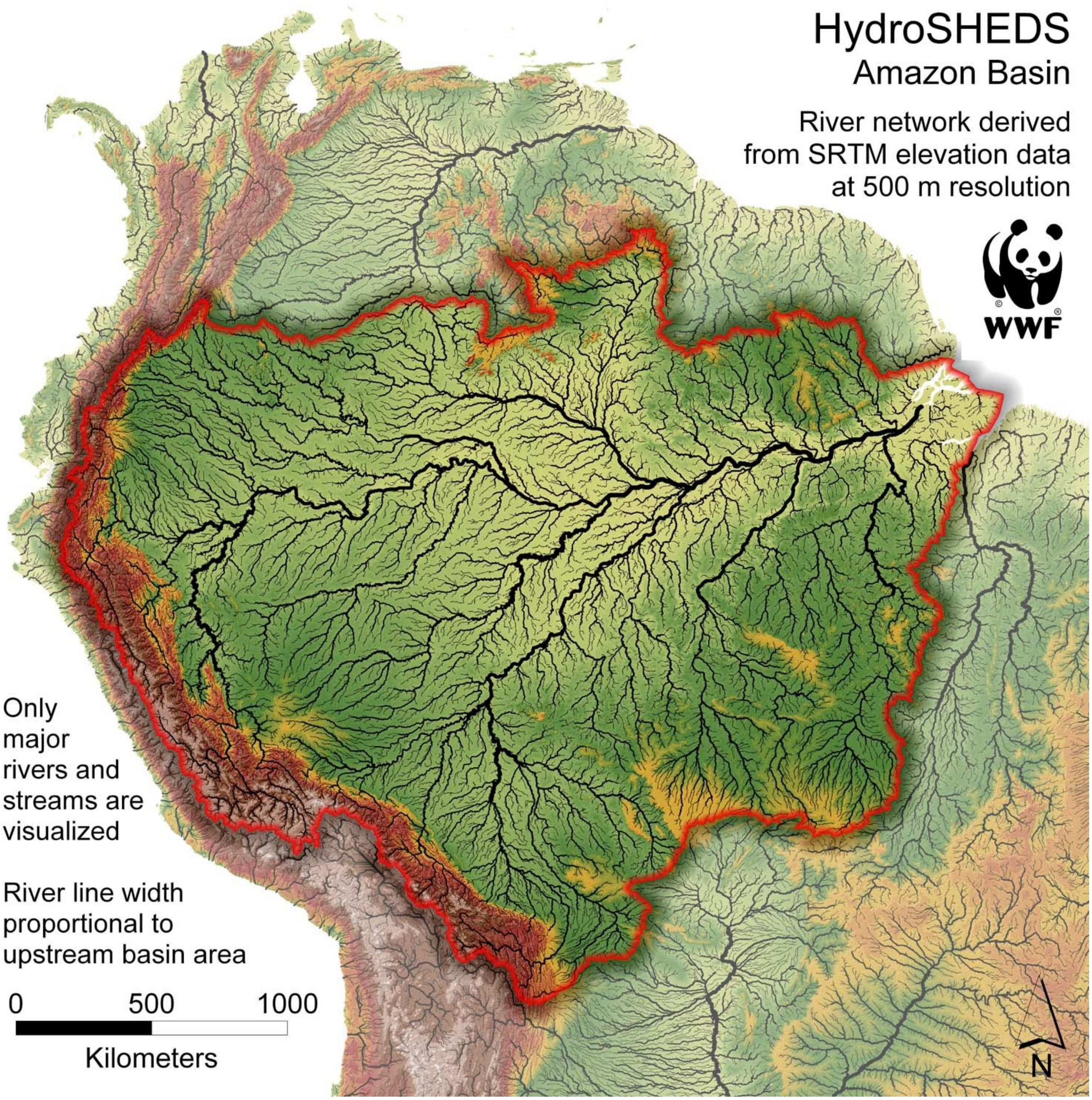
Tectonic Networks: PZM distributions of earthquake energy size are observed for all regions of the globe



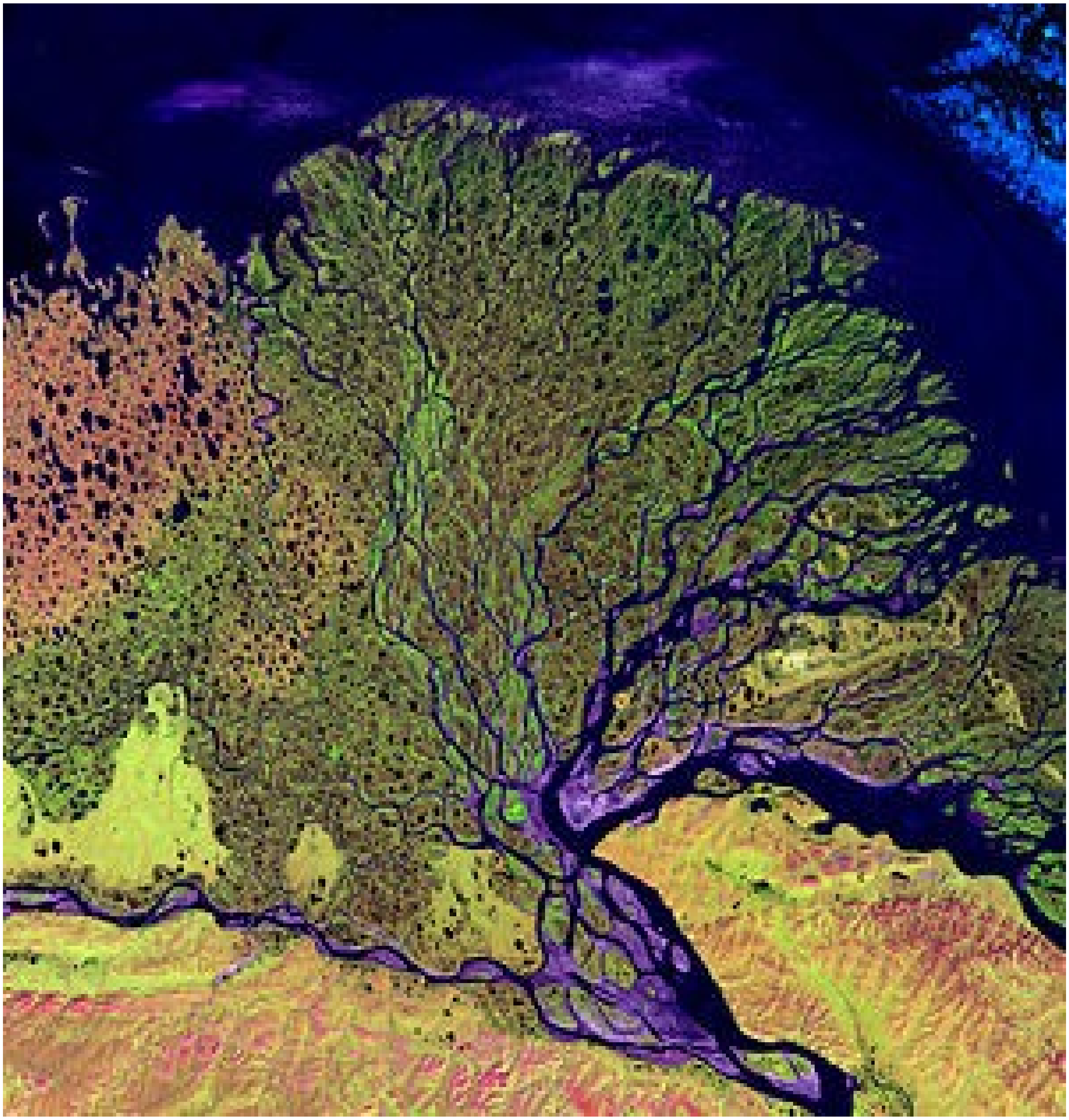
Geothermal network: volcanic eruption sizes show PZM distributions



Tectonic network: island size distributions are of the PZM type



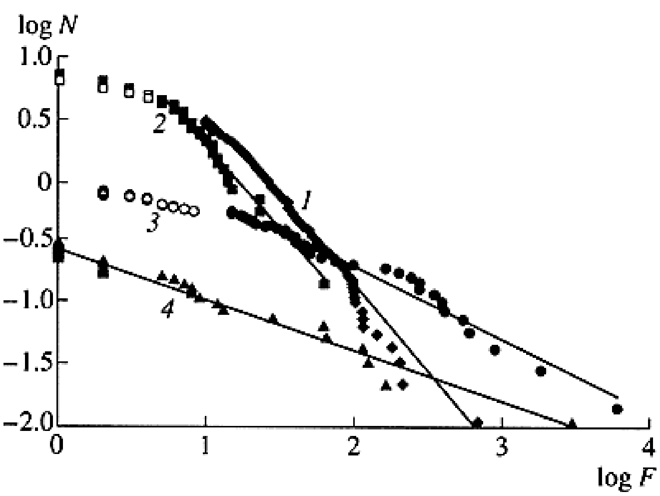
Water network: river size distributions are of the PZM type within a river basin



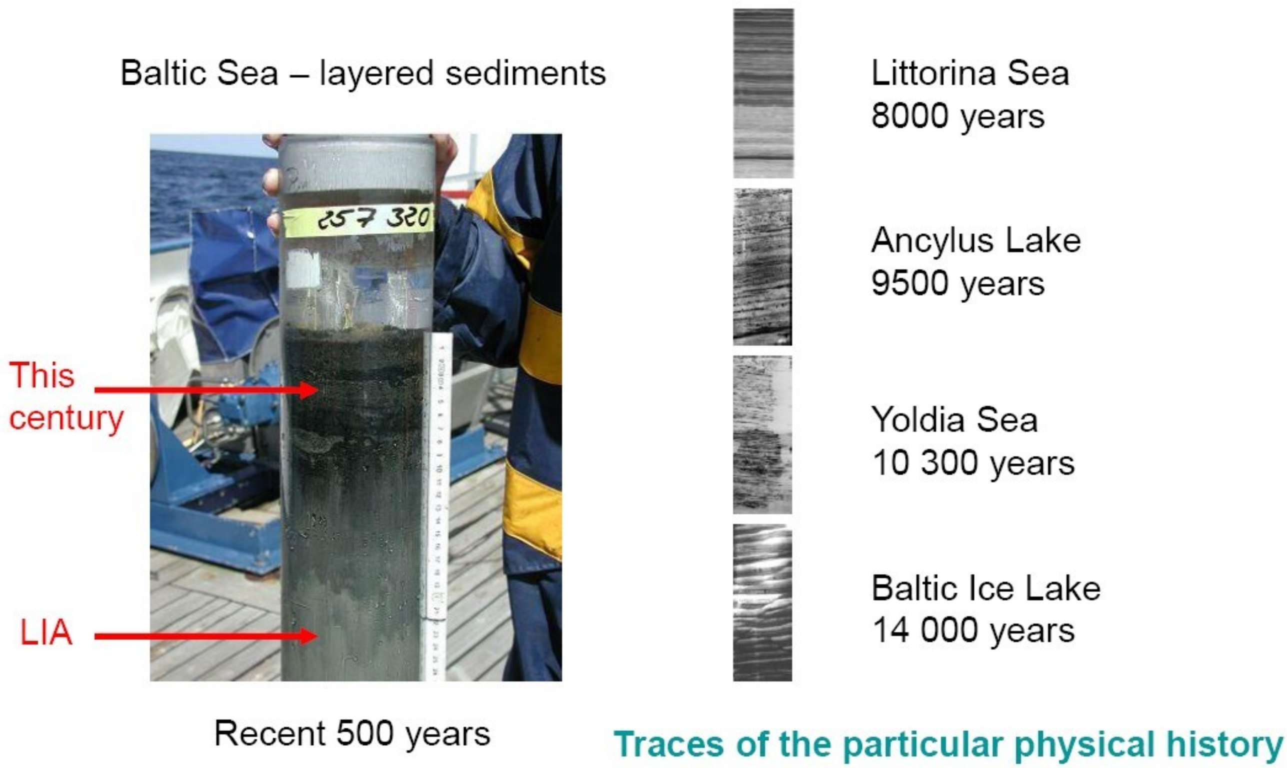
Water network: River Delta PZM distribution of river size

Water network: flood size and lake size distributions are of the PZM type

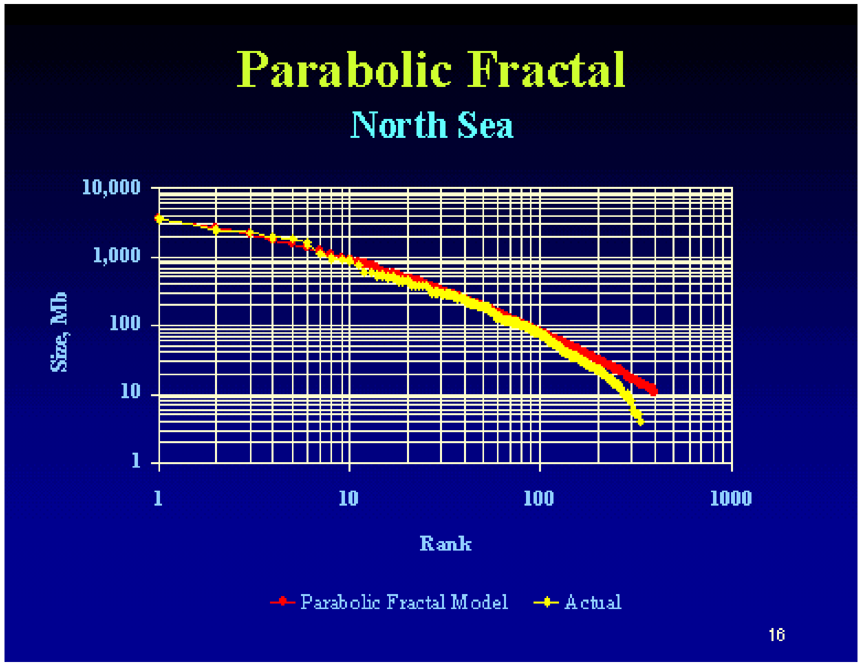
Atmosphere network: hurricane energy size distributions are of PZM type



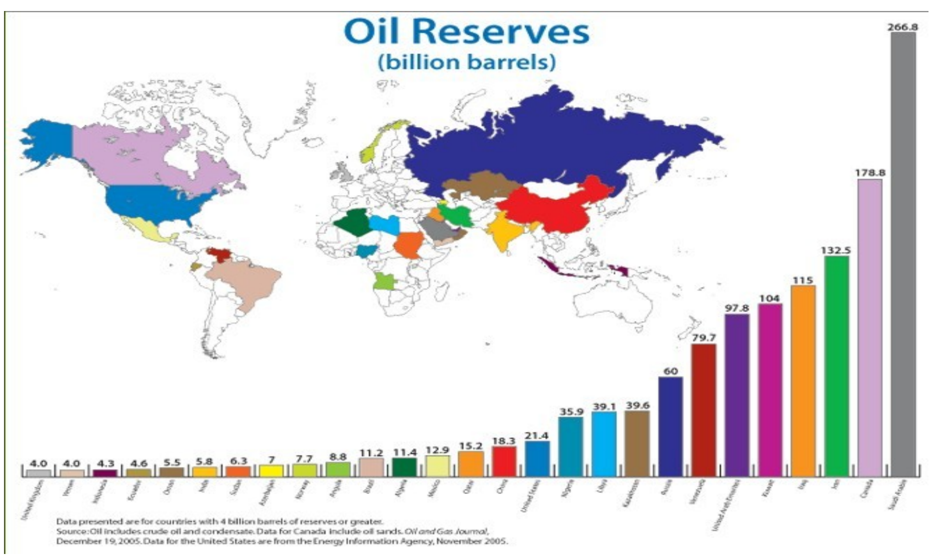
The fatality distribution of tornadoes (1), floods (2), hurricanes (3), earthquakes (4) in the 20th century in the United States show PZM regularity



Sediment network: cosmic and terrestrial dust size distributions are of the PZM type



Energy network: field size distributions of oil reserves (geologically transformed vegetation networks) are of PZM type



Prebiotic chemical networks (Hypercycles)

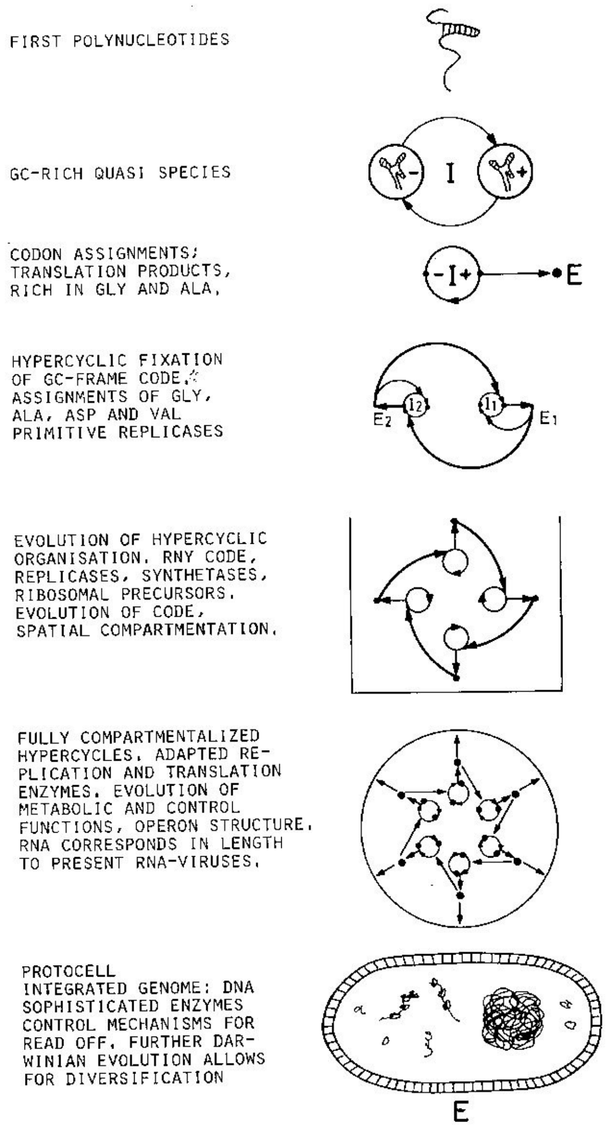
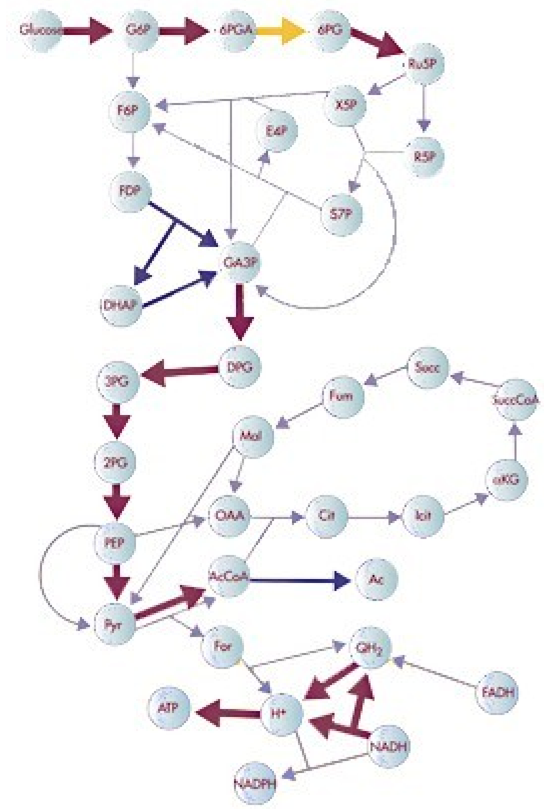


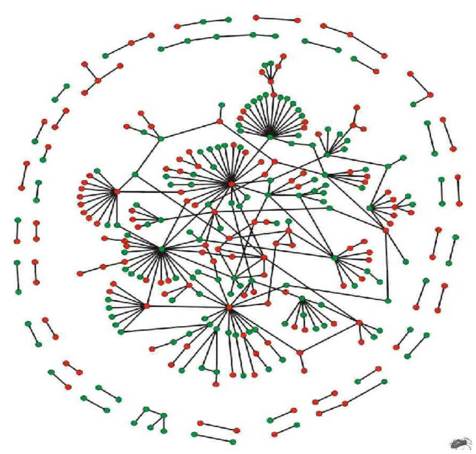
Fig. 63. Hypothetical scheme of evolution from single macromolecules to integrated cell structures

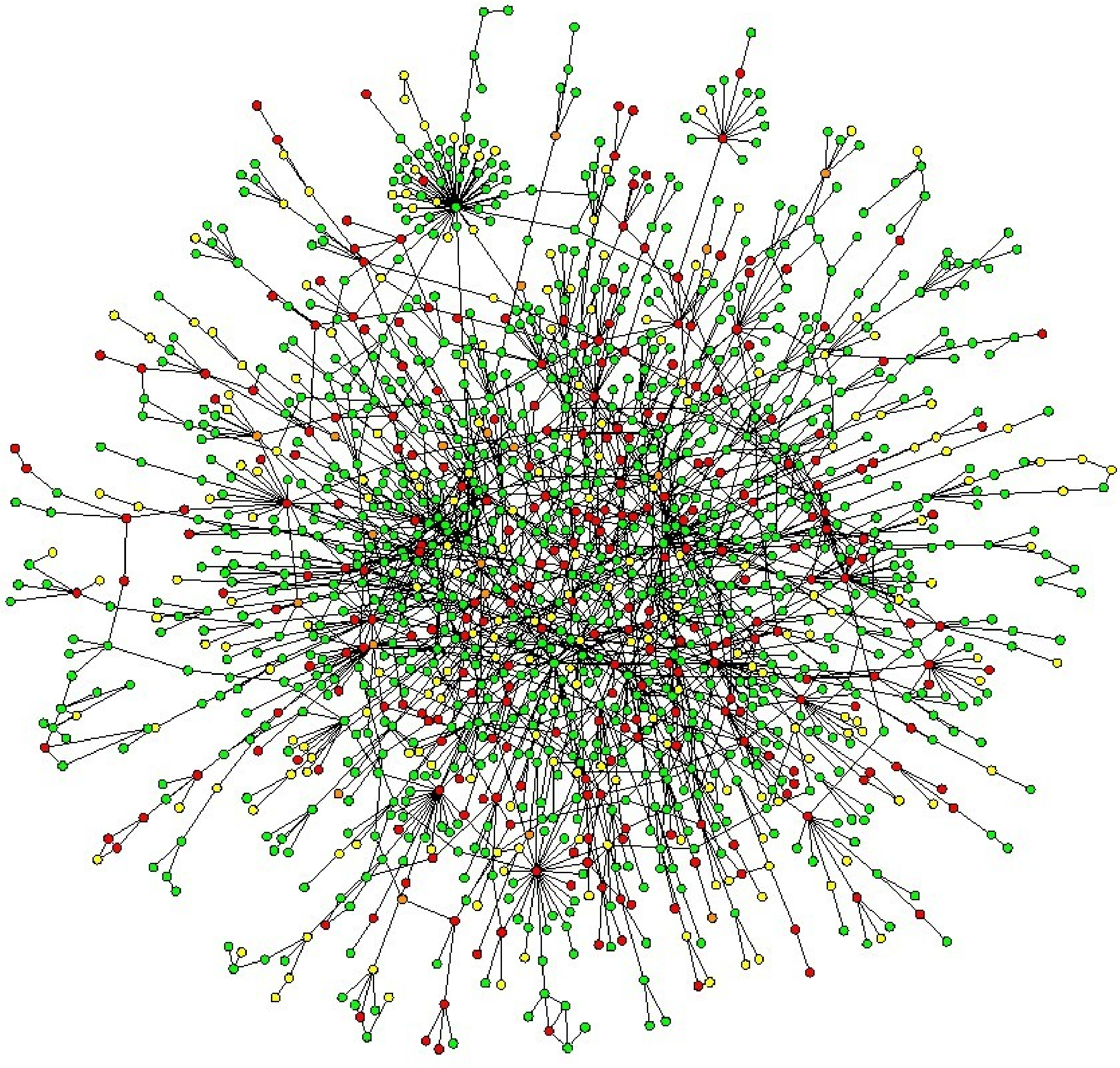
Chemical network: compounds of hypercycles show PZM distribution

Biophysics, Biochemistry : protein and metabolic networks



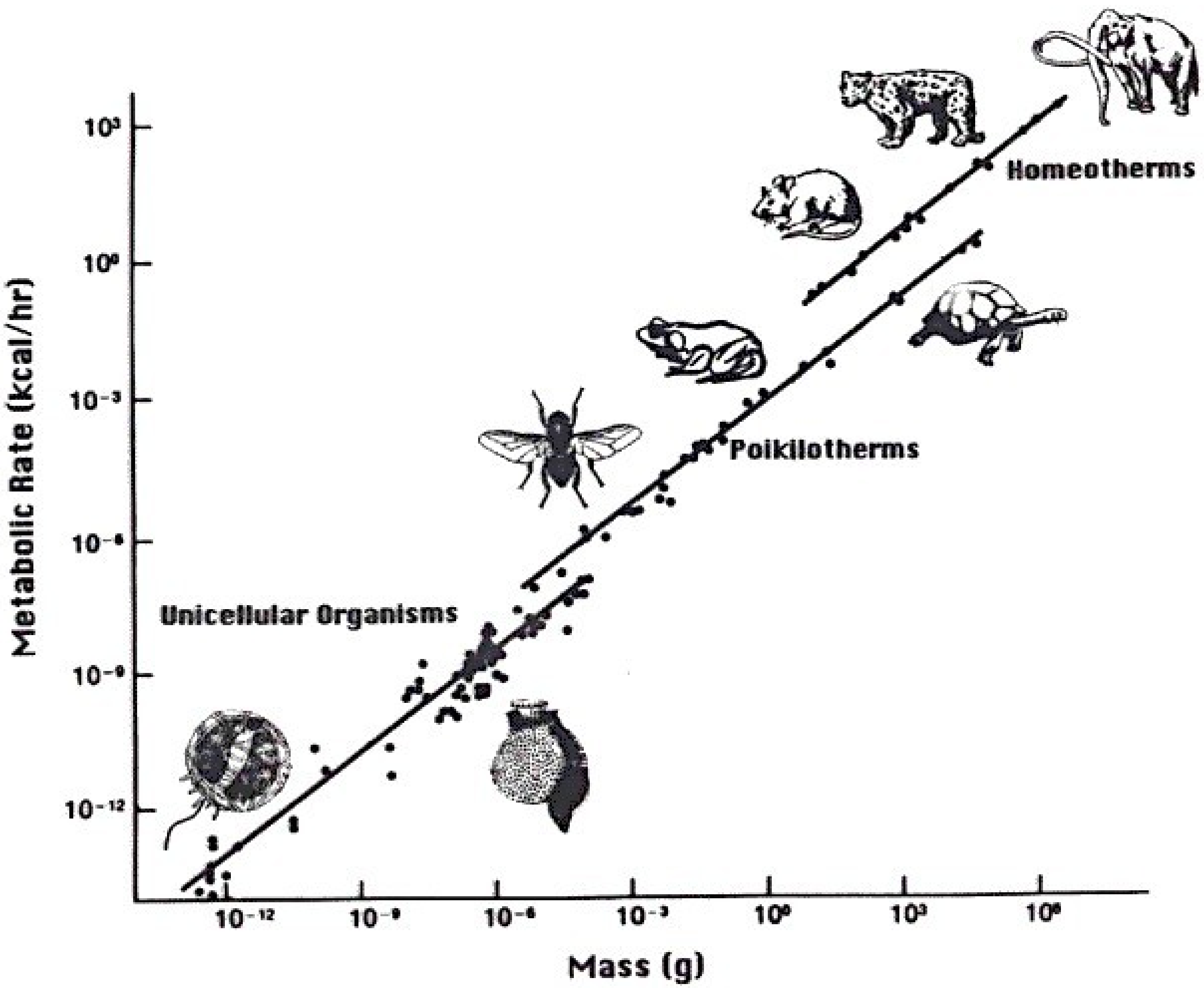
Metabolic network: the degree distribution of E. coli metabolic network is of PZM type





The yeast protein interaction network has a scalefree topology (PZM distribution) The scale-free nature of protein interaction networks is supposed to be a generic feature of all organisms.

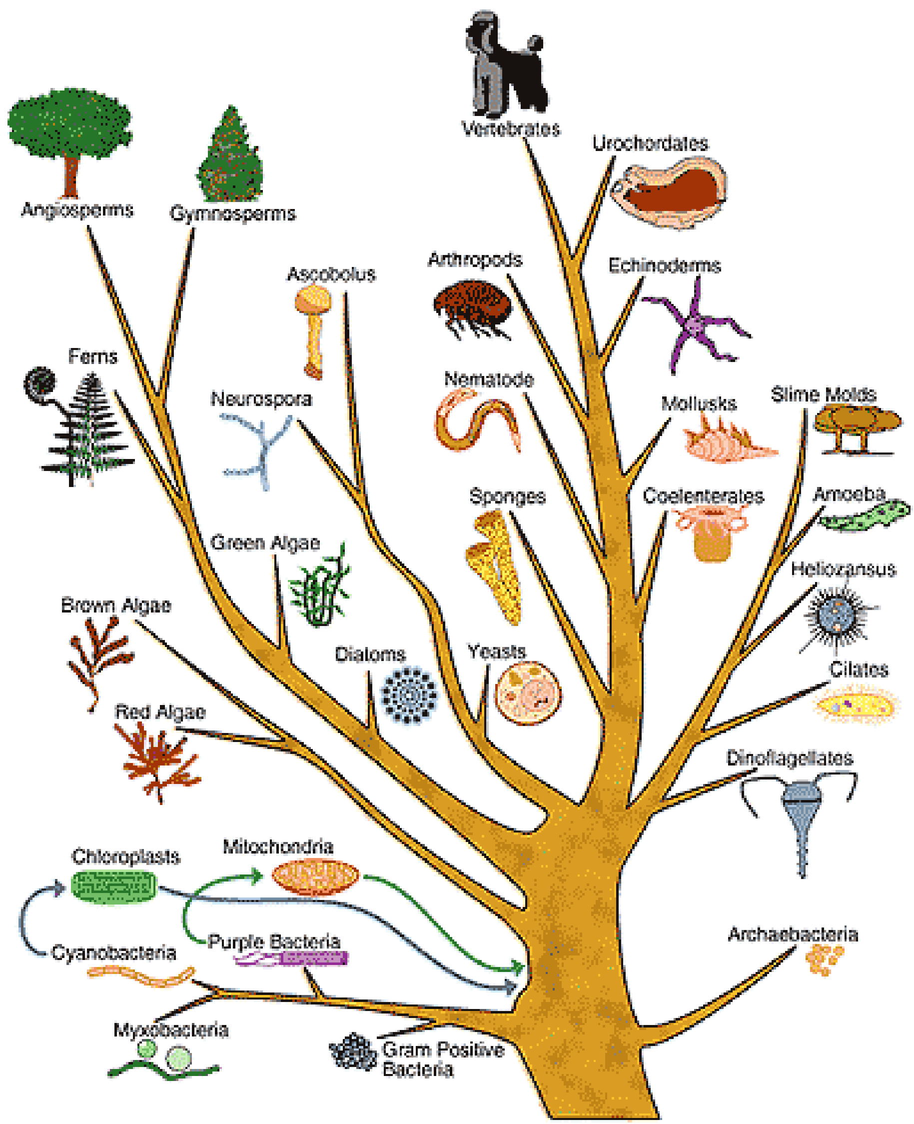
HEMMINGSEN ~1950'S



Allometric scaling of metabolic rate for a selection of homeotherms (birds and mammals), poikilotherms (fish, reptiles, amphibians, and invertebrates), and unicellular organisms. The solid lines all have a slope of .75. Modified from Hemmingsen, 1960.

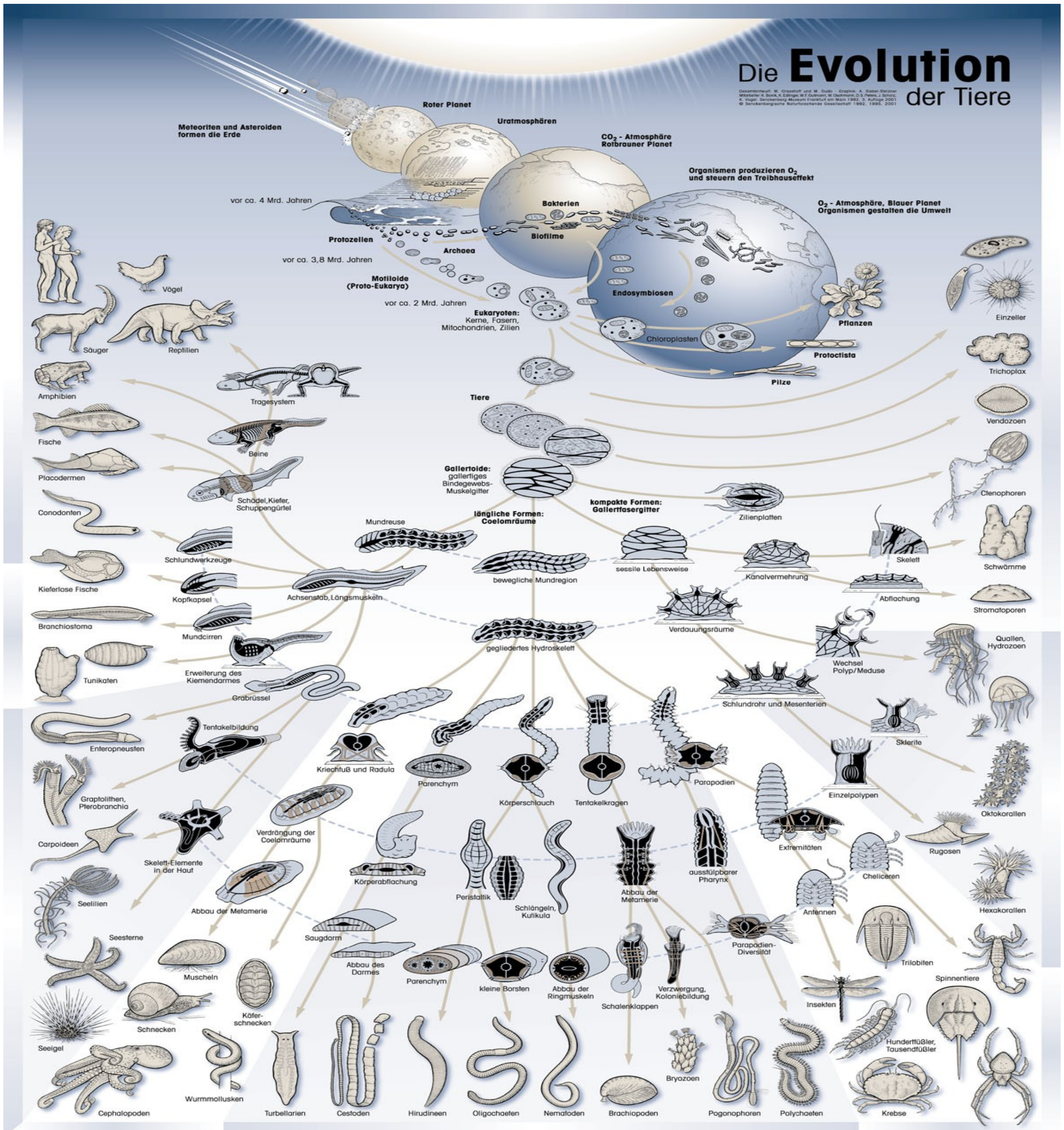
Biological energy transformation systems: the same scaling law is observed over 27 orders of magnitude

Biology Phylogeny: procariotes, eucariotes, genetic networks



Genetic network: ***population size distribution of species, species size distribution of genus, genera size distribution of biological family are of the PZM Pareto-Zipf-Mandelbrot type (hyperbolic fractal)***

"In terms of genetic evolution mankind is close to big apes, in terms of social evolution mankind is much closer to ants , termites and bees." Peter Winiwarter

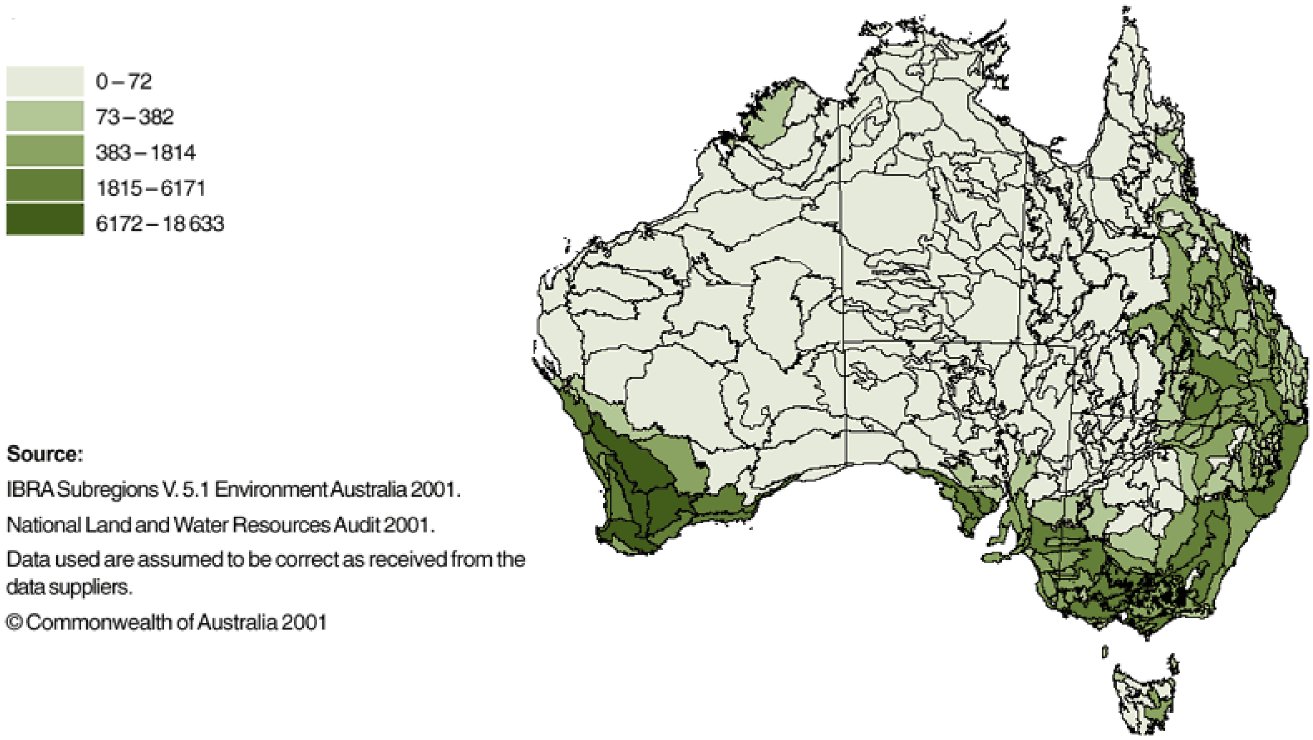


PZM (Pareto-Zipf-Mandelbrot, parabolic fractal) distributions are observed for all species at <u>all times of biological evolution</u>

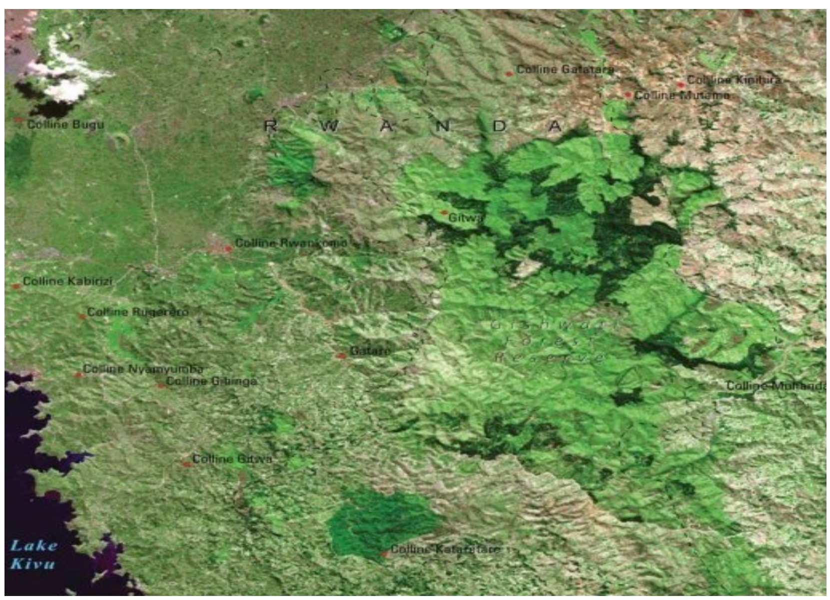
"its the same underlying computational algorithms which drive evolution. Mutations are not random, they are computed." Peter Winiwarter

Biology Ontogeny: trophic ecosystems, trophic networks

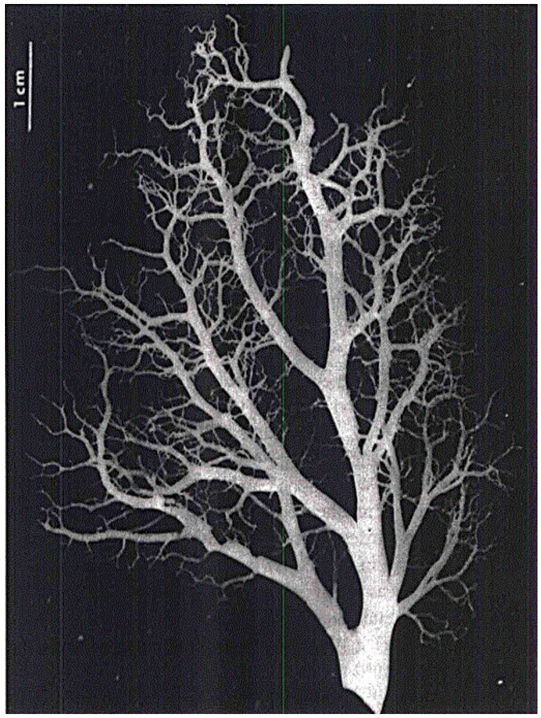
Forest network: branch size distributions, leave size distributions and the distribution of tree stem size are of PZM type



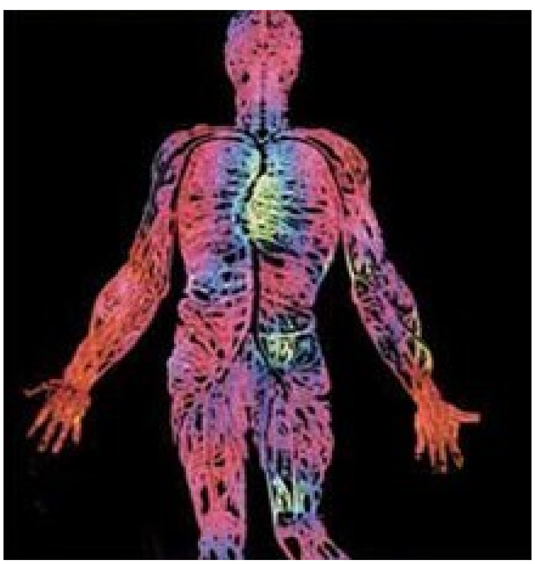
Solar energy transformation network: patches of vegetation size distribution are of the PZM type

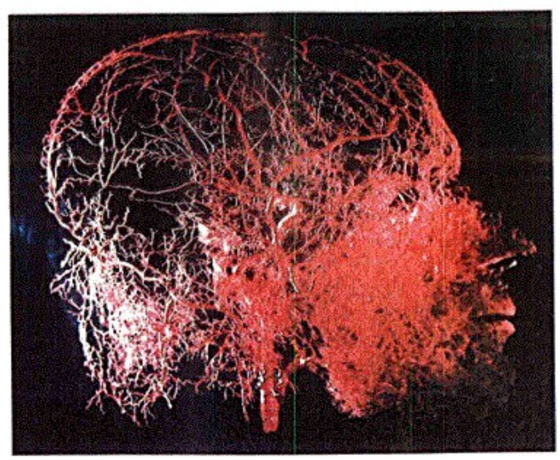


Forest network: distribution of areas burnt in forest fires are of the PZM Pareto-Zipf-Mandelbrot type (parabolic fractal)

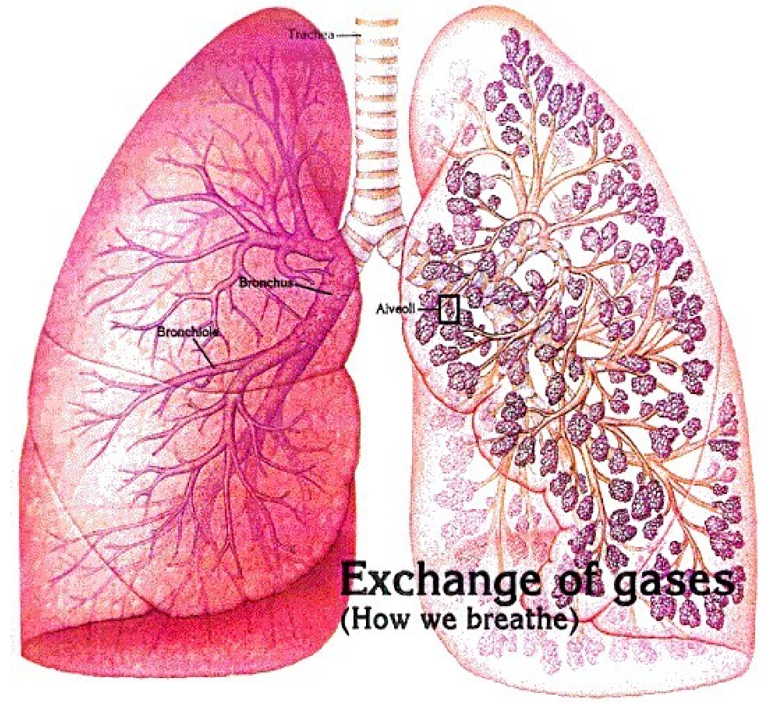


Blood vascular network: blood vessel shows PZM regularity

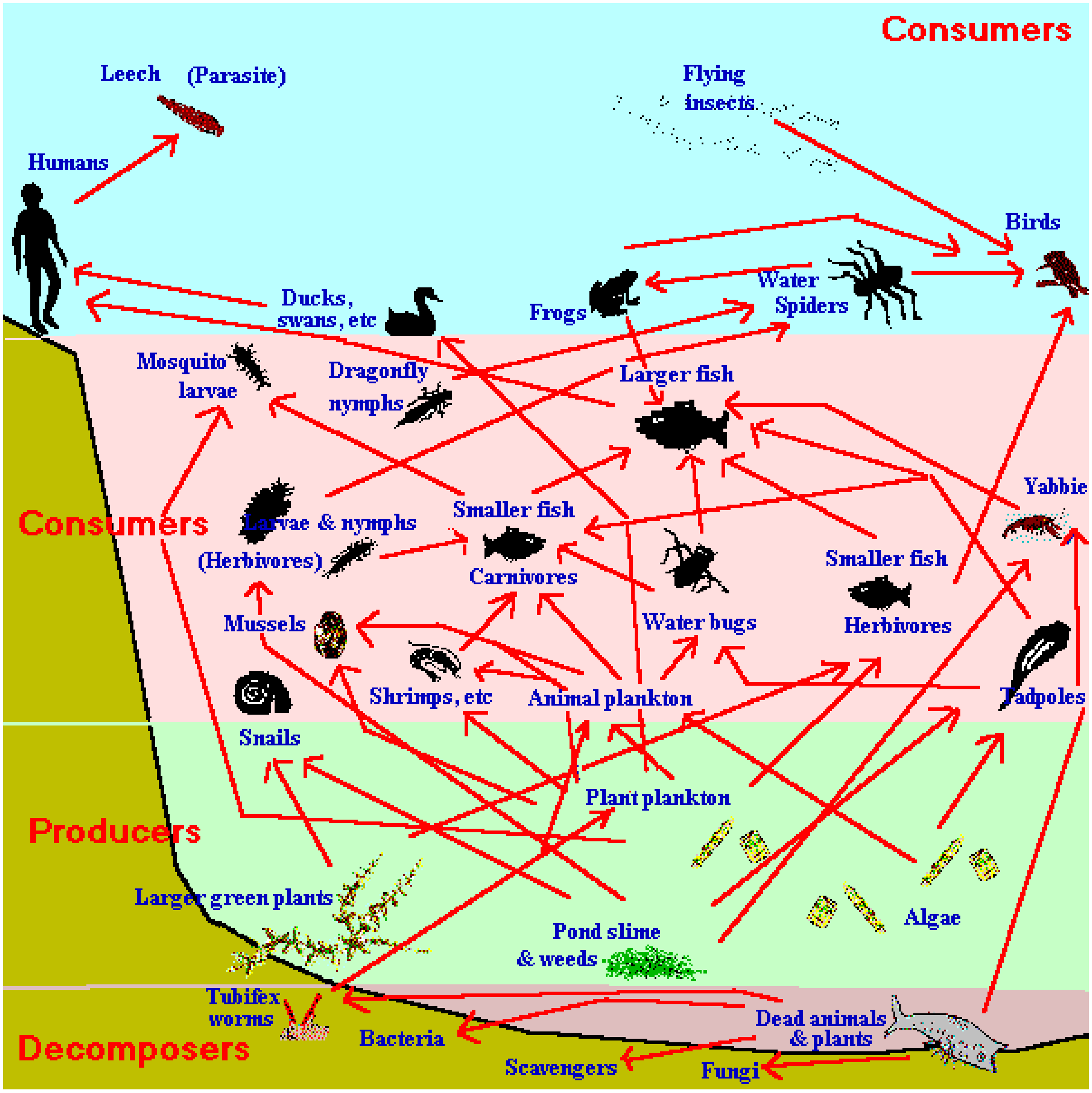




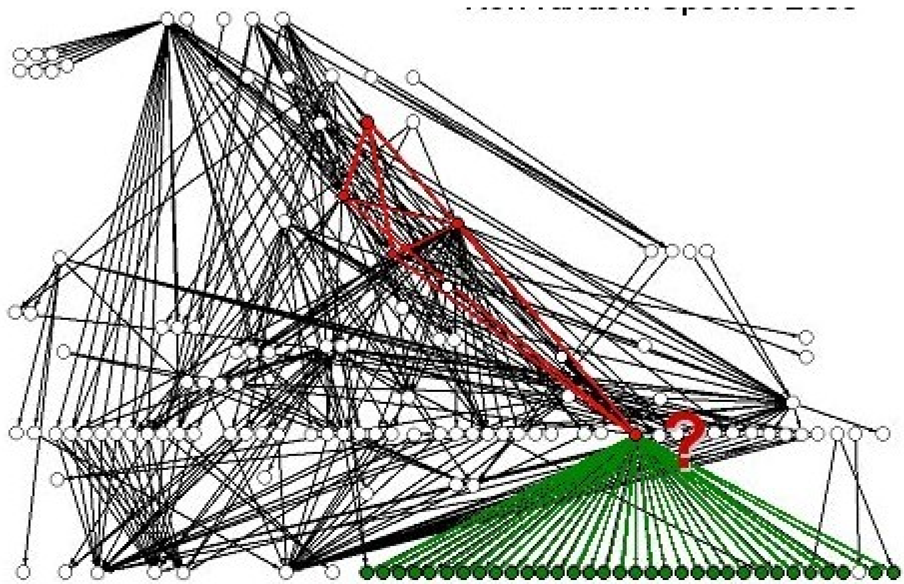
30.18 Blood vascular network in the human head The human face, particularly the area around the lips, contains a dense array of capillaries.



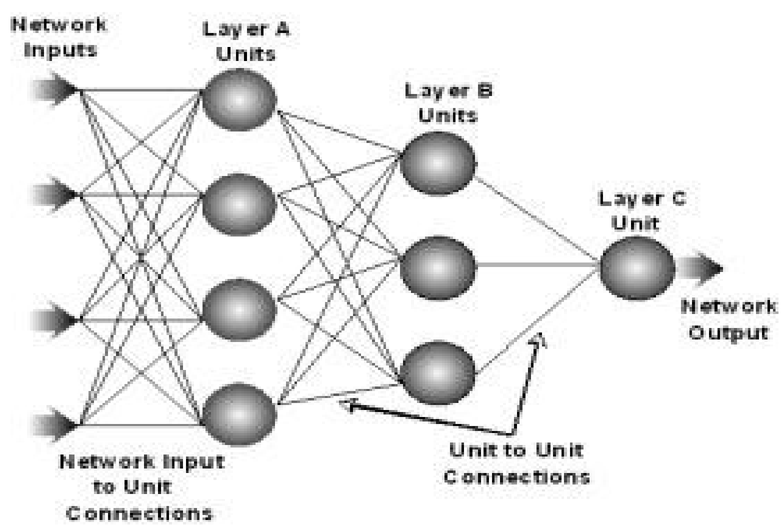
Metabolic network: the structure of the lung is fractal of th PZM type



Food web network:
biomass-size distribution of aquatic ecosystems (trophic web or foodweb) show PZM regularity



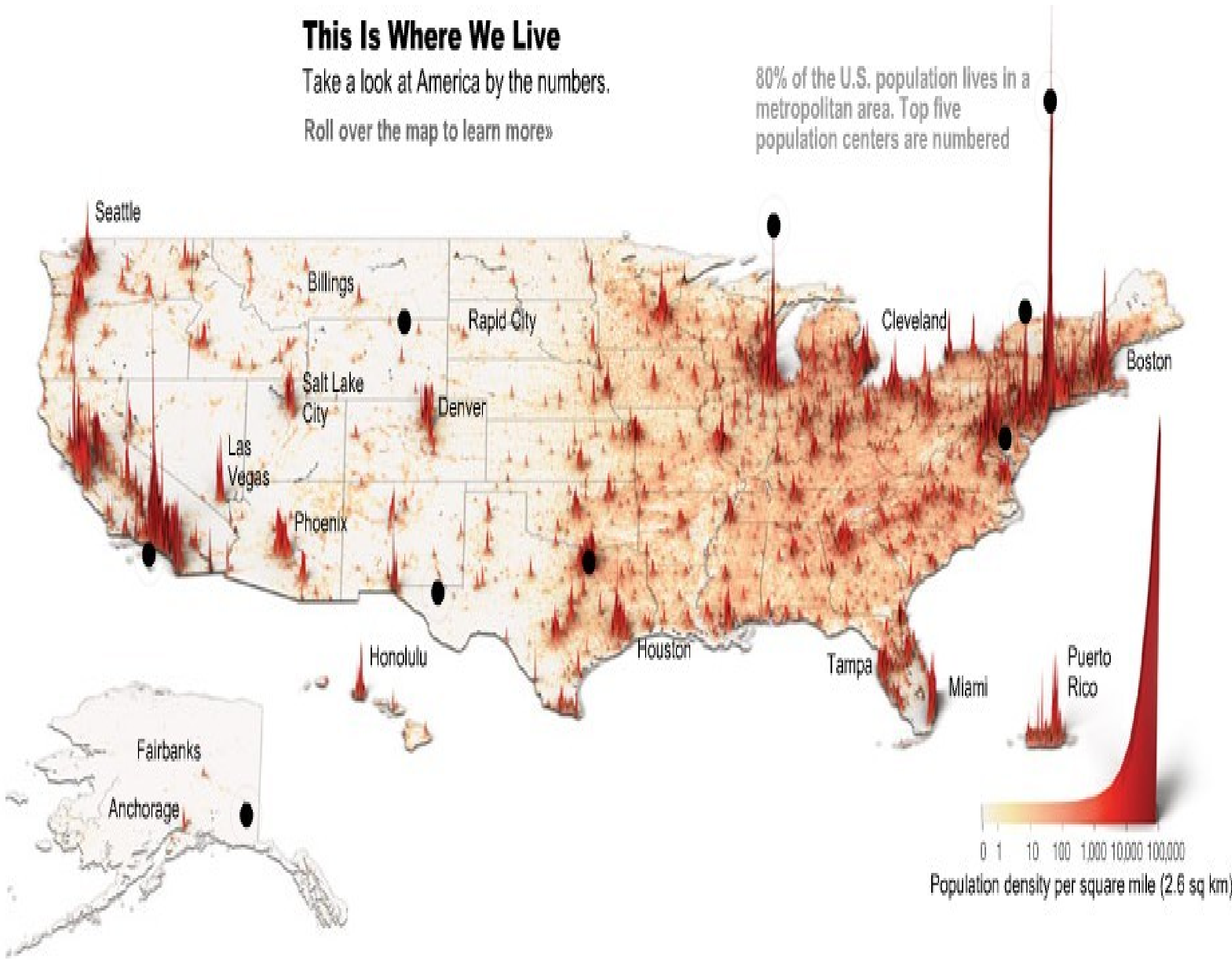
Winiwarter and Vidondo modelled the ecosystem evolution of the lake Constance by a neural network of the multilayer feed forward type with back-propagatio (multilayer perceptron)



Input layer: time series of daily solar energy input to the lake d, d-1,d-2, .. d-365

Output layer: a single constant, the slope of the PZM biomass size distribution at day d

Social networks: the small world of scalefree networks

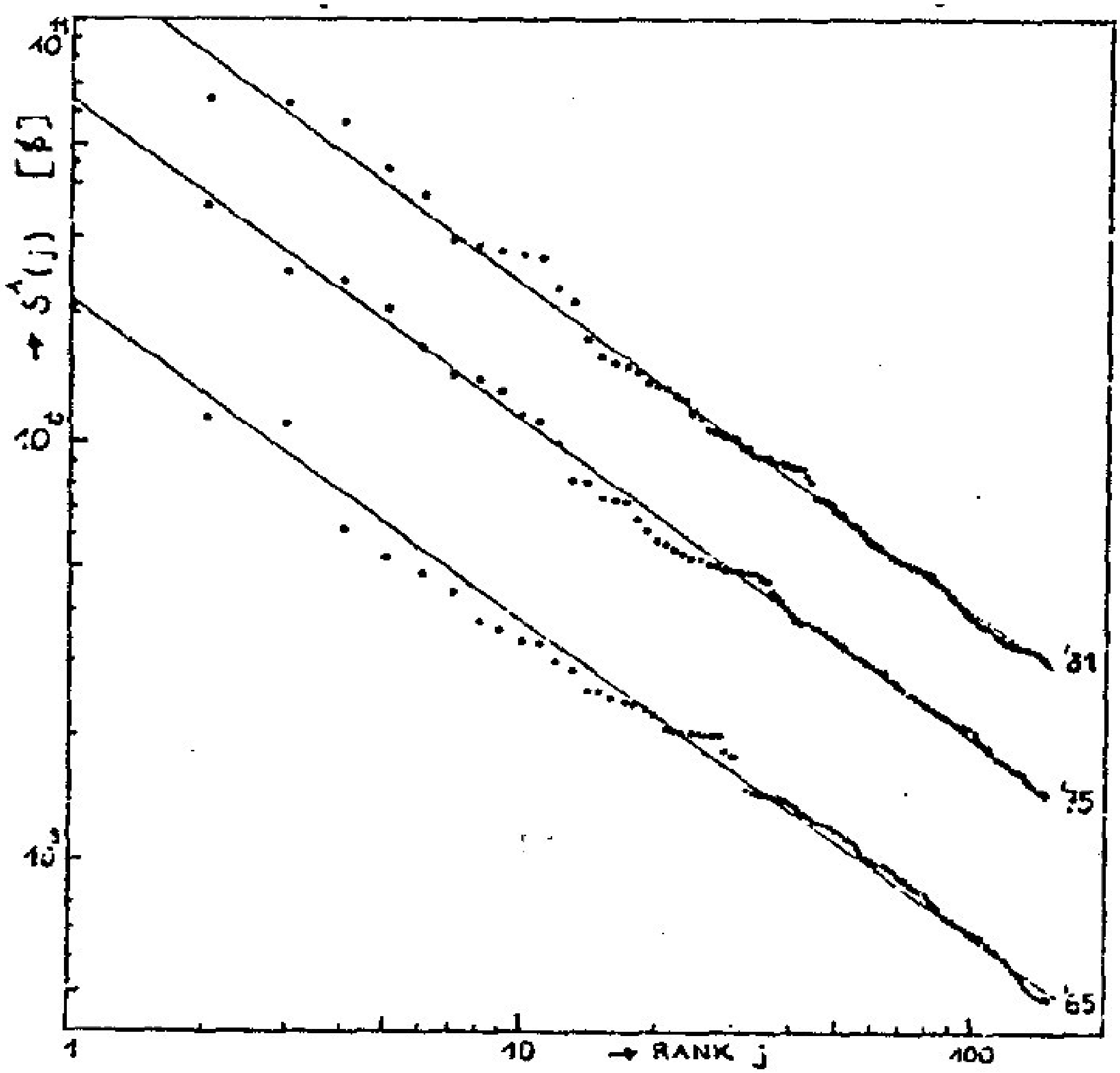


Population network: city size distributions of all countries follow a PZM regularity (rank size rule)

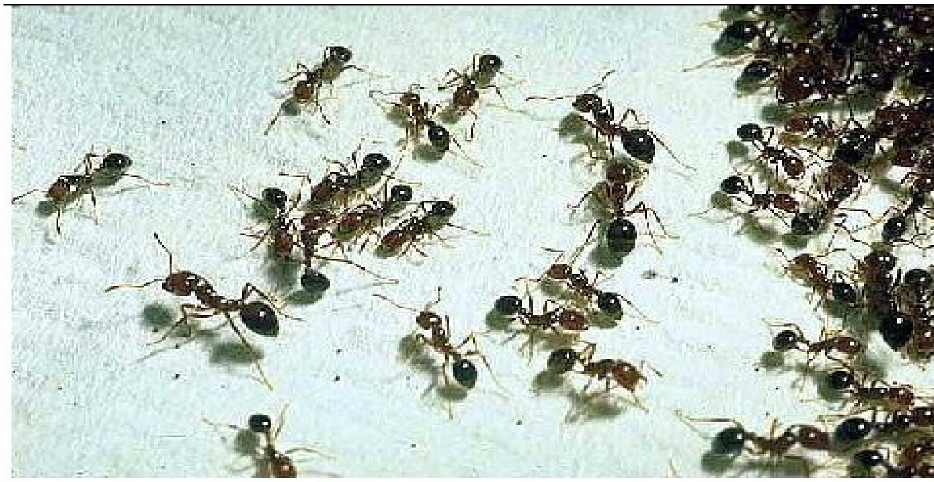
"The objective of social sciences does not consist any more in the reduction of complex to simple, but in the translation of complex into theory."
Peter Winiwarter

Network of personal wealth: the fortunes of individuals follow a PZM distribution

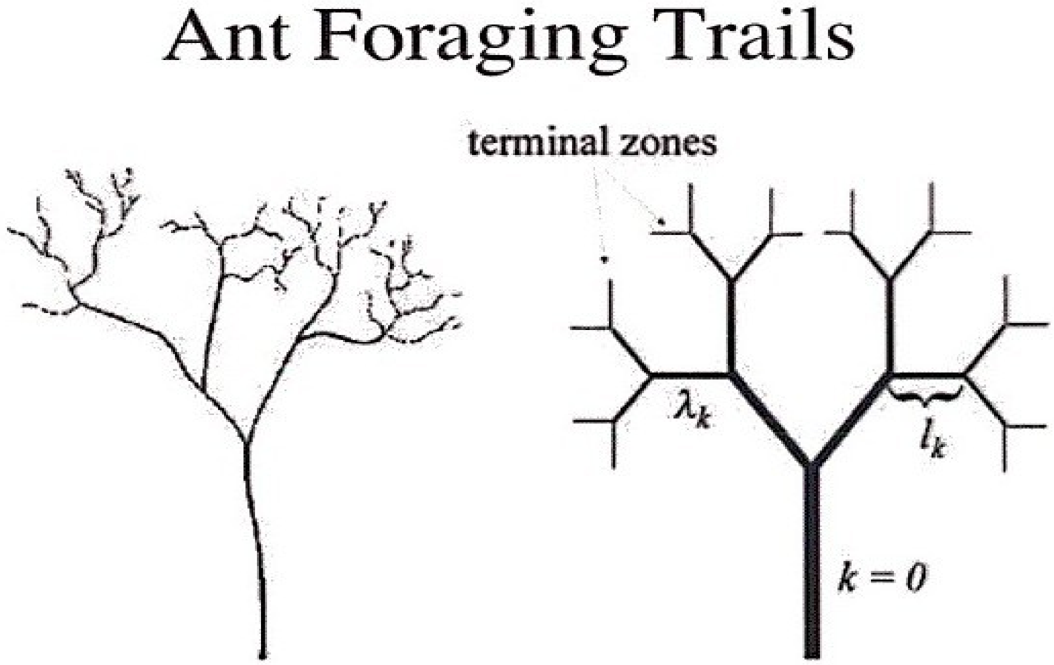
There are very few very rich, few rich, many small and and awful lot of poor



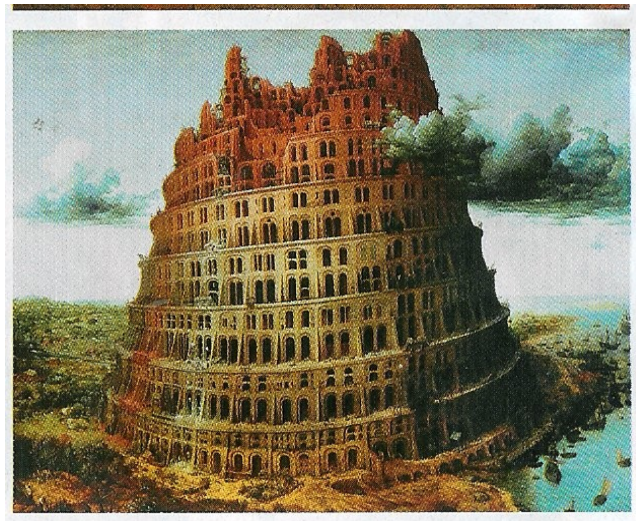
Network of business: the firm size distribution of the Fortune 500 follow a PZM regularity. Note that the slope of the diostribution is almost constant over time, only the size of the overall system grows (data from 1965, 1975 and 1981, the time of the study).



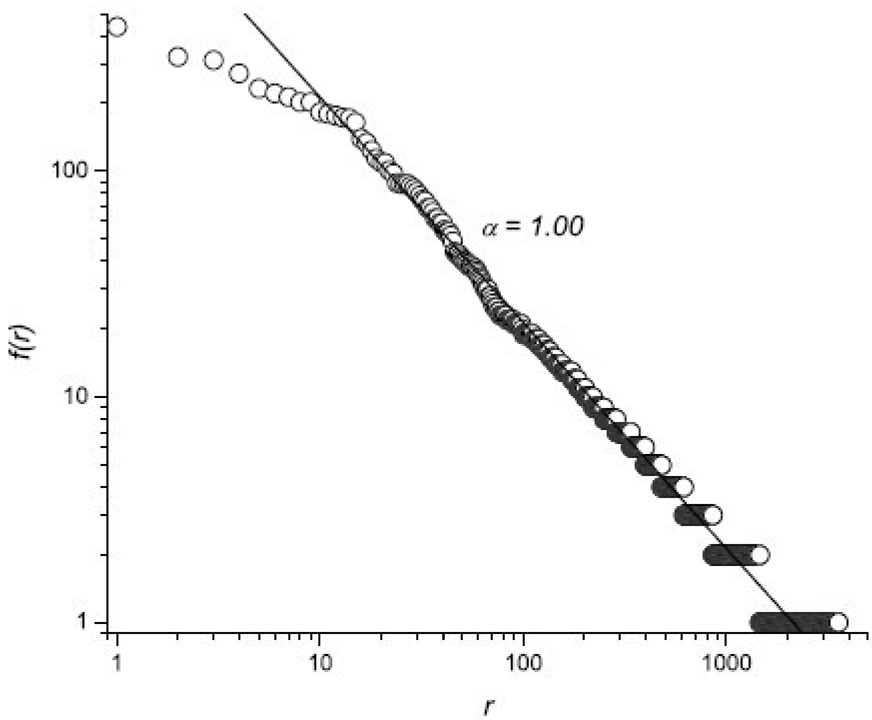
antswarm



Social networks: ant foraging trails show a PZM regularity

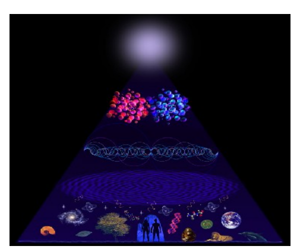


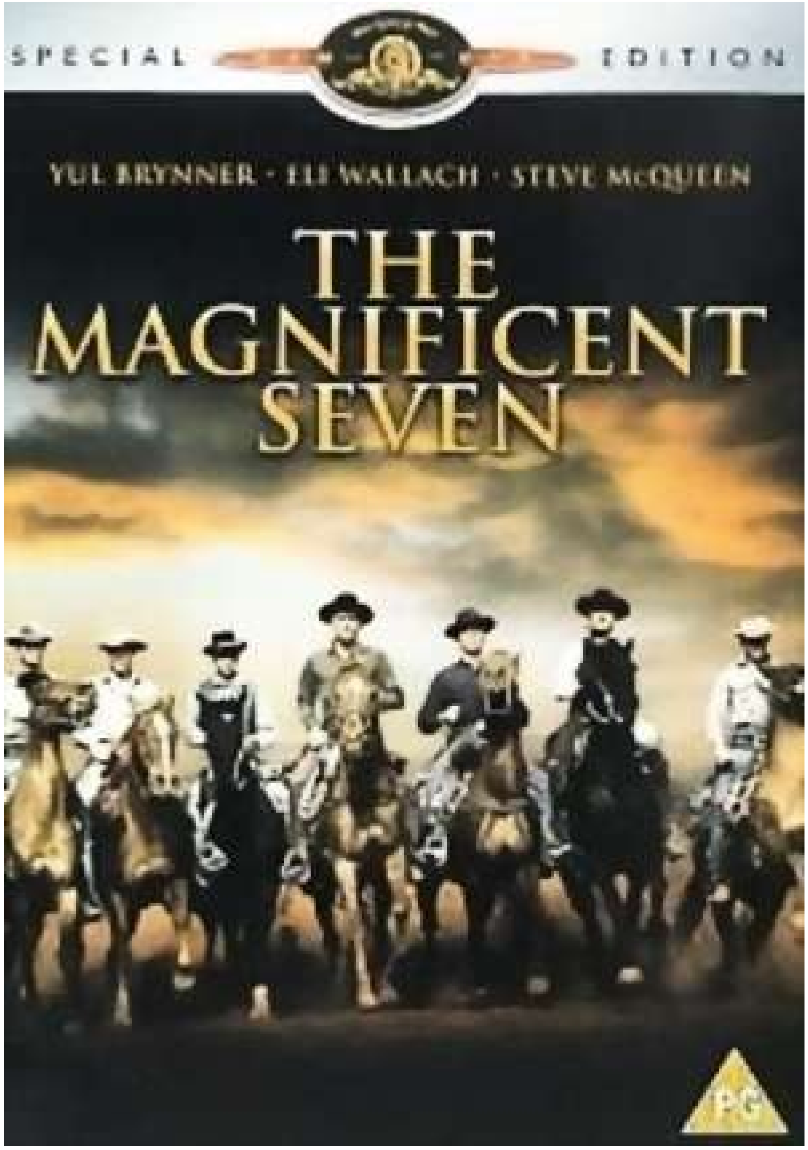
The network of language: all languages of the world and texts from all times follow a PZM regularity (Zipf's law)



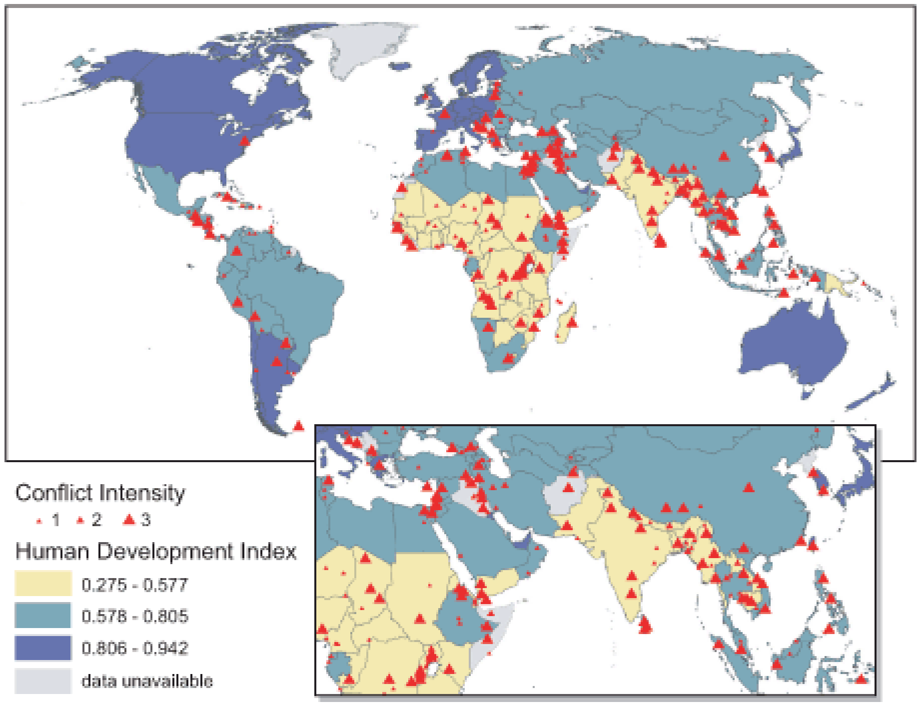
Genealogical networks: the degree distribution of genealogical networks are of the PZM type

Mythology and Religion network: the network of figures in Greek and Roman mythology and in Christian religion reveal a degree distribution of PZM type





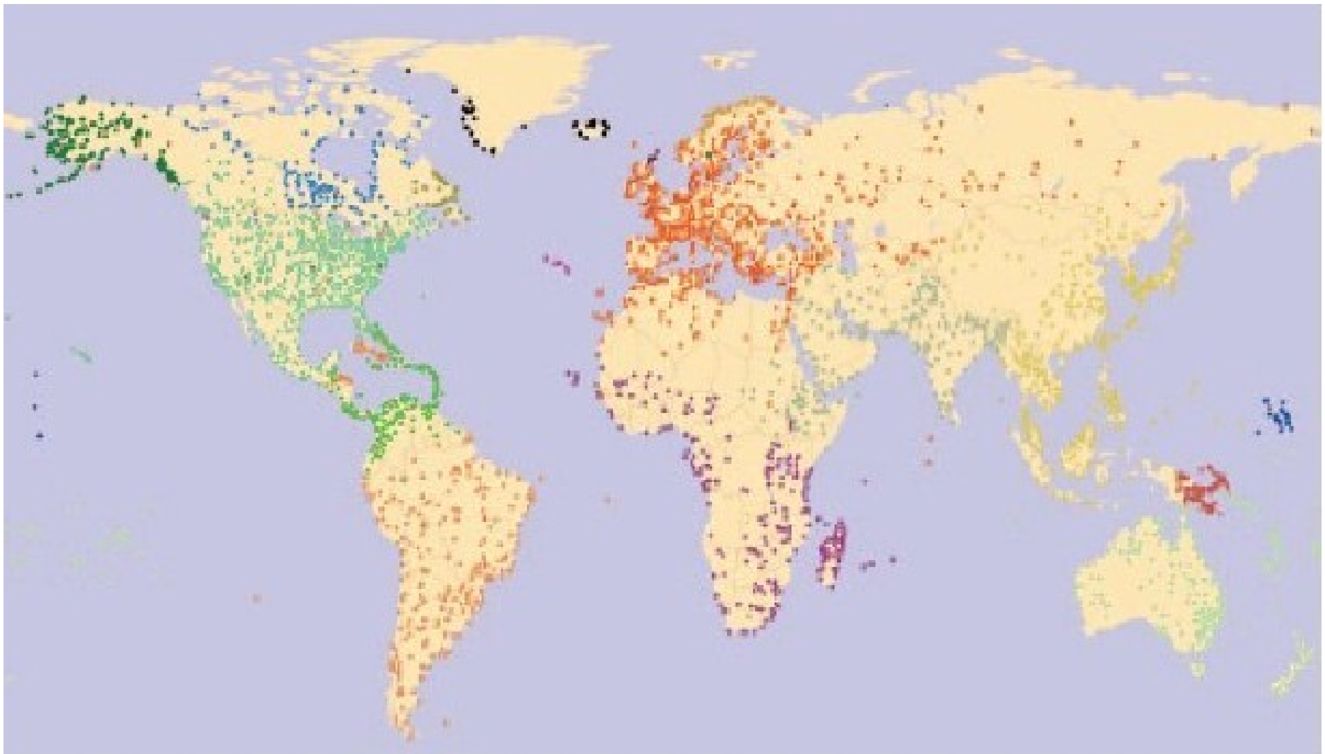
Hollywood network: the movie co-actors of the Internet Movie Database show a degree distribution of the PZM type. Any actor of any movie is not more than four movie links away from Kevin Bacon. (Actor A played with actor B in movie X, B played with C in movie Y, C played with Kevin Bacon in 'a few good man', hence actor A is only three links away from Kevin Bacon)



Network of national and international conflicts: the battle deaths per war distribution are of PZM type

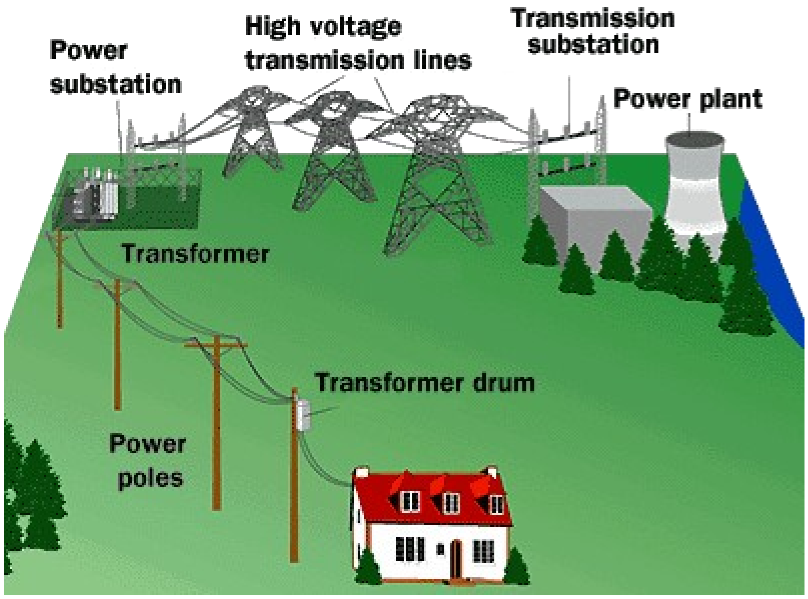
Technology networks: from stone tools to the internet

Transportation network: the Los Angeles Public Transportation Network consist of 1881 routes and 44629 stations (nodes) revealing PZM regularity. Similar regularities are observed for all major metropolitan areas of the world.

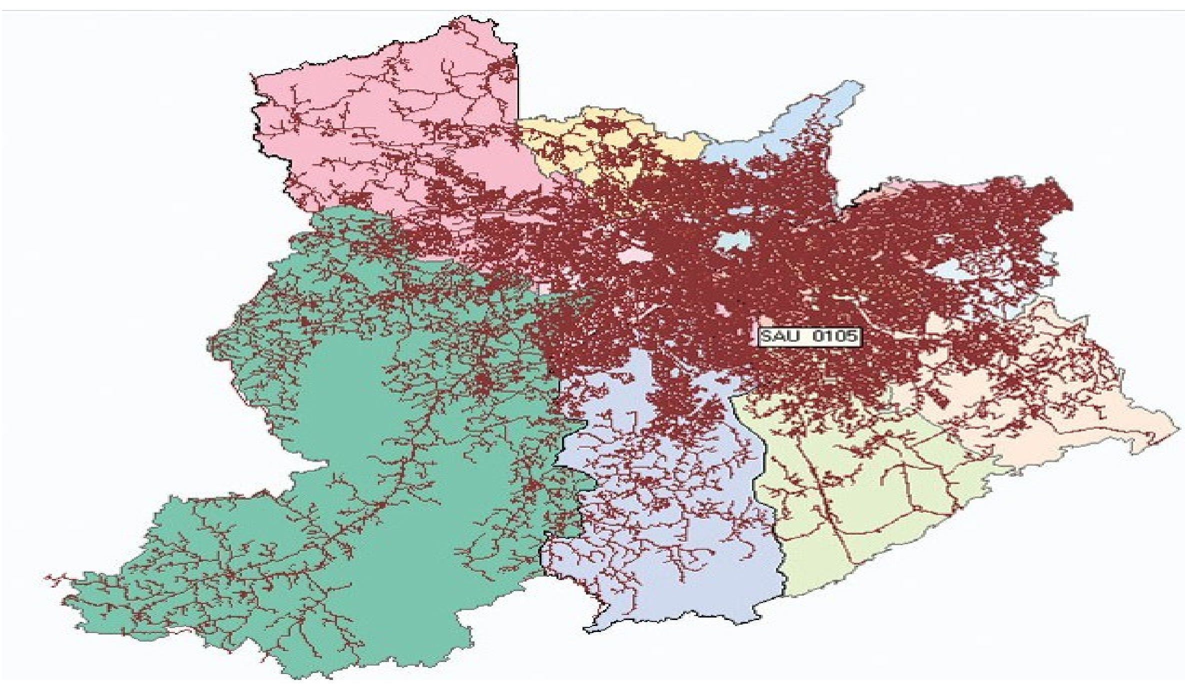


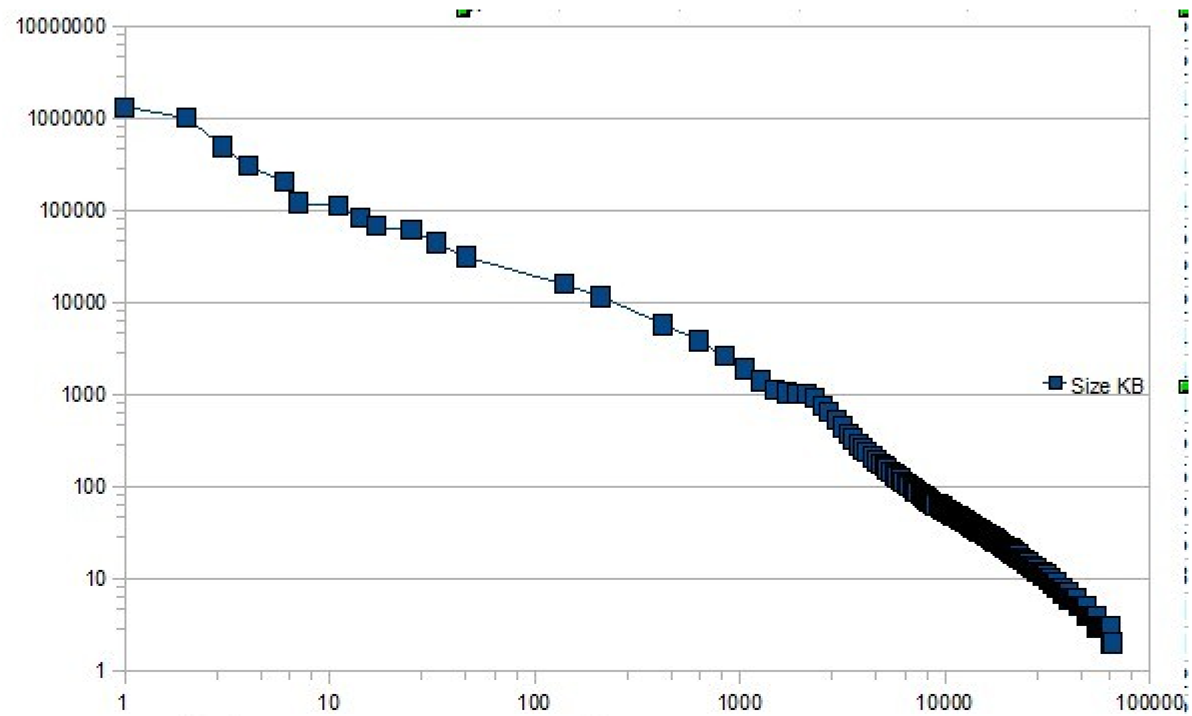
Airport network: the worldwide and national air transportation networks reveil small world property with a dregree distribution of the PZM type

network for which the number of direct connections
k to a given city (degree) has a cumulative distribution
P(> k) that decays as a truncated power-law
$P(> k) \propto k^{-\alpha} f(k/k_\times)$,
where $\alpha = 1.0 \pm 0.1$ *is the power-law exponent,* $f(u)$ *is*
a truncation function, and $k_\times$ *is a crossover value that*
depends on the size S of the network as $k_\times \sim S^{0.4}$.



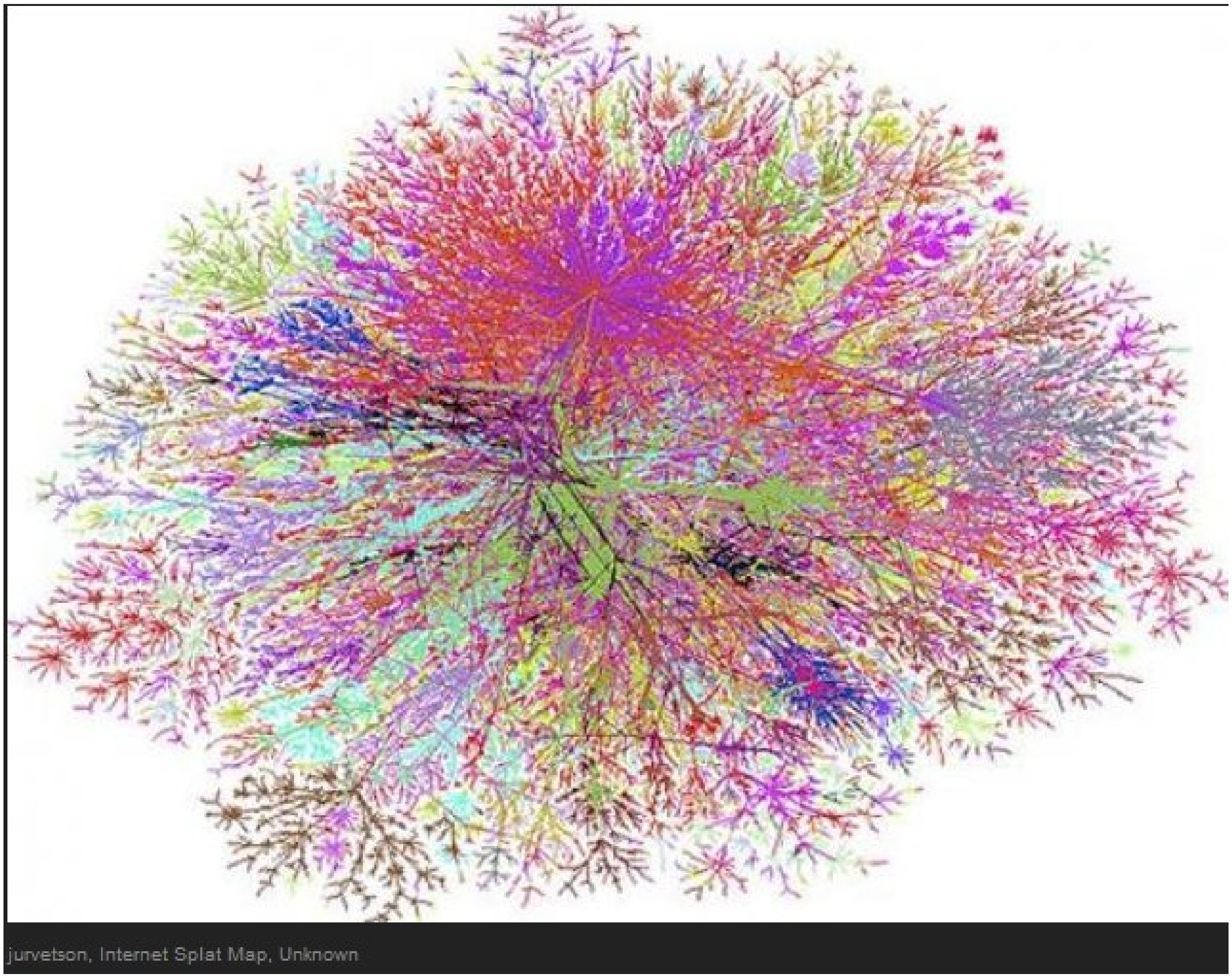
Electric energy distribution network: power grid distribution lines are of the PZM type



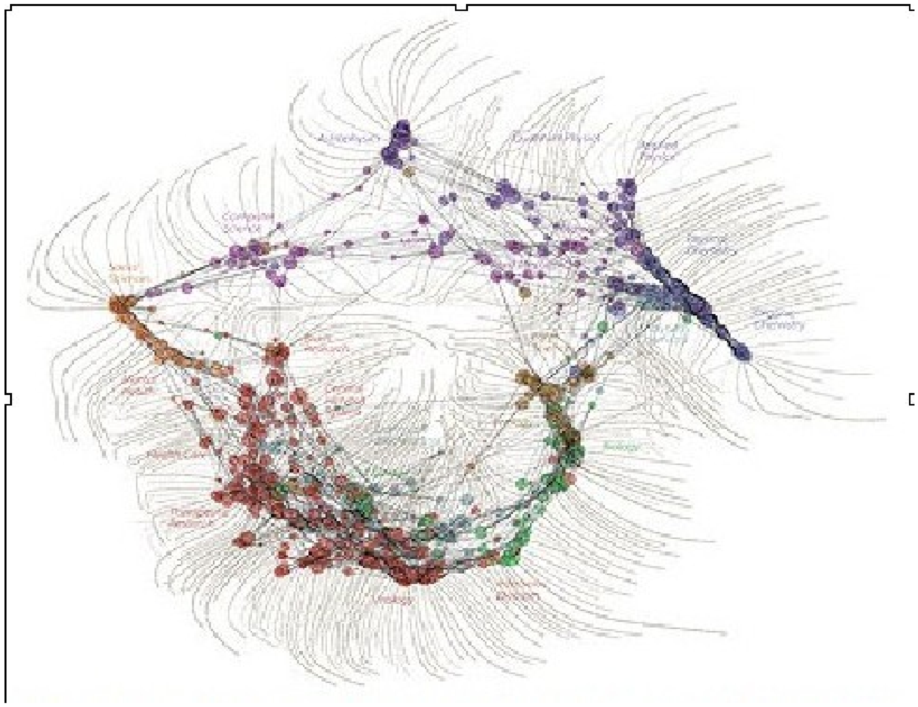


Computer file network: the file size distributions on the hard drive of a PC on a UNIX system and on the WEB are of PZM type

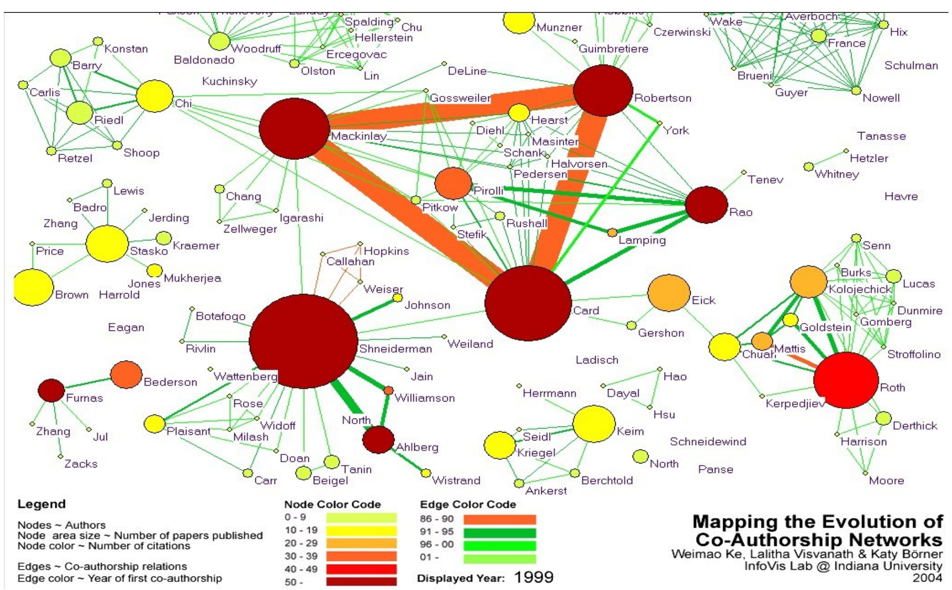
Nuclear energy network: the size of nuclear explosions follow a PZM regularity



Information network: the World Wide Web reveals scalefree power laws of the PZM type for site size distribution, incoming links, outgoing links ...



Research network: the citation network of research reveal PZM degree distribution



What do all these illustrated regularities have in common?

"It is an interesting possibility that the power laws followed by so many different kinds of systems might be the result of downward constraintes exerted by encompassing supersystems."

Stanley N. Salthe, ***Entropy*** 2004, **6**, 335

- All the systems for which we observe Pareto-Zipf-Mandelbrot (PZM) regularities are **networks.**
- All networks have a '**small world**' topology between complete randomness and complete order.
- The networks belong to one of the three categories:
  1) **Matter** transportation networks (e.g. public transportation network, water network)
  2) **Energy** transformation and transportation networks (e.g. Electricity network)
  3) **Information** networks (e.g. Telephone Network, the Internet and WWW)
- All small world networks reveal a cyclic hierarchical **feed forward** process : bottom up flow of information from singular interaction units over local modules (threshold automata / switches) to global hubs.
- All small world networks reveal a **back-propagatio** process : top down flow of information from central hubs over local modules (threshold automata / switches) down to individual interaction units.
- As will be shown in the following chapter, any small world network can be mapped on an Artificial Neural Network of the multilayer feed forward type with back-propagatio (**multilayer perceptron**)
- Such the features of multilayer perceptrons like memory, learning and universal mapping capability are inherent to all systems/networks for which we observe Pareto-Zipf-Mandelbrot (PZM) regularities.

For example let us consider the world airport network. A passenger wants to travel from one desert airport El Centro in Imperial county in southern California to another desert airport Tamanrasset in the south of Algeria.

The travel schedule will bring him from the El Centro / Imperial county airport IPL to LAX, the hub of the Los Angeles airport. From LAX he will take a flight to PAR, the Paris Roissy airport hub of France. From PAR he will take a flight to the smaller modular hub of the Algiers airport ALG. Finally from ALG he will take a small plane to arrive at the Tamanrasset airport TMR in the deep south desert of the Sahara close to the Hoggar mountain chain where he intended to admire prehistoric wall paintings of our ancestors. It took our traveler only four flights and three correspondence changes to arrive from one distant location to another distant location of the world. It's a small world in the true sense of complex graph theory.

Let us take another axample, the network of the Internet and the world wide web. From my location (www.bordalierinstitute.com) I want to visit the home page of www.evodevouniverse.com. First the packages of my request will travel to the local hub of my French Internet service provider www.1and1.fr, from there it will be routed the nameserver of www.1and1.com in the US, which looks up the IP number of evodevouniverse, which is hosted by a US service provider and from there the packages of my request will be routed down to the final IP adress of my correspondent. Once reached a similar bottom up and top down direction will send the requested page through the routers of the web back to my site. In the case of the Internet the traveling packages of my message can take different routes, but they will be re-assembled at arrival. Again we observe a cyclic feed forward of information with subsequent back-propagation through the complex web of the network.

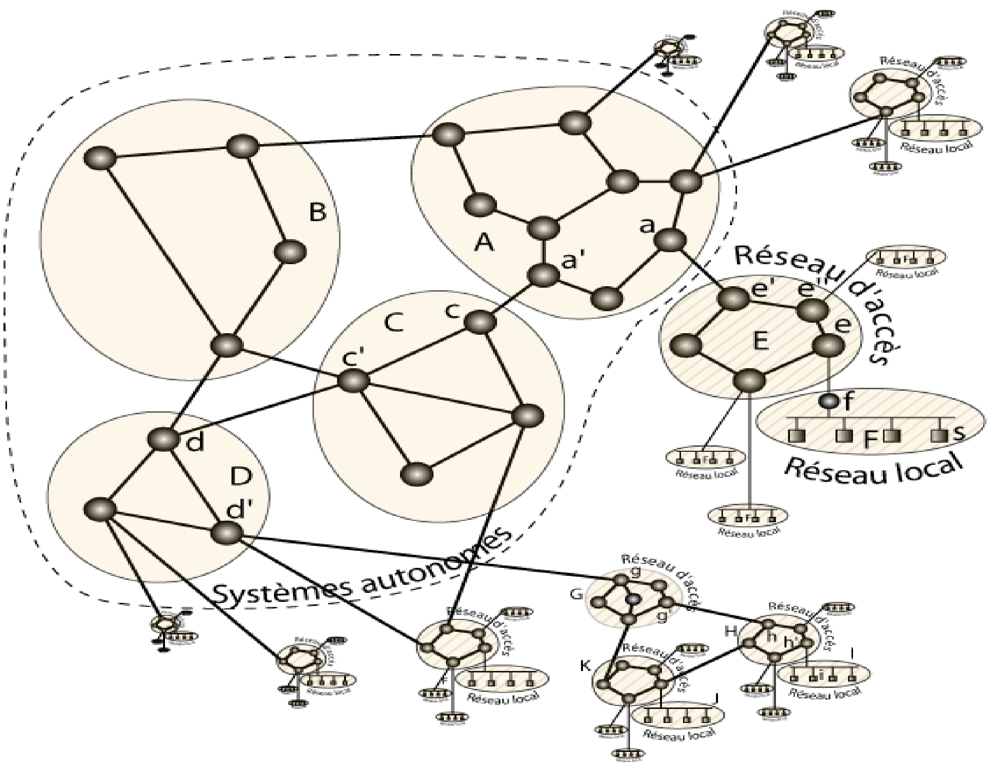


Figure: routing through the Internet: PC of local network, module of Internet access, autonomous Routers of Internet, module of Internet access, local network of PCs, individual interaction unit PC.

As third example we propose to examine the case of Citation webs. According to Lotka's law Citation webs exhibit a Pareto-Zipf-Mandelbrot distribution for citation frequencies.
A citation web can be modeled as a hierarchical multilayer feed forward Neural Network with cyclic back-propagatio (*Multilayer Perceptron*).
In a first approximation the hierarchical feed forward levels of a citation web are:

- single citation (a pointer to an idea) feeding into the
- paper level, which feeds into the
- referee level (peer-system) which feeds into the
- conference proceedings or journal publication level, which feeds into the
- book-editor level, which feeds into the
- book level, feeding the
- institute library level and finally feeding the
- Library of Congress level.

back-propagatio are the respective <u>reference lists of citations</u> fed down the levels of the hierarchy.
On each level there are binary threshold processors, which can be only **on** or **off**:

- Paper in progress vs. paper written,
- Paper submitted for conference vs. paper accepted for conference,
- Paper submitted to journal vs. paper published by journal,
- Book submitted to editor vs. book published by editor,
- Book proposed to library vs. book acquired by library
- Book on shelf vs. book scrapped from library ...

All these binary threshold automata of the citation web are interlinked in a hierarchical way and undergo a cyclic feed forward process with consecutive back-propagatio. If the Neural Network analogy holds we come to the following conclusions:

2. **Citation webs have «memory»**.
   It is the topology of the web's authors and their respective links in the citation webgraph which constitute the memory of the self-organized system.
3. **Citation webs are «learning».**
   Through a cyclic feedback process of *reference lists* through the different hierarchical levels of the system. An author's «weight» is proportional to the number of citations in the citation index.
4. **Citation webs are «intelligent».**
   The cyclic self-organization process (feed forward and consecutive back-propagatio) optimizes the overall coherence (synergy) of the system. Thus the system is striving to an extremal value of an objective function (goal).

Self-similarity and the beauty of Fractals

A fractal is generally "a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole,"[1] a property called self-similarity. The term was coined by Benoît Mandelbrot in 1975 and was derived from the Latin fractus meaning "broken" or "fractured." A mathematical fractal is based on an equation that undergoes iteration, a form of feedback based on recursion.

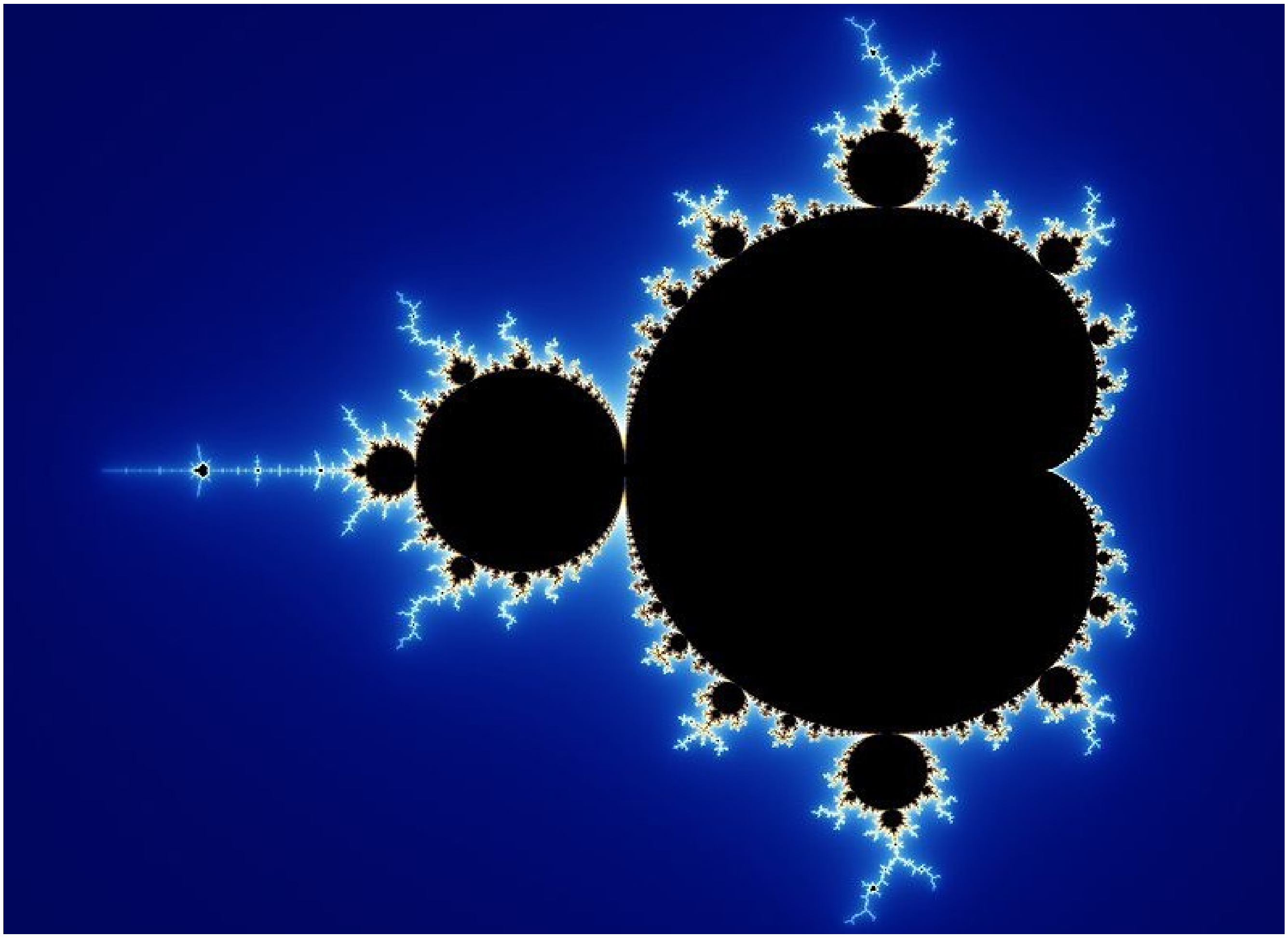
fractals, Wikipedia

A fractal often has the following features:[3]

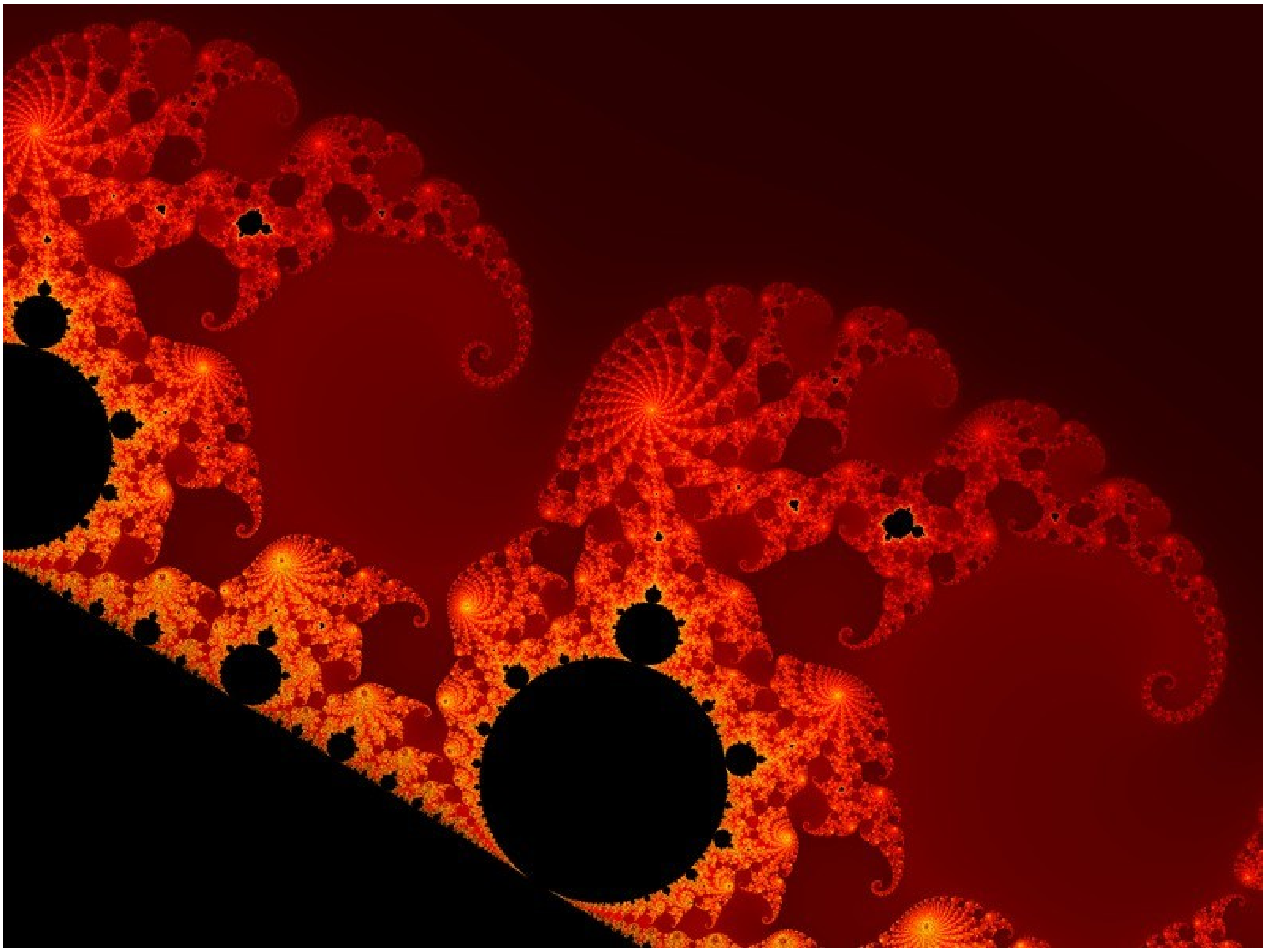
- It has a fine structure at arbitrarily small scales.
- It is too irregular to be easily described in traditional Euclidean geometric language.
- It is self-similar (at least approximately or stochastically).
- It has a Hausdorff dimension which is greater than its topological dimension (although this requirement is not met by space-filling curves such as the Hilbert curve).
- It has a simple and recursive definition.

Because they appear similar at all levels of magnification, fractals are often considered to be infinitely complex (in informal terms). Natural objects that approximate fractals to a degree include clouds, mountain ranges, lightning bolts, coastlines, and snow flakes. However, not all self-similar objects are fractals—for example, the real line (a straight Euclidean line) is formally self-similar but fails to have other fractal characteristics; for instance, it is regular enough to be described in Euclidean terms.

Images of fractals can be created using fractal generating software. Images produced by such software are normally referred to as being fractals even if they do not have the above characteristics, as it is quite possible to zoom into a region of the image that does not exhibit any fractal properties.



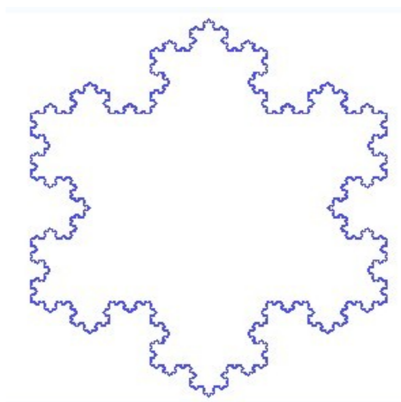
The Mandelbrot set is a famous example of a fractal.



A closer view of the Mandelbrot set

History

Animated construction of a Sierpiński Triangle, only going four generations of infinite



To create a Koch snowflake, one begins with an equilateral triangle and then replaces the middle third of every line segment with a pair of line segments that form an equilateral "bump." One then performs the same replacement on every line segment of the resulting shape, ad infinitum. With every iteration, the perimeter of this shape increases by one third of the previous length. The Koch snowflake is the result of an infinite number of these iterations, and has an infinite length, while its area remains finite. For this reason, the Koch snowflake and similar constructions were sometimes called "monster curves."

The mathematics behind fractals began to take shape in the 17th century when mathematician and philosopher Leibniz considered recursive self-similarity (although he made the mistake of thinking that only the straight line was self-similar in this sense).

It took until 1872 before a function appeared whose graph would today be considered fractal, when Karl Weierstrass gave an example of a function with the non-intuitive property of being everywhere continuous but nowhere differentiable. In 1904, Helge von Koch, dissatisfied with Weierstrass's very abstract and analytic definition, gave a more geometric definition of a similar function, which is now called the Koch snowflake. In 1915, Waclaw Sierpinski constructed his triangle and, one year later, his carpet. Originally these geometric fractals were described as curves rather than the 2D shapes that they are known as in their modern constructions. In 1918, Bertrand Russell had recognized a "supreme beauty" within the mathematics of fractals that was then emerging.[2] The idea of self-similar curves was taken further by Paul Pierre Lévy, who, in his 1938 paper *Plane or Space Curves and Surfaces Consisting of Parts Similar to the Whole* described a new fractal curve, the Lévy C curve. Georg Cantoralso gave examples of subsets of the real line with unusual properties—these Cantor sets are also now recognized as fractals.

Iterated functions in the complex plane were investigated in the late 19th and early 20th centuries by Henri Poincaré, Felix Klein, Pierre Fatouand Gaston Julia. However, without the aid of modern computer graphics, they lacked the means to visualize the beauty of many of the objects that they had discovered.

In the 1960s, Benoît Mandelbrot started investigating self-similarity in papers such as *How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension*, which built on earlier work by Lewis Fry Richardson. Finally, in 1975 Mandelbrot coined the word "fractal" to denote an object whose Hausdorff–Besicovitch dimension is greater than its topological dimension. He illustrated this mathematical definition with striking computer-constructed visualizations. These images captured the popular imagination; many of them were based on recursion, leading to the popular meaning of the term "fractal".

Examples

A class of examples is given by the Cantor sets, Sierpinski triangle and carpet, Menger sponge, dragon curve, space-filling curve, and Koch curve. Additional examples of fractals include the Lyapunov fractal and the limit sets of Kleinian groups. Fractals can be deterministic (all the above) or stochastic (that is, non-deterministic). For example, the trajectories of the Brownian motion in the plane have a Hausdorff dimension of 2.

Chaotic dynamical systems are sometimes associated with fractals. Objects in the phase space of a dynamical system can be fractals (see attractor). Objects in the parameter space for a family of systems may be fractal as well. An interesting example is the Mandelbrot set. This set contains whole discs, so it has a Hausdorff dimension equal to its topological dimension of 2—but what is truly surprising is that the boundary of the Mandelbrot set also has a Hausdorff dimension of 2 (while the topological dimension of 1), a result proved by Mitsuhiro Shishikura in 1991. A closely related fractal is the Julia set.

Generating fractals

Four common techniques for generating fractals are:

- **Escape-time fractals** — (also known as "orbits" fractals) These are defined by a formula or recurrence relation at each point in a space (such as the complex plane).Examples of this type are the Mandelbrot set, Julia set, the Burning Ship fractal, the Nova fractal and the Lyapunov fractal. The 2d vector fields that are generated by one or two iterations of escape-time formulae also give rise to a fractal form when points (or pixel data) are passed through this field repeatedly.
- **Iterated function systems** — These have a fixed geometric replacement rule. Cantor set, Sierpinski carpet, Sierpinski gasket, Peano curve, Koch snowflake, Harter-Heighway dragon curve, T-Square, Menger sponge, are some examples of such fractals.
- **Random fractals** — Generated by stochastic rather than deterministic processes, for example, trajectories of the Brownian motion, Lévy flight,fractal landscapes and the Brownian tree. The latter yields so-called mass- or dendritic fractals, for example, diffusion-limited aggregation orreaction-limited aggregation clusters.
- **Strange attractors** — Generated by iteration of a map or the solution of a system of initial-value differential equations that exhibit chaos.

Classification of fractals

Fractals can also be classified according to their self-similarity. There are three types of self-similarity found in fractals:

- **Exact self-similarity** — This is the strongest type of self-similarity; the fractal appears identical at different scales. Fractals defined by iterated function systems often display exact self-similarity.
- **Quasi-self-similarity** — This is a loose form of self-similarity; the fractal appears approximately (but not exactly) identical at different scales. Quasi-self-similar fractals contain small copies of the entire fractal in distorted and degenerate forms. Fractals defined by recurrence relations are usually quasi-self-similar but not exactly self-similar.
- **Statistical self-similarity** — This is the weakest type of self-similarity; the fractal has numerical or statistical measures which are preserved across scales. Most reasonable definitions of "fractal" trivially imply some form of statistical self-similarity. (Fractal dimension itself is a numerical measure which is preserved across scales.) Random fractals are examples of fractals which are statistically self-similar, but neither exactly nor quasi-self-similar.

Fractals in nature

Approximate fractals are easily found in nature. These objects display self-similar structure over an extended, but finite, scale range. Examples include clouds, snow flakes, crystals, mountain ranges, lightning, river networks, cauliflower or broccoli, and systems of blood vessels and pulmonary vessels. Coastlines may be loosely considered fractal in nature.

Figure. A fractal fern and Photograph of a cauliflower, showing naturally a occuring fractal

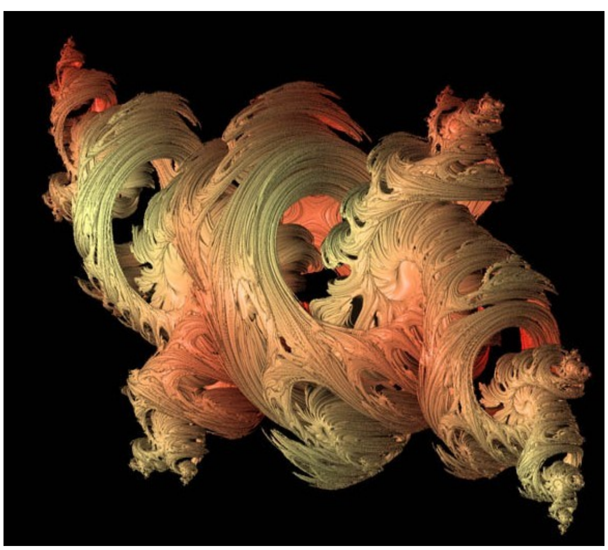


Figure. Fractal shell and fractal tree

Trees and ferns are fractal in nature and can be modeled on a computer by using a recursive algorithm. This recursive nature is obvious in these examples — a branch from a tree or a frond from a fern is a miniature replica of the whole: not identical, but similar in nature. The connection between fractals and leaves are currently being used to determine how much carbon is really contained in trees. This connection is hoped to help determine and solve the environmental issue of carbon emission and control.

Applications of fractals

Main article: Fractal analysis

As described above, random fractals can be used to describe many highly irregular real-world objects. Other applications of fractals include:[10]

- Classification of histopathology slides in medicine
- Fractal landscape or Coastline complexity
- Enzyme/enzymology (Michaelis-Menten kinetics)
- Generation of new music
- Generation of various art forms
- Signal and image compression
- Creation of digital photographic enlargements
- Seismology
- Fractal in soil mechanics
- Computer and video game design, especially computer graphics for organic environments and as part of procedural generation
- Fractography and fracture mechanics
- Fractal antennas — Small size antennas using fractal shapes
- Small angle scattering theory of fractally rough systems
- T-shirts and other fashion
- Generation of patterns for camouflage, such as MARPAT
- Digital sundial
- Technical analysis of price series (see Elliott wave principle)

Fractal dynamics

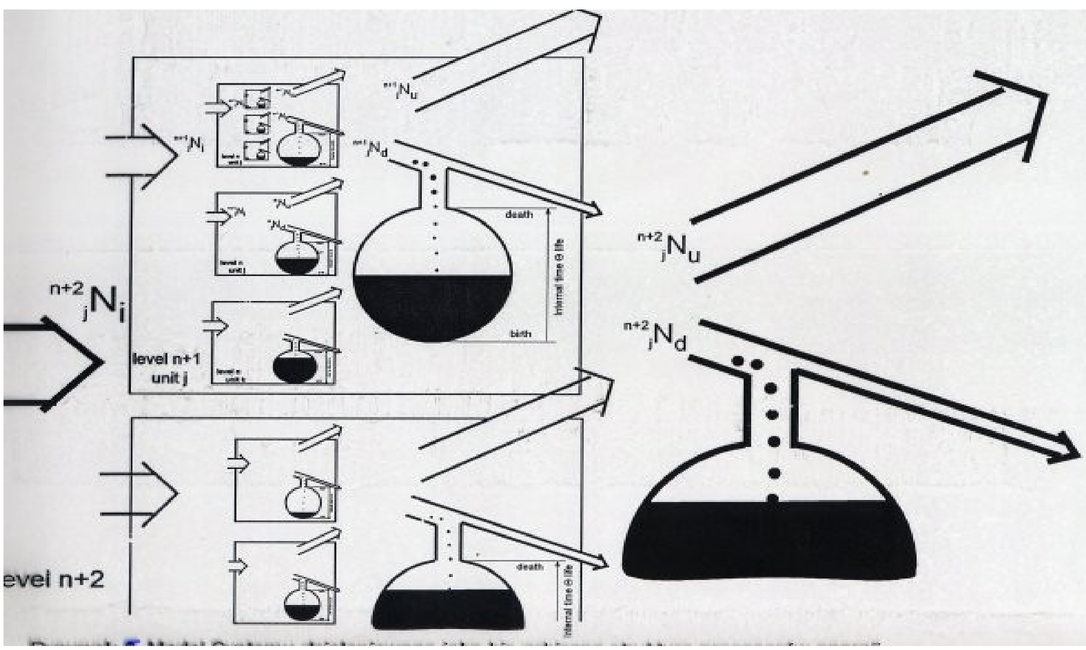
Almost all fractals with the exception of attractors in chaotic dynamical systems are topological fractals, where the geometry of the structure reveals a self-similarity.

In the chapter on theory we will also have a look at self-similar fractal like processes like the self-similar hierarchy of energy transformation processors called birth and death processors and the self-similar structure of feed forward Artificial Neural Networks with backpropagation called Perceptrons.

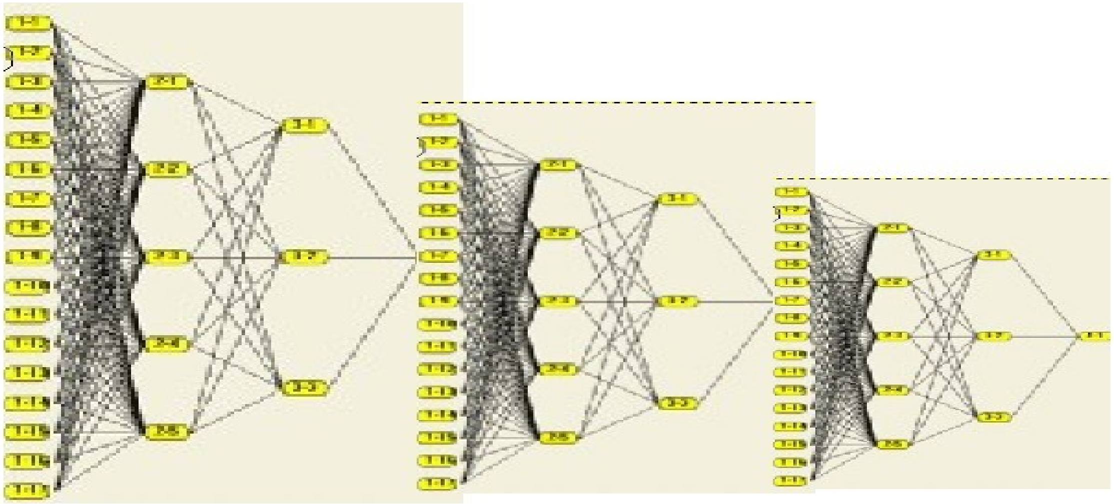
In these fractal hierarchies we observe a **self-similar process** of energy / information transformation on each level of the hierarchy when zooming in or zooming out.

A birth and death processor accumulates internal dissipation energy until its death or breakdown.

This process takes place on any level of the hierarchy.



The self-similar hierarchy of energy / information transformation processors [Winiwarter 1992]



The self-similar hierarchy of information transformation processors (multilayer perceptrons)

In the case of a formal neuron, each neuron accumulates weighted inputs until it fires at threshold.

This process takes place on any level of the hierarchy.

The approach of self-similar recursion is called recursionism and is at the basis of the metaphysical background of this book.

Networks everywhere

Networks are ubiquitous. They serve to model any type of physical realm. A network in general is an interconnected group or system, or a fabric or structure of fibrous elements attached to each other at regular intervals, or formally: a graph. The network approach is at the core of the ideas put forward in this book.

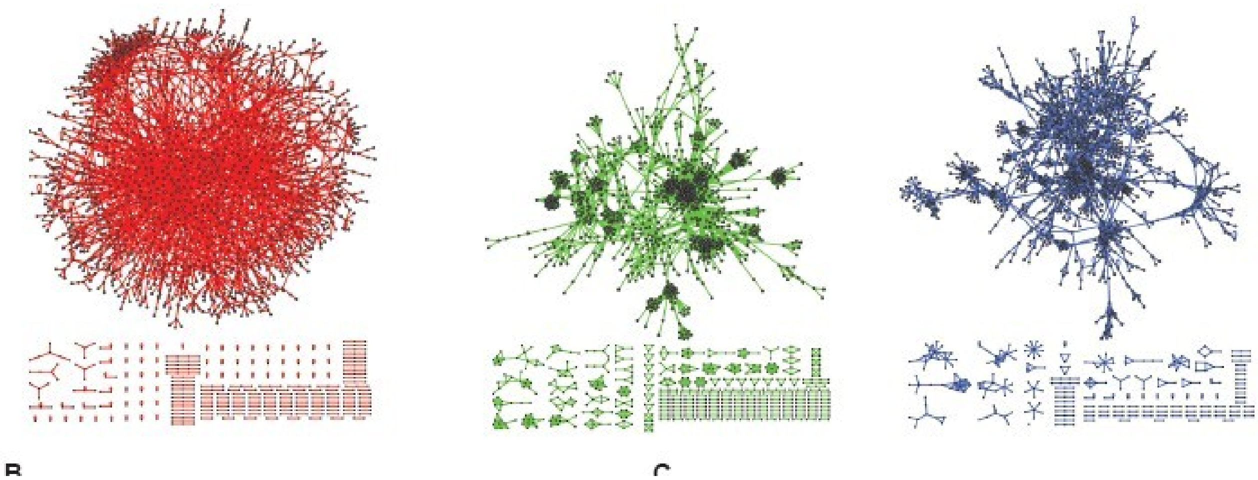


Figure. Network analysis of different data sets

Networks, Wikipedia

A network diagrams is a special kind of cluster diagram, which even more general represents any cluster or small group or bunch of something, structured or not. Both the flow diagram and the tree diagram can be seen as a specific type of network diagram.

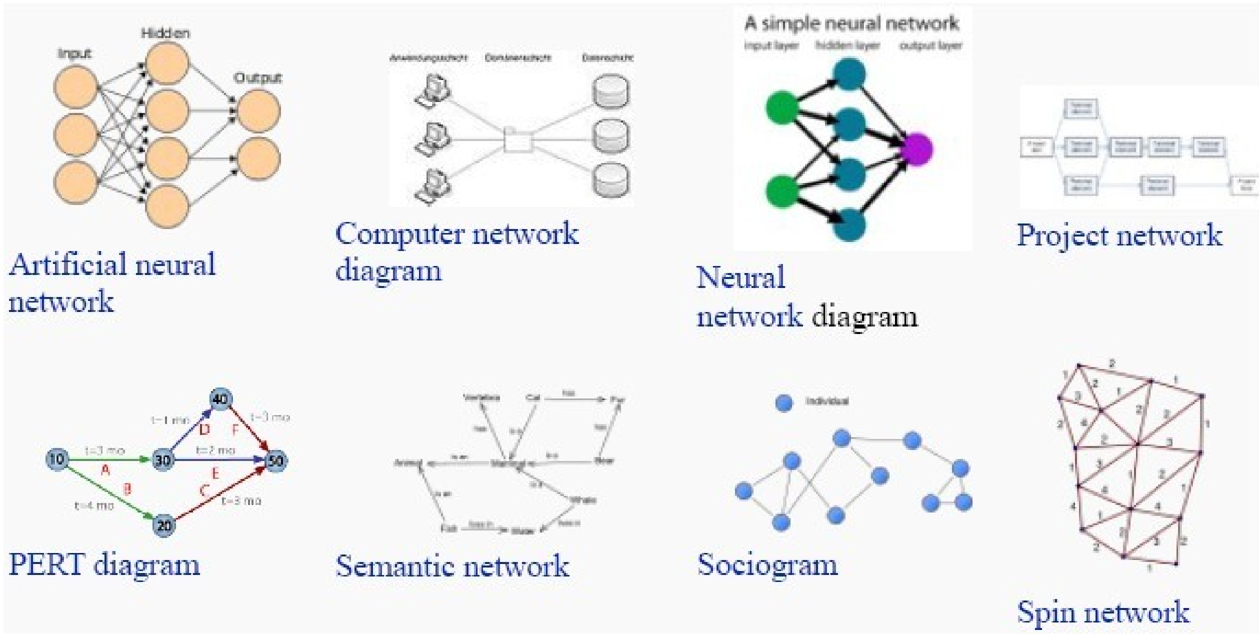
Types of network diagrams

There are different types network diagrams:

- Artificial neural network or "neural network" (NN), is a mathematical model or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation.

- Computer network diagram is a schematic depicting the nodes and connections amongst nodes in a computer network or, more generally, any telecommunications network.
- In project management a network diagram is the logical representation of activities, that defines the sequence or the work of a project. It shows the path of a project, lists starting and completion dates , and names the responsibilities for each task. At a glance it explains how the work of the project goes together. A network for a simple project might consist one or two pages, and on a larger project several network diagrams may exist.[1] Specific diagrams here are
    - Project network: a general flow chart depicting the sequence in which a project's terminal elements are to be completed by showing terminal elements and their dependencies.
    - PERT network
- Neural network diagram: is a network or circuit of biological neurons or artificial neural networks, which are composed of artificial neurons or nodes.
- A semantic network is a network or circuit of biological neurons. The modern usage of the term often refers to artificial neural networks, which are composed of artificial neurons or nodes]].[2]
- A sociogram is a graphic representation of social links that a person has. It is a sociometric chart that plots the structure of interpersonal relations in a group situation.

Gallery



Network topologies

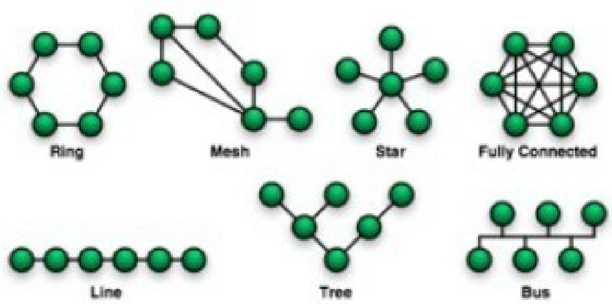


Figure. Diagram of different network topologies.

In computer science the elements of a network are arranged in certain basic shapes (see figure):

- Ring: The ring network connects each node to exactly two other nodes, forming a circular pathway for activity or signals –a ring. The interaction or data travels from node to node, with each node handling every packet.

- Mesh is a way to route data, voice and instructions between nodes. It allows for continuous connections and reconfiguration around broken or blocked paths by "hopping" from node to node until the destination is reached.
- Star: The star network consists of one central element, switch, hub or computer, which acts as a conduit to coordinate activity or transmit messages.
- Fully connected: Self Explanatory
- Line –Everything connected in a single line.
- Tree: This consists of tree-configured nodes connected to switches/concentrators, each connected to a linear bus backbone. Each hub rebroadcasts all transmissions received from any peripheral node to all peripheral nodes on the network, sometimes including the originating node. All peripheral nodes may thus communicate with all others by transmitting to, and receiving from, the central node only.
- Bus: In this network architecture a set of clients are connected via a shared communications line, called a bus.

Network theory

Network theory is an area of applied mathematics and part of graph theory. It has application in many disciplines including particle physics, computer science, biology, economics, operations research, and sociology. Network theory concerns itself with the study of graphs as a representation of either symmetric relations or, more generally, of asymmetric relations between discrete objects. Examples of which include logistical networks, the World Wide Web, gene regulatory networks, metabolic networks, social networks, epistemological networks, etc.

Network optimization

Network problems that involve finding an optimal way of doing something are studied under the name of combinatorial optimization. Examples include network flow, shortest path problem,transport problem, transshipment problem, location problem, matching problem, assignment problem, packing problem, routing problem, Critical Path Analysis and PERT (Program Evaluation & Review Technique).

Centrality measures

Information about the relative importance of nodes and edges in a graph can be obtained through centrality measures, widely used in disciplines like sociology. For example, eigenvector centrality uses the eigenvectors of the adjacency matrix to determine nodes that tend to be frequently visited.

Social network analysis maps relationships between individuals in social networks.[1] Such individuals are often persons, but may be groups (including cliques), organizations, nation-states,web sites, or citations between scholarly publications (scientometrics).

Network analysis, and its close cousin traffic analysis, has significant use in intelligence. By monitoring the communication patterns between the network nodes, its structure can be established. This can be used for uncovering insurgent networks of both hierarchical and leaderless nature.

Link analysis is a subset of network analysis, exploring associations between objects. An example may be examining the addresses of suspects and victims, the telephone numbers they have dialed and financial transactions that they have partaken in during a given timeframe, and the familial relationships between these subjects as a part of police investigation. Link analysis here provides the crucial relationships and associations between very many objects of different types that are not apparent from isolated pieces of information. Computer-assisted or fully automatic computer-based link analysis is increasingly employed by banks and insurance agencies in fraud detection, by telecommunication operators in telecommunication network analysis, by medical sector in epidemiology and pharmacology, in law enforcement investigations, by search engines for relevance rating (and conversely by the spammers for spamdexing and by business owners for search engine optimization), and everywhere else where relationships between many objects have to be analyzed.

Web link analysis

Several Web search ranking algorithms use link-based centrality metrics, including (in order of appearance) Marchiori's Hyper Search, Google's PageRank, Kleinberg's HITS algorithm, and theTrustRank algorithm. Link analysis is also conducted in information science and communication science in order to understand and extract information from the structure of collections of web pages. For example the analysis might be of the interlinking between politicians' web sites or blogs.

Spread of content in networks

Content in a complex network can spread via two major methods: conserved spread and non-conserved spread. In conserved spread, the total amount of content that enters a complex network remains constant as it passes through. The model of conserved spread can best be represented by a pitcher containing a fixed amount of water being poured into a series of funnels connected by tubes. Here, the pitcher represents the original source and the water is the content being spread. The funnels and connecting tubing represent the nodes and the connections between nodes, respectively. As the water passes from one funnel into another, the water disappears instantly from the funnel that was previously exposed to the water. In non-conserved spread, the amount of content changes as it enters and passes through a complex network. The model of non-conserved spread can best be represented by a continuously running faucet running through a series of funnels connected by tubes. Here, the amount of water from the original source is infinite. Also, any funnels that have been exposed to the water continue to experience the water even as it passes into successive funnels. The non-conserved model is the most suitable for explaining the transmission of most [[infectious diseases].

The origins: the seven bridges of Königsberg

The Seven Bridges of Königsberg is a famous historical problem in mathematics. Its 1736 negative resolution by Leonhard Euler laid the foundations of graph theory and presaged the idea of topology.

Description

The city of Königsberg in Prussia (now Kaliningrad, Russia) was set on both sides of the Pregel River, and included two large islands which were connected to each other and the mainland by seven bridges.

The problem was to find a walk through the city that would cross each bridge once and only once. The islands could not be reached by any route other than the bridges, and every bridge must have been crossed completely every time (one could not walk halfway onto the bridge and then turn around to come at it from another side).

Euler's analysis

It turns out that the problem has no solution.

To start with, Euler pointed out that the choice of route inside each landmass is irrelevant. The only important feature of a route is the sequence of bridges crossed. This allowed him to reformulate the problem in abstract terms (laying the foundations of graph theory), eliminating all features except the list of landmasses and the bridges connecting them. In modern terms, one replaces each landmass with an abstract "vertex" or node, and each bridge with an abstract connection, an "edge", which only serves to record which pair of vertices (landmasses) is connected by that bridge. The resulting mathematical structure is called a graph.

Figure. The seven bridges of Königsberg

Since only the connection information is relevant, the shape of pictorial representations of a graph may be distorted in any way without changing the graph itself. Only the existence (or lack) of an edge between each pair of nodes is significant. For example, it does not matter whether the edges drawn are straight or curved, or whether one node is to the left or right of another.

Next, Euler observes that (except at the endpoints of the walk) whenever one enters a vertex by a bridge, one leaves the vertex by a bridge. In other words, during any walk in the graph, the number times one enters a non-terminal vertex equals the number of times one leaves it. Now if every bridge is traversed exactly once it follows that for each landmass (except possibly for the ones chosen for the start and finish), the number of bridges touching that landmass is **even** (half of them will be traversed "toward" the landmass, the other half "away" from it). On the other hand, all the four landmasses in the original problem are touched by an **odd** number of bridges (one is touched by 5 bridges and the other three by 3). Since at most two landmasses can serve as the endpoints of a putative walk, the existence of a walk traversing each bridge once leads to a contradiction.

In modern language, Euler shows that the existences of a walk in a graph which traverses each edge once depends on the degrees of the nodes. The degree of a node is the number of edges touching it. Euler's argument shows that a walk of the desired form exists if and only if the graph is connected, and there are exactly zero or two nodes of odd degree. Such a walk is now called an *Eulerian path* or *Euler walk* in his honor. Further, if there are nodes of odd degree, all Eulerian paths start at one of them and end at the other. Since the graph corresponding to historical Königsberg has four nodes of odd degree, it cannot have an Eulerian path.

An alternative form of the problem asks for a path that traverses all bridges and also has the same starting and ending point. Such a walk is called an *Eulerian circuit* or an *Euler tour*. Such a circuit exists if and only if the graph is connected and there are no nodes of odd degree at all. Clearly Eulerian circuits are also Eulerian paths.

Euler's work was presented to the St. Petersburg Academy on August 26, 1735, and published as *Solutio problematis ad geometriam situs pertinentis* (The solution of a problem relating to the geometry of position) in the journal *Commentarii academiae scientiarum Petropolitanae* in 1741.[1] It is available in English in The World of Mathematics.

Significance in the history of mathematics

In the history of mathematics, Euler's solution of the Königsberg bridge problem is considered to be the first theorem of graph theory, a subject now generally regarded as a branch of combinatorics. Combinatorial problems of other types had been considered since antiquity.

In addition, Euler's recognition that the key information was the number of bridges and the list of their endpoints (rather than their exact positions) presaged the development of topology. The difference between the actual layout and the graph schematic is a good example of the idea that topology is not concerned with the rigid shape of objects.

Present state of the bridges

Two of the seven original bridges were destroyed by bombs during World War II. Two others were later demolished by the Russians and replaced by a modern highway. The three other bridges remain, although only two of them are from Euler's time (one was rebuilt by the Germans in 1935). Thus, in all, there are five bridges in modern-day Königsberg (modern name Kaliningrad).

In terms of graph theory, two of the nodes now have degree 2, and the other two have degree 3. Therefore, an Eulerian path is now possible, but since it must begin on one island and end on the other, it is impractical for tourists.

The random Networks of Erdős and Rényi

Paul Erdős (occasionally spelled Erdos or Erdös; Hungarian: Erdős Pál; March 26, 1913 – September 20, 1996) was an immensely prolific (and famously eccentric) Hungarian mathematician. With hundreds of collaborators, he worked on problems in combinatorics, graph theory, number theory, classical analysis, approximation theory, set theory, and probability theory.
His colleague Alfréd Rényi said, "a mathematician is a machine for turning coffee into theorems", and Erdős drank copious quantities. (This quotation is often attributed incorrectly to Erdős.) After 1971 he also took amphetamines, despite the concern of his friends, one of whom (Ron Graham) bet him $500 that he could not stop taking the drug for a month. Erdős won the bet, but complained during his abstinence that mathematics had been set back by a month: "Before, when I looked at a piece of blank paper my mind was filled with ideas. Now all I see is a blank piece of paper." After he won the bet, he promptly resumed his amphetamine habit.
Because of his prolific output, friends created the **Erdős number** as a humorous tribute; Erdős alone was assigned the Erdős number of 0 (for being himself), while his immediate collaborators could claim an Erdős number of 1, their collaborators have Erdős number at most 2, and so on. Some have estimated that 90% of the world's active mathematicians have an Erdős number smaller than 8 (not surprising in the light of the small world phenomenon). It is jokingly said that Baseball Hall of Famer Hank Aaron has an Erdős number of 1 because they both autographed the same baseball when Emory University awarded them honorary degrees on the same day. Erdős numbers have also been humorously assigned to an infant, a horse and several actors. For details see the "Extended Erdős Number Project".

Complex networks describe a variety of systems found in nature and society. Traditionally these

systems have been modeled as random graphs, a relatively primitive and brutal approach. These traditional models do not produce topological and structural properties featured in real network examples. In recent years many new models have been developed, to correctly describe the scale- free structure of real networks.
Traditionally the study of complex networks has been the territory of mathematics, especially the graph theory. Initially the graph theory focused on regular graphs, with no apparent design principles were described as random graphs, proposed as the simplest and most straightforward realization of a complex network.
The pioneer of the theory was Leonhard Euler, who studied first regular graphs in 18^{th} century. In the 20_{th} century the theory became much more statistically and algorithmically oriented.
Later in 1950's graph theory was used to describe large networks, with no particular distributions of nodes and link, whose organization principles were not easily definable. These networks were first studied by Paul Erdős and Alfred Rényi and were called "random graphs", due to their generating method: we start with N nodes and connect every pair of them with probability p. Obtained graph has on average $p\ (N(N\ 1))\ 2$ edges distributed randomly. The degree distribution of such graph is Poisson with peak at $P(\ k\ )$. This model has guided our thinking for decades after it has been presented.
The topology of **real world** large networks (i.e. Internet, WWW, telephone networks, ecological networks) substantially differs from the topology of random graphs produced by the simple Erdős-Rényi (ER) model, therefore new methods, tools and models needed to be developed.

In past years we witnessed dramatic advances in this direction. The computerisation of data acquisition has led to the emergence of large databases on the topology of various real networks. Wide availability of computer power allows to investigate networks containing millions of nodes, exploring questions that could not be answered before as well as the slow but noticeable breakdown between different science disciplines allows scientists to access different databases, allowing to uncover the generic properties of large networks.
Networks found in nature show degree distribution that greatly differs from the Poisson degree distribution of random graphs. Because of existence of a few vertices with high degree, the distribution of real networks has a power-law tail $P(k)\ k$, which indicates scale free properties.

The small Worlds of Watts and Strogatz, the six degrees of separation

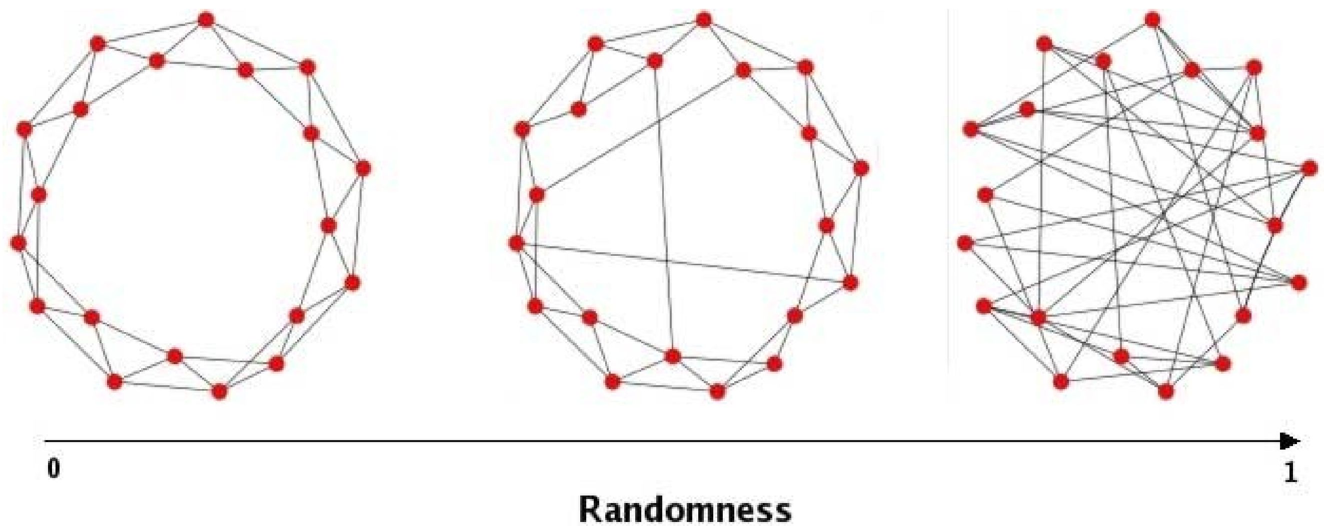


Figure. ***Small worlds*, between perfect order and chaos;**
the first graph is completely ordered, the graph in the middle is a "small world" graph, the graph at the right is complete random.

A network is called a small-world network by analogy with the small-world phenomenon (popularly known as six degrees of separation).The small world hypothesis, which was first described by the Hungarian writer Frigyes Karinthy in 1929, and tested experimentally by Stanley Milgram (1967), is the idea that two arbitrary people are connected by only six degrees of separation, i.e. the diameter of the corresponding graph of social connections is not much larger than six. In 1998, Duncan J. Watts and Steven Strogatz published the first small-world network model, which through a single parameter smoothly interpolates between a random graph to a lattice. Their model demonstrated that with the addition of only a small number of long-range links, a regular graph, in which the diameter is proportional to the size of the network, can be transformed into a "small world" in which the average number of edges between any two vertices is very small (mathematically, it should grow as the logarithm of the size of the network), while the clustering coefficient stays large. It is known that a wide variety of abstract graphs exhibit the small-world property, e.g., random graphs and scale-free networks. Further, real world networks such as the World Wide Web and the metabolic network also exhibit this property.

In the scientific literature on networks, there is some ambiguity associated with the term "small world." In addition to referring to the size of the diameter of the network, it can also refer to the co-occurrence of a small diameter and a high clustering coefficient. The clustering coefficient is a metric that represents the density of triangles in the network. For instance, sparse random graphs have a vanishingly small clustering coefficient while real world networks often have a coefficient significantly larger. Scientists point to this difference as suggesting that edges are correlated in real world networks.

Despite the large network size, it commonly happens that there is relatively short distance among any pair of nodes. Path length is defined by minimum number of edges needed to pass from first point to the other (in case of weighted edges, the path length is defined by minimal sum of weights). This phenomena is called the small world effect and can be observed in society and nature: all chemicals inside a living cell are at average 3 reactions away from each other, there is a path of acquaintances between most pairs of people in USA with typical length of about six and the actors in Hollywood are on average within three costars from each other.

All networks of the chapter 'illustrated regularities of the Pareto-Zipf-Mandelbrot (PZM) type' reveal small world properties. As shown below, any small world graph can be mapped on an Artificial Neural Network of the multilayer Perceptron with hidden layers and links arranged properly.

This allows us to put forward our main hypothesis, the equivalence between energy transformation systems (Birth and Death processor networks) and information transformation systems (networks of multilayer Perceptrons).

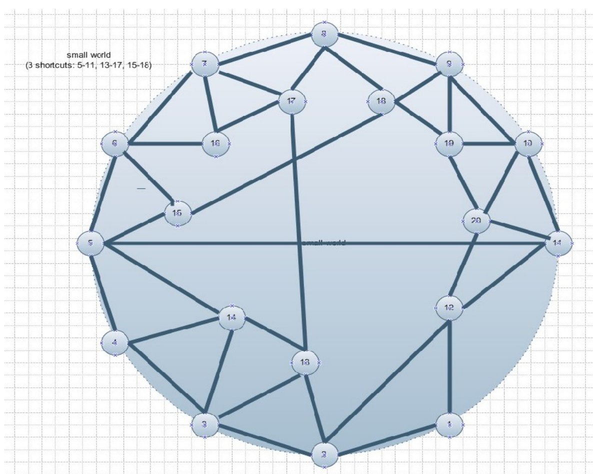


Figure. Example of a small world network mapped on a multilayer Perceptron see graph below.

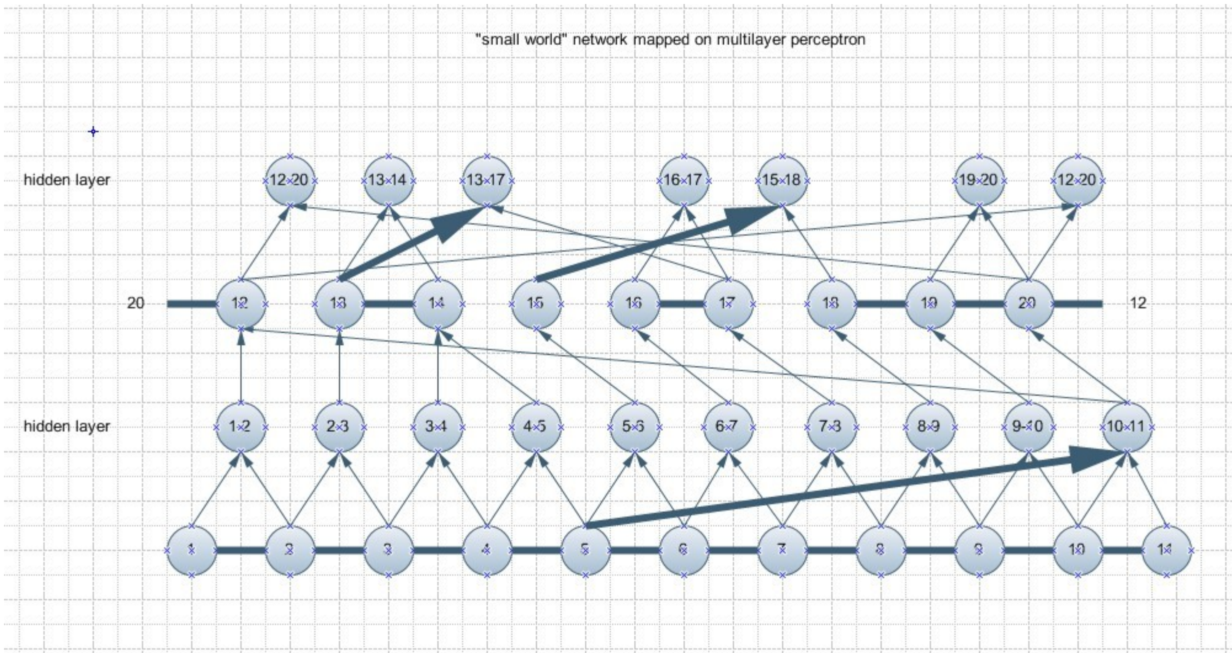


Figure. The shortcuts of the small world graph are mapped as fat arrow links in the multilayer Perceptron.

Barabási's scalefree networks from cells to the Internet

A network is named scale-free if its degree distribution, i.e., the probability that a node selected uniformly at random has a certain number of links (degree), follows a particular mathematical function called a power law. The power law implies that the degree distribution of these networks has no characteristic scale. In contrast, network with a single well-defined scale are somewhat similar to a lattice in that every node has (roughly) the same degree. Examples of networks with a single scale include the Erdős–Rényi random graph and hypercubes. In a network with a scale-free degree distribution, some vertices have a degree that is orders of magnitude larger than the average - these vertices are often called "hubs", although this is a bit misleading as there is no inherent threshold above which a node can be viewed as a hub. If there were, then it wouldn't be a scale-free distribution!

Interest in scale-free networks began in the late 1990s with the apparent discovery of a power-law degree distribution in many real world networks such as the World Wide Web, the network of Autonomous systems (ASs), some network of Internet routers, protein interaction networks, email networks, etc. Although many of these distributions are not unambiguously power laws, their breadth, both in degree and in domain, shows that networks exhibiting such a distribution are clearly very different from what you would expect if edges existed independently and at random (a Poisson distribution). Indeed, there are many different ways to build a network with a power-law degree distribution. The Yule process is a canonical generative process for power laws, and has been known since 1925. However, it is known by many other names due to its frequent reinvention, e.g., The Gibrat principle by Herbert Simon, the Matthew effect, cumulative advantage and, most recently, preferential attachment by Barabási and Albert for power-law degree distributions.

Networks with a power-law degree distribution can be highly resistant to the random deletion of vertices, i.e., the vast majority of vertices remain connected together in a giant component. Such networks can also be quite sensitive to targeted attacks aimed at fracturing the network quickly. When the graph is uniformly random except for the degree distribution, these critical vertices are the ones with the highest degree, and have thus been implicated in the spread of disease (natural and artificial) in social and communication networks, and in the spread of fads (both of which are modeled by a percolation or branching process).

Clustering

In many real examples of networks or graphs fully connected subgraphs emerge. Such structures are called cliques. A typical example of such feature are circles of friends or acquaintances in social networks where every member of a clique knows every other member. This inherent tendency of clustering is quantified by the clustering coefficient [Watts and Strogatz, 1998] and is defined for a single node in the network.

E_i is the number of all edges that actually exist among all first neighbor of selected node. If all the neighbors were connected, there would be $(1)/2 \; _{i\,i}k\;k$ edges among them. The ratio between the actual number of edges E_i and maximum number of edges is the clustering coefficient of a node.

The clustering coefficient of all the network is the average of all individual Ci's:
For random graphs the clustering coefficient is equal to graph generating connection probability ($C\ p$), since the probability of first neighbors being connected is constant for all nodes.
In real networks the clustering coefficient is much larger than in case of random graphs of equal size (equal number of nodes and edges).

Degree distribution

The number of edges a node has is called node degree. The spread of node degrees is characterized by a distribution function *P(k*), which gives the probability that randomly selected node has exactly k edges. Since in the random graph the edges are placed randomly, the majority of nodes have approximately the same degree, close to the average k of the network. The degree distribution of a random graph is a Poisson distribution with a peak at $P\ k$. On the other hand the empirical results for most large networks show distribution that significantly deviates from Poisson distribution. This degree distribution has a power-law tail.
Such network are called **scale free**. While some real networks still display an exponential tail, often the functional form of *P(k*) still deviates from Poisson distribution expected for a random graph.

Scale free model

Many large networks are scale free: their degree distribution follows a power law for large *k*. Even for those real networks for which *P(k)* has an exponential tail, the degree distribution significantly deviates from a Poisson. Random graph theory and the WS model are unable to reproduce this feature.
What is the mechanism responsible for the emergence of scale free networks?
A shift from modeling network topology to modeling the **network assembly and evolution** is required to get insight into mechanisms responsible to create scale-free networks.
While the goal of the other models (random graphs and small - world models) is to construct a graph with correct topological features, modeling scale free networks puts the emphasis on capturing the **network dynamics.** The assumption behind evolving or dynamic networks is that if we capture correctly the processes that assembled the networks that we see today, then we will obtain their topology correctly as well. Dynamics takes the driving role, topology being only a by product of this modeling philosophy.

Scale-free networks, Wikipedia

A **scale-free network** is a network whose degree distribution follows a power law, at least asymptotically. That is, the fraction $P(k)$ of nodes in the network having k connections to other nodes goes for large values of k as $P(k)$

~ $k^{-\gamma}$ where γ is a constant whose value is typically in the range $2<\gamma<3$, although occasionally it may lie outside these bounds.

Scale-free networks are noteworthy because many empirically observed networks appear to be scale-free, including the world wide web, protein networks, citation networks, and some social networks.

- Scale-free networks show a power law degree distribution like many real networks.
- The mechanism of preferential attachment has been proposed as a mechanism to explain power law degree distributions in some networks.

History

In studies of the networks of citations between scientific papers, Derek de Solla Price showed in 1965 that the number of links to papers—i.e., the number of citations they receive—had a heavy-tailed distribution following a Pareto distribution or power law, and thus that the citation network was scale-free. He did not however use the term "scale-free network" (which was not coined until some decades later). In a later paper in 1976, Price also proposed a mechanism to explain the occurrence of power laws in citation networks, which he called "cumulative advantage" but which is today more commonly known under the name preferential attachment.

Recent interest in scale-free networks started in 1999 with work by Albert-László Barabási and colleagues at the University of Notre Dame who mapped the topology of a portion of the Web (Barabási and Albert 1999), finding that some nodes, which they called "hubs", had many more connections than others and that the network as a whole had a power-law distribution of the number of links connecting to a node.

After finding that a few other networks, including some social and biological networks, also had heavy-tailed degree distributions, Barabási and collaborators coined the term "scale-free network" to describe the class of networks that exhibit a power-law degree distribution. Soon after, Amaral et al. showed that most of the real-world networks can be classified into two large categories according to the decay of P(k) for large k.

Barabási and Albert proposed a mechanism to explain the appearance of the power-law distribution, which they called "preferential attachment" and which is essentially the same as that proposed by Price. Analytic solutions for this mechanism (also similar to the solution of Price) were presented in 2000 by Dorogovtsev, Mendes and Samukhin and independently by Krapivsky, Redner, and Leyvraz, and later rigorously proved by mathematician Béla Bollobás. Notably, however, this mechanism only produces a specific subset of networks in the scale-free class, and many alternative mechanisms have been discovered since.

Although the scientific community is still debating the usefulness of the scale-free term in reference to networks, Li et al. (2005) recently offered a potentially more precise "scale-free metric". Briefly, let g be a graph with edge-set ε, and let the degree (number of edges) at a vertex i be d_i. Define

$$s(g) = \sum_{(i,j)\in\epsilon} d_i d_j.$$

This is maximised when high-degree nodes are connected to other high-degree nodes. Now define

$$S(g) = \frac{s(g)}{s_{max}}$$

where $S_{\max}$ is the maximum value of $s(h)$ for h in the set of all graphs with an identical degree distribution to g. This gives a metric between 0 and 1, such that graphs with low $S(g)$ are "scale-rich", and graphs with $S(g)$ close to 1 are "scale-free". This definition captures the notion of self-similarity implied in the name "scale-free".

Characteristics and examples

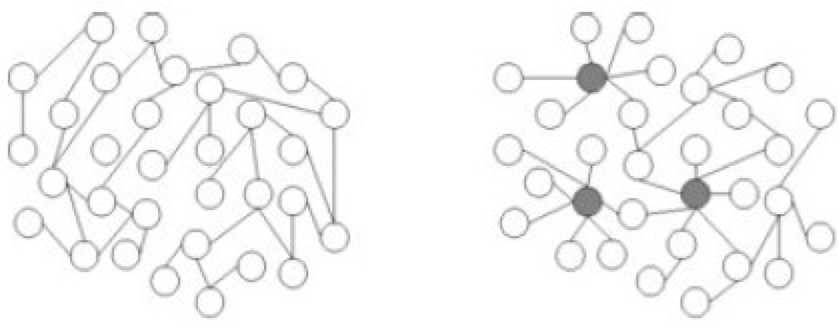


Figure. Random network left, scale-free network right. In the scale-free network, the larger hubs are highlighted.

As with all systems characterized by a power law distribution, the most notable characteristic in a scale-free network is the relative commonness of vertices with a degree that greatly exceeds the average. The highest-degree nodes are often called "hubs", and are thought to serve specific purposes in their networks, although this depends greatly on the domain.

The power law distribution highly influences the network topology. It turns out that the major hubs are closely followed by smaller ones. These ones, in turn, are followed by other nodes with an even smaller degree and so on. This hierarchy allows for a fault tolerant behavior. Since failures occur at random and the vast majority of nodes are those with small degree, the likelihood that a hub would be affected is almost negligible. Even if such event occurs, the network will not lose its connectedness, which is guaranteed by the remaining hubs. On the other hand, if we choose a few major hubs and take them out of the network, it simply falls apart and is turned into a set of rather isolated graphs. Thus hubs are both the strength of scale-free networks and their Achilles' heel.

Another important characteristic of scale-free networks is the clustering coefficient distribution, which decreases as the node degree increases. This distribution also follows a power law. That means that the low-degree nodes belong to very dense sub-graphs and those sub-graphs are connected to each other through hubs. Consider a social network in which nodes are people and links are acquaintance relationships between people. It is easy to see that people tend to form communities, i.e., small groups in which everyone knows everyone (one can think of such community as a complete graph). In addition, the members of a community also have a few acquaintance relationships to people outside that community. Some people, however, are so related to other people (e.g., celebrities, politicians) that they are connected to a large number of communities. Those people may be considered the hubs responsible for the small world phenomenon.

At present, the more specific characteristics of scale-free networks can only be discussed in either the context of the generative mechanism used to create them, or the context of a particular real-world network thought to be scale-free. For instance, networks generated by preferential attachment typically place the high-degree vertices in the middle of the network, connecting them together to form a core, with progressively lower-degree nodes making up the regions between the core and the periphery. Many interesting results are known for this subclass of scale-free networks. For instance, the random removal of even a large fraction of vertices impacts the overall connectedness of the network very little, suggesting that such topologies could be useful for security, while targeted attacks destroys the connectedness very quickly. Other scale-free networks, which place the high-degree vertices at the periphery, do not exhibit these properties; notably, the structure of the Internet is more like this latter kind of network than the kind built by preferential attachment. Indeed, many of the results about scale-free networks have been claimed to apply to the Internet, but are disputed by Internet researchers and engineers.

As with most disordered networks, such as the small world network model, the average distance between two vertices in the network is very small relative to a highly ordered network such as a lattice. The clustering coefficient of scale-free networks can vary significantly depending on other topological details, and there are now generative mechanisms that allow one to create such networks that have a high density of triangles.

It is interesting that Cohen and Havlin proved that uncorrelated power-law graph having $2 < \gamma < 3$ will also have ultrasmall diameter $d \sim \ln \ln N$. So from the practical point of view, the diameter of a growing scale-free network might be considered almost constant.

Although many real-world networks are thought to be scale-free, the evidence remains inconclusive, primarily because the generative mechanisms proposed have not been rigorously validated against the real-world data. As such, it is too early to rule out alternative hypotheses. A few examples of networks claimed to be scale-free include:

- Social networks, including collaboration networks. An example that has been studied extensively is the collaboration of movie actors in films.
- Protein-Protein interaction networks.
- Sexual partners in humans, which affects the dispersal of sexually transmitted diseases.
- Many kinds of computer networks, including the World Wide Web.
- Semantic networks. [1]

Generative models

These scale-free networks do not arise by chance alone. Erdős and Rényi (1960) studied a model of growth for graphs in which, at each step, two nodes are chosen uniformly at random and a link is inserted between them. The properties of these random graphs are not consistent with the properties observed in scale-free networks, and therefore a model for this growth process is needed.

The scale-free properties of the Web have been studied, and its distribution of links is very close to a power law, because there are a few Web sites with huge numbers of links, which benefit from a good placement in search engines and an established presence on the Web. Those sites are the ones that attract more of the new links. This has been called the winner takes allphenomenon.

The mostly widely known generative model for a subset of scale-free networks is Barabási and Albert's (1999) rich get richer generative model in which each new Web page creates links to existing Web pages with a probability distribution which is not uniform, but proportional to the current in-degree of Web pages. This model was originally discovered by Derek J. de Solla Pricein 1965 under the term **cumulative advantage**, but did not reach popularity until Barabási rediscovered the results under its current name (BA Model).According to this process, a page with many in-links will attract more in-links than a regular page. This generates a power-law but the resulting graph differs from the actual Web graph in other properties such as the presence of small tightly connected communities. More general models and networks characteristics have been proposed and studied (for a review see the book by Dorogovtsev and Mendes).

A different generative model is the **copy** model studied by Kumar et al. (2000), in which new nodes choose an existent node at random and copy a fraction of the links of the existent node. This also generates a power law.

However, if we look at communities of interests in a specific topic, discarding the major hubs of the Web, the distribution of links is no longer a power law but resembles more a normal distribution, as observed by Pennock et al. (2002) in the communities of the home pages of universities, public companies, newspapers and scientists. Based on these observations, they propose a generative model that mixes preferential attachment with a baseline probability of gaining a link.

The growth of the networks (adding new nodes) is not a necessary condition for creating a scale-free topology. Dangalchev (2004) gives examples of generating static scale-free networks. Another possibility (Caldarelli et al. 2002) is to consider the structure as static and draw a link between vertices according to a particular property of the two vertices involved. Once specified the statistical distribution for these vertices properties (fitnesses), it turns out that in some circumstances also static networks develop scale-free properties.

Recently, Manev and Manev (Med. Hypotheses, 2005) proposed that small world networks may be operative in adult brain neurogenesis. Adult neurogenesis has been observed in mammalian brains, including those of humans, but a question remains: how do new neurons become functional in the adult brain? It is proposed that the random addition of only a few new neurons functions as a maintenance system for the brain's "small-world" networks. Randomly added to an orderly network, new links enhance signal propagation speed and synchronizability. Newly generated neurons are ideally suited to become such links: they are immature, form more new connections compared to mature ones, and their number but not their precise location may be maintained by continuous proliferation and dying off. Similarly, it is envisaged that the treatment of brain pathologies by cell transplantation would also create new random links in small-world networks and that even a small number of successfully incorporated new neurons may be functionally important.

Real Networks: Empirical results

The study of most complex networks has been initiated by a desire to understand various real systems.
Complex systems that have been studied are:
1. **World Wide Web (WWW)**: Nodes are web pages and link are hyperlinks. The network is directed, but in some researches is made undirected Some of the researches are made on site level: All the pages in a site are merged into a supernode.
2. **Internet**: topology is studied at two different levels: at the router level the nodes are routers and edges are physical connections between them; at the interdomain level each domain, containing hundreds of routers, is represented as a single node. This is an undirected network.
3. **Cellular networks**: metabolisms of different species from all three domains of life are studied and organized into networks in which the substrates (ATP, ADP, H2O) are nodes and edges represent the predominantly directed chemical reactions in which these substrates can participate.
4. **Ecological networks or food webs**: the nodes are species and the edges represent predator-prey relationships among them. Food webs are directed networks.
5. **Protein folding**: Different states of single protein are represented by different nodes. Conformations are linked if they can be obtained from each other by an elementary move. This is an undirected network.
6. **Citation networks**: Published articles are represented by nodes and a directed edge represents a reference to a previously published article. This is an undirected network.
7. **Co authorship networks**: Collaboration network exists of scientists represented by nodes and two nodes are connected if two scientists have written an article together.
8. **Movie actor collaboration networks**: In this network the nodes are actors and two nodes have a common edge if two actors have acted in a movie together. This is an undirected network.
9. **The web of human sexual contacts**: Many sexually transmitted diseases spread on a network of sexual relationships. This is an undirected network.
10. **Phone-call networks**: A large directed graph can be constructed using telephone numbers as nodes and completed phone calls as edges, directed from caller to receiver.
11. **Networks in linguistics**: The complexity of human language offers several possibilities

to define and study complex networks. One way of building a network is to describe words as nodes and connect them with edges if they appear one word form each other inside sentences of the literature of certain language. This is an undirected network. The other way to construct a network is to link words bases on their meaning: words are represented as nodes and are linked by an edge id they are known to be synonyms. This is an undirected network as well.

12. **Power networks**: Power grid is described as an undirected network where nodes are generators, transformers and substations and the edges are high-voltage transmission lines.

13. **Neural networks**: Nerve systems of different animal species are studied. An undirected network nodes are neurons joined together by an edge if connected by either synapse or gap-junction.

Studies of complex systems stated above were performed by different scientists on different datasets of different network sizes, ranging from small networks with only few hundred nodes (ecological networks) to large networks with as many as 10^9 nodes like WWW. Studied networks are of both directed and undirected type. In researches the average path length among the nodes of a graph, clustering coefficient and degree distribution were measured and compared to the same properties of random graphs. For a estimation of clustering coefficient the directed networks need to be turned into undirected, since coefficient can only be calculated for undirected webs.

All the real networks mentioned in this section feature short average path lengths, large clustering coefficients and many of them have power-tail degree distribution and are scale free (WWW, cellular networks, Internet, some social networks and the citation networks). However, others like the power grid or the neural network appear to feature exponential or a coherent mixture of scale-free and exponential degree distributions. As it is shown these networks are far from being random like ER random graphs, these systems are best described by evolving networks and can therefore develop both power law and exponential degree distributions or a mixture of them. While the power law regime appears to be robust, sublinear preferential attachment, aging effects, growth constraints lead to crossovers to exponential decay.

The mysteries of Artificial Neural Networks

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.

Artificial Neural Networks have developed in a highly specialized technical field with thousands of publications. Below the rather technical article from Wikipedia. If you have to retain a single feature of ANNs its the one the Artificial Neural Networks can **learn**.

In more practical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

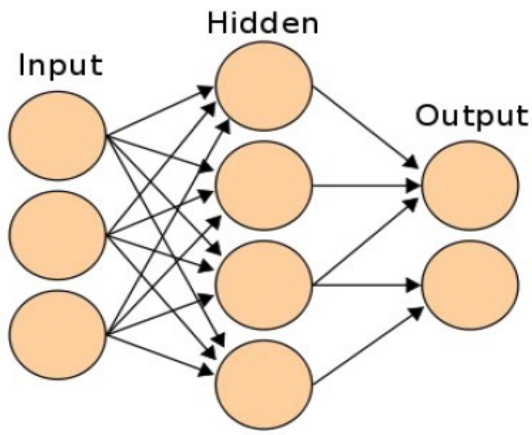


Figure. A neural network is an interconnected group of nodes, akin to the vast network of neurons in the human brain.

Background

There is no precise agreed-upon definition among researchers as to what a neural network is, but most would agree that it involves a network of simple processing elements (neurons), which can exhibit complex global behavior, determined by the connections between the processing elements and element parameters. The original inspiration for the technique was from examination of the central nervous system and the neurons (and their axons, dendrites and synapses) which constitute one of its most significant information processing elements (see Neuroscience). In a neural network model, simple nodes (called variously "neurons", "neurodes", "PEs" ("processing elements") or "units") are connected together to form a network of nodes — hence the term "neural network." While a neural network does not have to be adaptive per se, its practical use comes with algorithms designed to alter the strength (weights) of the connections in the network to produce a desired signal flow.

These networks are also similar to the biological neural networks in the sense that functions are performed collectively and in parallel by the units, rather than there being a clear delineation of subtasks to which various units are assigned (see also connectionism). Currently, the term Artificial Neural Network (ANN) tends to refer mostly to neural network models employed in statistics, cognitive psychology and artificial intelligence. Neural network models designed with emulation of the central nervous system (CNS) in mind are a subject of theoretical neuroscience (computational neuroscience).

In modern software implementations of artificial neural networks the approach inspired by biology has more or less been abandoned for a more practical approach based on statistics and signal processing. In some of these systems neural networks, or parts of neural networks (such as artificial neurons) are used as components in larger systems that combine both adaptive and non-adaptive elements. While the more general approach of such adaptive systems is more suitable for real-world problem solving, it has far less to do with the traditional artificial intelligence connectionist models. What they do, however, have in common is the principle of non-linear, distributed, parallel and local processing and adaptation.

Models

Neural network models in artificial intelligence are usually referred to as artificial neural networks (ANNs); these are essentially simple mathematical models defining a function $f: X \rightarrow Y$. Each type of ANN model corresponds to a *class* of such functions.

The *network* in *artificial neural network*

$$f(x) = K\left(\sum_i w_i g_i(x)\right)$$

The word *network* in the term 'artificial neural network' arises because the function $f(x)$ is defined as a composition of other functions $g_i(x)$, which can further be defined as a composition of other functions. This can be conveniently represented as a network structure, with arrows depicting the dependencies between variables. A widely used type of composition is the *nonlinear weighted sum*, where , where K is some predefined function, such as the hyperbolic tangent. It will be convenient for the following to refer to a collection of functions g_i as simply a vector $g = (g_1, g_2, \ldots g_n)$.

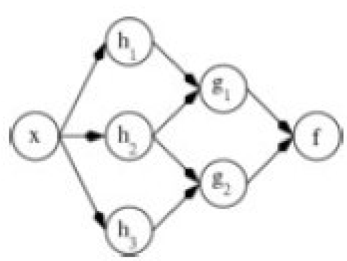


Figure. ANN dependency graph

This figure depicts such a decomposition of f, with dependencies between variables indicated by arrows. These can be interpreted in two ways.

The first view is the functional view: the input x is transformed into a 3-dimensional vector h, which is then transformed into a 2-dimensional vector g, which is finally transformed into f. This view is most commonly encountered in the context of optimization.

The second view is the probabilistic view: the random variable $F = f(G)$ depends upon the random variable $G = g(H)$, which depends upon $H = h(X)$, which depends upon the random variable X. This view is most commonly encountered in the context of graphical models.

The two views are largely equivalent. In either case, for this particular network architecture, the components of individual layers are independent of each other (e.g., the components of g are independent of each other given their input h). This naturally enables a degree of parallelism in the implementation.

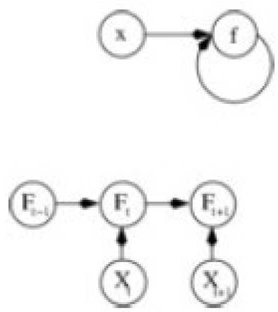


Figure. Recurrent ANN dependency graph

Networks such as the previous one are commonly called feedforward, because their graph is a directed acyclic graph. Networks with cycles are commonly called recurrent. Such networks are commonly depicted in the manner shown at the top of the figure, where f is shown as being dependent upon itself. However, there is an implied temporal dependence which is not shown.

Learning

However interesting such functions may be in themselves, what has attracted the most interest in neural networks is the possibility of ***learning***, which in practice means the following:

Given a specific *task* to solve, and a *class* of functions F, learning means using a set of *observations*, in order to find a function which solves the task in an *optimal sense*.

The cost function C is an important concept in learning, as it is a measure of how far away we are from an optimal solution to the problem that we want to solve. Learning algorithms search through the solution space in order to find a function that has the smallest possible cost.

$$\hat{C} = \frac{1}{N}\sum_{i=1}^{N}(f(x_i) - y_i)^2$$

For applications where the solution is dependent on some data, the cost must necessarily be a *function of the observations*, otherwise we would not be modeling anything related to the data. It is frequently defined as a statistic to which only approximations can be made. As a simple example consider the problem of finding the model f which minimizes C, for data pairs (x,y) drawn from some distribution D. In practical situations we would only have N samples from D and thus, for the above example, we would only minimize . Thus, the cost is minimized over a sample of the data rather than the true data distribution.

When N $\rightarrow \infty$ some form of on line learning must be used, where the cost is partially minimized as each new example is seen. While on line learning is often used when D is fixed, it is most useful in the case where the distribution changes slowly over time. In neural network methods, some form of online learning is frequently also used for finite datasets.

Choosing a cost function

While it is possible to arbitrarily define some ad hoc cost function, frequently a particular cost will be used either because it has desirable properties (such as convexity) or because it arises naturally from a particular formulation of the problem (i.e., In a probabilistic formulation the posterior probability of the model can be used as an inverse cost). **Ultimately, the cost function will depend on the task we wish to perform**. The three main categories of learning tasks are over viewed below.

Learning paradigms

There are three major learning paradigms, each corresponding to a particular abstract learning task. These are supervised learning, unsupervised learning and reinforcement learning. Usually any given type of network architecture can be employed in any of those tasks.

1. Supervised learning

In supervised learning, we are given a set of example pairs (x,y) with $x \in X$ and $y \in Y$ and the aim is to find a function $f: X \rightarrow Y$ in the allowed class of functions that matches the examples. In other words, we wish to *infer* the mapping implied by the data; the cost function is related to the mismatch between our mapping and the data and it implicitly contains prior knowledge about the problem domain.

A commonly used cost is the mean-squared error which tries to minimize the average squared error between the network's output, f(x), and the target value y over all the example pairs. When one tries to minimise this cost using gradient descent for the class of neural networks called Multi-Layer Perceptrons, one obtains the common and well-known backpropagation algorithm for training neural networks.

Tasks that fall within the paradigm of supervised learning are pattern recognition (also known as classification) and regression (also known as function approximation). The supervised learning paradigm is also applicable to sequential data (e.g., for speech and gesture recognition). This can be thought of as learning with a "teacher," in the form of a function that provides continuous feedback on the quality of solutions obtained thus far.

2. Unsupervised learning

In unsupervised learning we are given some data x, and the cost function to be minimized can be any function of the data x and the network's output, f.

The cost function is dependent on the task (what we are trying to model) and our *a priori* assumptions (the implicit properties of our model, its parameters and the observed variables).

As a trivial example, consider the model $f(x) = a$, where a is a constant and the cost $C = E[(x - f(x))2]$. Minimizing this cost will give us a value of a that is equal to the mean of the data. The cost function can be much more complicated. Its form depends on the application: For example in compression it could be related to the mutual information between x and y. In statistical modeling, it could be related to the posterior probability of the model given the data. (Note that in both of those examples those quantities would be maximized rather than minimized).

Tasks that fall within the paradigm of unsupervised learning are in general estimation problems; the applications include clustering, the estimation of statistical distributions, compression andfiltering.

3. Reinforcement learning

In reinforcement learning, data x is usually not given, but generated by an agent's interactions with the environment. At each point in time t, the agent performs an action yt and the environment generates an observation xt and an instantaneous cost ct, according to some (usually unknown) dynamics. The aim is to discover a *policy* for selecting actions that minimizes some measure of a long-term cost, i.e. the expected cumulative cost. The environment's dynamics and the long-term cost for each policy are usually unknown, but can be estimated.

ANNs are frequently used in reinforcement learning as part of the overall algorithm.

Tasks that fall within the paradigm of reinforcement learning are control problems, games and other sequential decision making tasks.

Learning algorithms

Training a neural network model essentially means selecting one model from the set of allowed models (or, in a Bayesian framework, determining a distribution over the set of allowed models) that minimises the cost criterion. There are numerous algorithms available for training neural network models; most of them can be viewed as a straightforward application of optimization theory and statistical estimation.

Most of the algorithms used in training artificial neural networks are employing some form of gradient descent. This is done by simply taking the derivative of the cost function with respect to the network parameters and then changing those parameters in a gradient-related direction.

Evolutionary methods, simulated annealing, and expectation-maximization and non-parametric methods are among other commonly used methods for training neural networks. See alsomachine learning.

Temporal perceptual learning relies on finding temporal relationships in sensory signal streams. In an environment, statistically salient temporal correlations can be found by monitoring the arrival times of sensory signals. This is done by the perceptual network.

Employing artificial neural networks

Perhaps the greatest advantage of ANNs is their ability to be used as an arbitrary function approximation mechanism which 'learns' from observed data. However, using them is not so straightforward and a relatively good understanding of the underlying theory is essential.

- Choice of model: This will depend on the data representation and the application. Overly complex models tend to lead to problems with learning.
- Learning algorithm: There are numerous tradeoffs between learning algorithms. Almost any algorithm will work well with the *correct hyperparameters* for training on a particular fixed dataset. However selecting and tuning an algorithm for training on unseen data requires a significant amount of experimentation.
- Robustness: If the model, cost function and learning algorithm are selected appropriately the resulting ANN can be extremely robust.

With the correct implementation ANNs can be used naturally in online learning and large dataset applications. Their simple implementation and the existence of mostly local dependencies exhibited in the structure allows for fast, parallel implementations in hardware.

Applications

The utility of artificial neural network models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical.

Real life applications

The tasks to which artificial neural networks are applied tend to fall within the following broad categories:

- Function approximation, or regression analysis, including time series prediction and modelling.
- Classification, including pattern and sequence recognition, novelty detection and sequential decision making.
- Data processing, including filtering, clustering, blind source separation and compression.

Application areas include system identification and control (vehicle control, process control), game-playing and decision making (backgammon, chess, racing), pattern recognition (radar systems, face identification, object recognition and more), sequence recognition (gesture, speech, handwritten text recognition), medical diagnosis, financial applications (automated trading systems), data mining (or knowledge discovery in databases, "KDD"), visualization and e-mail spam filtering.

Neural network software

Neural network software is used to simulate, research, develop and apply artificial neural networks, biological neural networks and in some cases a wider array of adaptive systems. See also logistic regression.

Types of neural networks

Feedforward neural network

The feedforward neural network was the first and arguably simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.

A multilayer feedforward network -also called multilayer Perceptron -with gradient descent for error backpropagation has been used in our case study on the ecosystem of lake Constance.

There exists a panoply of other network types to cite just a few:
Radial basis function (RBF) network, Kohonen self-organizing network or SOMs (self-organizing maps), Recurrent network, Hopfield network, Echo state network, Long short term memory network, Stochastic neural networks, Boltzmann machine, Modular neural networks, Committee of machines, Associative neural network (ASNN), Holographic associative memory, Instantaneously trained networks, Spiking neural networks, Dynamic neural networks, Cascading neural networks, Neuro-fuzzy networks, Compositional pattern-producing networks, One-shot associative memory ...

Theoretical properties

Computational power, universal Turing Machine

The multi-layer perceptron (MLP) is a **universal function approximator**, as proven by the Cybenko theorem. However, the proof is not constructive regarding the number of neurons required or the settings of the weights.

Work by Hava Siegelmann and Eduardo D. Sontag has provided a proof that a specific recurrent architecture with rational valued weights (as opposed to the commonly used floating point approximations) has the full power of a Universal Turing Machine[2] using a finite number of neurons and standard linear connections. They have further shown that the use of irrational values for weights results in a machine with trans-Turing power.

Capacity

Artificial neural network models have a property called 'capacity', which roughly corresponds to their ability to model any given function. It is related to the amount of information that can be stored in the network and to the notion of complexity.

Convergence

Nothing can be said in general about convergence since it depends on a number of factors. Firstly, there may exist many local minima. This depends on the cost function and the model. Secondly, the optimization method used might not be guaranteed to converge when far away from a local minimum. Thirdly, for a very large amount of data or parameters, some methods become impractical. In general, it has been found that theoretical guarantees regarding convergence are an unreliable guide to practical application.

Generalization and statistics

In applications where the goal is to create a system that generalizes well in unseen examples, the problem of overtraining has emerged. This arises in overcomplex or overspecified systems when the capacity of the network significantly exceeds the needed free parameters. There are two schools of thought for avoiding this problem: The first is to use cross-validation and similar techniques to check for the presence of overtraining and optimally select hyperparameters such as to minimize the generalization error. The second is to use some form of regularization. This is a concept that emerges naturally in a probabilistic (Bayesian) framework, where the regularization can be performed by selecting a larger prior probability over simpler models; but also in statistical learning theory, where the goal is to minimize over two quantities: the 'empirical risk' and the 'structural risk', which roughly correspond to the error over the training set and the predicted error in unseen data due to overfitting.

We have cited the whole “zoo” of neural networks just to illustrate how rich this field of research is.

In our case study of the ecosystem of lake Constance we used only the first mentioned type of neural network, a multilayer feed forward network with gradient descent error backpropagation also called multilayer Perceptron.

The **multilayer Perceptron** is also the main paradigm when mapping energy transformation networks of a trophic web to information transformation networks of an ANN.

To map hypercycle structure networks, complex network theory has not much to say on the subject. The models are limited to networks growing under the mechanisms of preferential linking. In the field of artificial neural networks Kohonen's self-organizing maps (SOMs) might be a way to get hypercycle structure into a harder grip.

Theoretical attempts to explain the PZM regularities: Birth and Death processors and Artificial Neural Networks

"We are all agreed that your theory is crazy. The question which divides us is whether it is crazy enough to have a chance of being correct. My own feeling is that is not crazy enough. "

Niels Bohr on Pauli's theory of elementary particles
from Arne A. Wyller's *The Planetary Mind*

In every discipline, for example in geography there exist a great number of theoretical models to “explain” the observed PZM regularities. There are almost as many models as there are authors and many of them arrive at a Pareto-Zipf distribution. However the model's assumptions are specific to the field. One speaks of spatial fields, central place hierarchy and so on, but these concepts can not be transposed to other fields for which we observe PZM regularities.

We therefor constrained ourself to speak only about general models which cover a variety of different fields of science.

Self-organized critically, Wikipedia

I cite "Evolution of Networks" by [Dorogovtsev and Mendes, 2003] the bible of network theory.
"Thus the architecture that is based on fat-tailed degree distributions, with the key role of strongly connected vertices (hubs), is very important.
Where does it come from? Is it a result of the imposition of some external will, a lucky product of special design? Does somebody create intentionally such an architecture?
The answer is no.
These structures are the direct result of the self-organization of networks. Hence, the evolution of networks turns out to be among numerous growth processes which have been studied by phycisists for many years. One can say, by definition, that scalefree networks are in a critical state. so, the problems of the network growth are directly related to self-oranized critically."

Self-organized criticality is one of a number of important discoveries made in statistical physics and related fields over the latter half of the 20th century, discoveries which relate particularly to the study of complexity in nature. For example, the study of cellular automata, from the early discoveries of Stanislaw Ulam and John von Neumann through to John Conway's Game of Life and the extensive work of Stephen Wolfram, made it clear that complexity could be generated as an emergent feature of extended systems with simple local interactions. Over a similar period of time, Benoît Mandelbrot's large body of work on fractals showed that much complexity in nature could be described by certain ubiquitous mathematical laws, while the extensive study of phase transitions carried out in the 1960s and '70s showed how scale invariant phenomena such as fractals and power laws emerged at the critical point between phases.

Bak, Tang and Wiesenfeld's 1987 paper linked together these factors: a simple cellular automaton was shown to produce several characteristic features observed in natural complexity (fractal geometry,1/f noise and power laws) in a way that could be linked to critical-point phenomena. Crucially, however, the paper demonstrated that the complexity observed emerged in a robust manner that did not depend on finely-tuned details of the system: variable parameters in the model could be changed widely without affecting the emergence of critical behavior (hence, *self-organized* criticality). Thus, the key result of BTW's paper was its discovery of a mechanism by which the emergence of complexity from simple local interactions could be *spontaneous* — and therefore plausible as a source of natural complexity — rather than something that was only possible in the lab (or lab computer) where it was possible to tune control parameters to precise values. The publication of this research sparked considerable interest from both theoreticians and experimentalists, and important papers on the subject are among the most cited papers in the scientific literature.

Due to BTW's metaphorical visualization of their model as a "sandpile" on which new sand grains were being slowly sprinkled to cause "avalanches", much of the initial experimental work tended to focus on examining real avalanches in granular matter, the most famous and extensive such study probably being the Oslo rice pile experiment. Other experiments include those carried out on magnetic-domain patterns, the Barkhausen effect and vortices in superconductors. Early theoretical work included the development of a variety of alternative SOC-generating dynamics distinct from the BTW model, attempts to prove model properties analytically (including calculating the critical exponents), and examination of the necessary conditions for SOC to emerge. One of the important issues for the latter investigation was whether conservation of energy was required in the local dynamical exchanges of models: the answer in general is no, but with (minor) reservations, as some exchange dynamics (such as those of BTW) do require local conservation at least on average. In the long term, key theoretical issues yet to be resolved include the calculation of the possible universality classes of SOC behaviour and the question of whether it is possible to derive a general rule for determining if an arbitrary algorithm displays SOC.

Alongside these largely lab-based approaches, many other investigations have centered around large-scale natural or social systems that are known (or suspected) to display scale-invariant behavior. Although these approaches were not always welcomed (at least initially) by specialists in the subjects examined, SOC has nevertheless become established as a strong candidate for explaining a number of natural phenomena, including: earthquakes (which, long before SOC was discovered, were known as a source of scale-invariant behavior such as the Gutenberg-Richter law describing the statistical distribution of earthquake sizes, and the Omori law describing the frequency of aftershocks); solar flares; fluctuations in economic systems such as financial markets (references to SOC are common in econophysics); landscape formation; forest fires; landslides; epidemics; and biological evolution (where SOC has been invoked, for example, as the dynamical mechanism behind the theory of "punctuated equilibria" put forward by Niles Eldredge and Stephen Jay Gould). Worryingly, given the implications of a scale-free distribution of event sizes, some researchers have suggested that another phenomenon that should be considered an example of SOC is the occurrence of wars. These "applied" investigations of SOC have included both attempts at modelling (either developing new models or adapting existing ones to the specifics of a given natural system), and extensive data analysis to determine the existence and/or characteristics of natural scaling laws.

The recent excitement generated by scale-free networks has raised some interesting new questions for SOC-related research: a number of different SOC models have been shown to generate such networks as an emergent phenomenon, as opposed to the simpler models proposed by network researchers where the network tends to be assumed to exist independently of any physical space or dynamics.

West's MinMax priciple for scaling laws

Geoffrey West from the Santa Fee Institute studied scaling laws during an entire life time, see his book Scaling in Biology [Brown, West, 2000].

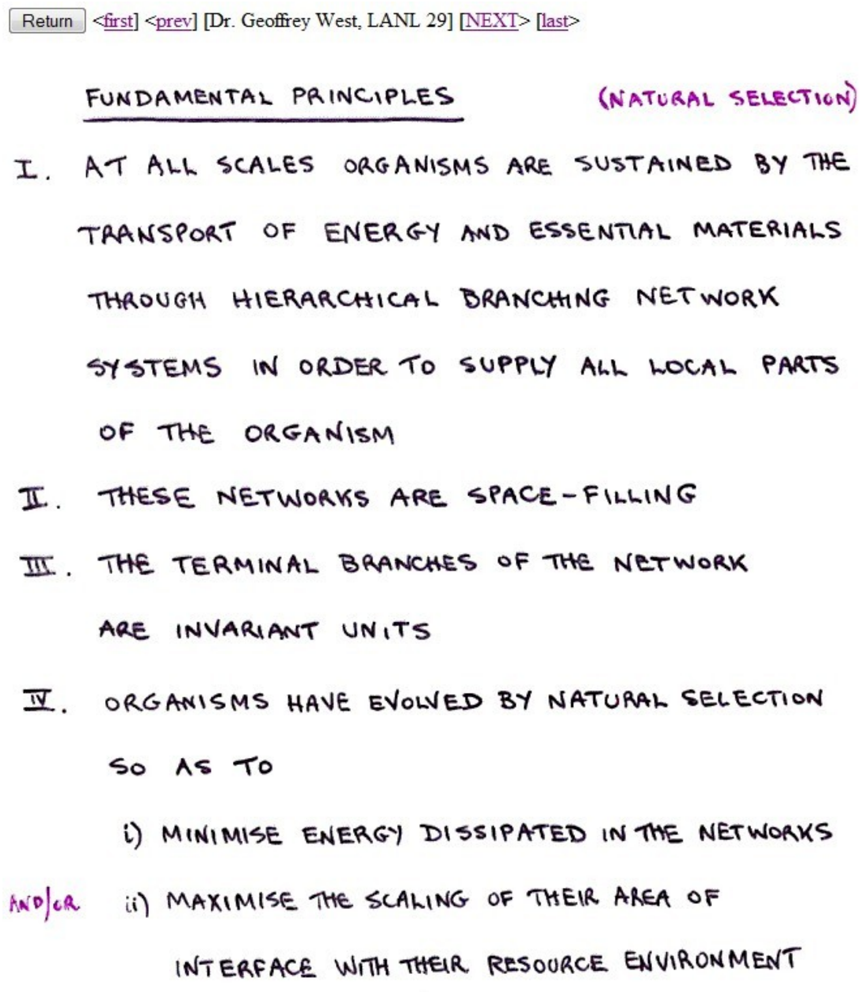
The first accurate measurements of body mass versus metabolic rate in 1932 shows that the metabolic rate R for all organisms follows exactly the 3/4 power-law of the body mass, i.e., $R \propto M^{3/4}$. This is known as the Kleiber's Law. It holds good from the smallest bacterium to the largest animal. The relation remains valid even down to the individual components of a single cell such as the mitochondrion, and the respiratory complexes (a subunit of the mitochondrion). It works for plants as well. This is one of the few all-encompassing principles in biology. But the law's universality is baffling: Why should so many species, with their variety of body plans, follow the same rules?
An explanation for this kind of relationship was proposed further back in 1883:

- Suppose the organism has a size of L, then the surface area $A \propto L^2$, while the volume $V \propto L^3$ assuming that it is in the shape of a sphere.
- If the density in the organism $\rho \propto M / L^3$ is constant, then $L \propto M^{1/3}$, where M is the total mass of the organism.
- Since the heat dissipation from an organism is proportional to its surface area, the total metabolic rate $R \propto L^2 \propto M^{2/3}$, which is close but not quite the same as the 3/4 power-law.

Then in 1997, a couple of physicist and biologists successfully derive the 3/4 power-law using the concept of fractal. The theory considers the fact that the tissues of large organisms have a supply problem. That is what blood systems in animals and vascular plants are all about: transporting materials to and from tissues. Small organisms don't face the problem to the same extent. A very small organism

has such a large surface area compared to its volume that it can get all the oxygen it needs through its body wall. Even if it is multicellular, none of its cells are very far from the outside body wall. But a large organism has a transport problem because most of its cells are far away from the supplies they need. Insects literally pipe air into their tissues in a branching network of tubes called tracheae. Mammals have richly branched air tubes, but they are confined to special organs, the lungs. Fish do a similar thing with gills. Trees use their richly dividing branches to supply their leaves with water and pump sugars back from the leaves to the trunk.
The 3/4-power law is derived in part from the assumption that mammalian distribution networks are "fractal like" and in part from the conjecture that natural selection has tended to maximize metabolic capacity "by maintaining networks that occupy a fixed percentage (6 - 7%) of the volume of the body". Effort has been made to derive the 3/4 power-law for a broader category that includes plants, animals, and even one-celled organisms lacking a vascular system. The latest derivation is based mostly on geometry, particularly the hierarchical nature of circulatory networks. It is argued that an organism's "internal area" -- the total area of its capillary walls -- fills up space so efficiently that it, in effect, adds a third dimension (similar to the compactification of extra dimensions in the Superstring Theory). Therefore, the "internal volume" of all the vessels feeding the capillaries acts as an extra dimension, scaling as the fourth power of internal length.

The figure below shows West's formulation of a minimum maximum principle which is at the base of the observed power law regularities in terms of hierarchical branching networks.



Holistic Extremum principle (Mandelbrot, Winiwarter)

Mandelbrot, the inventor of Fractals [Winiwarter, 1983b], has studied in detail the theory of coding and given an explanation for the regularities of word counts in terms of an extremum principle :
within a text, the quantity to be optimized (minimized) is the "average cost per word".
Assuming that the "cost" of a word depends on the "costs" of its constituting letters, Mandelbrot showed that the resulting "optimal" distribution is of the Pareto-Zipf type.
Based on a general principal of evolution or self-organization which states, that the complexity of a self-organized system can only grow or remain constant (first law of genesis)[Winiwarter], we put forward the hypothesis, that Pareto-Zipf type distributions are common to all processes of self-

organization (second law of genesis)[Winiwarter] .
Generalizing Mandelbrot's arguments from words to energy quanta, we speculated, that the observed Pareto-Zipf regularities are the result of a general extremum principle, which maximizes what we have called the energy redundancy (binding energy or synergy) within a self-organized system.
This approach seems very general and attractive, however - besides for systems of nucleons - it is difficult or impossible to verify.
Note that Mandelbrot's approach of a holistic extremum principle is very similar the the principle of maximum entropy production.

Pareto ⊕ Pareto = Pareto , stability under addition (Roehner, Winiwarter)

The Gaussian distribution is known to be a limit distribution of random variables .
It is well known, that the random sum ⊕ of two Gaussian distributions G1 and G2 yields a new distribution G3 which is also Gaussian.
$G1 \oplus G2 = G3$
It is too generally assumed, that this property is unique for the distributions called "normal", "bell-shaped" or Gaussian.
We have shown, that Pareto distributions are possible limit distributions of sums of random variables [Roehner, Winiwarter, 1985] .
The random sum ⊕ of two Paretian distributions P1 and P2 yields a new distribution P3 which is also Paretian .
$P1 \oplus P2 = P3$
Based on this statistical stability of Pareto distributions, we have explained the stability of empirical distributions as the result of a stochastic process :
$S_{t+1} = \alpha\ S_t \oplus \Delta$
The distribution at time t+l depends on the distribution at time t multiplied by a factor α characterizing the total growth of the system, plus a deviation Δ added at random.
If the initial distribution is Paretian and if the distribution of fluctuations Δ is Paretian, then the resulting distribution must also be Paretian.
This statistical stability is certainly an interesting and important feature, explaining the extreme perseverance of Pareto distributions over time, but it does not explain in a satisfactory way their origins. Stating that every observed regularity is the stochastic result of prior regularities, can be mathematically correct, but is not a very satisfying explanation.

Birth and Death processor, the basic interaction unit

It is the merit of Howard Odum [Odum 1988] to have focused our attention on the energy transformation aspect of hierarchically organized ecosystems. He already stated that " observing self-organization in nature suggests how energy is related to hierarchy and information .. . The details of the energy transformation mechanisms are quite different in ecosystems, chemical reaction systems,

turbulent hydrodynamical systems, social systems, and stars, but energy and mathematical characteristics are common to all".
All levels of the evolutionary hierarchy have two things in common: energy transformation and information transformation.
In a paper entitled “Life symptoms: the Behavior of Open Systems with Limited Energy Dissipation Capacity and Evolution” [Winiwarter and Cempel 1992], departing from a very specific model - describing tribo-vibro-acoustic processes in machines - we propose a generalized theoretical framework in terms of
energy transformation with limited internal energy dissipation capacity, which is applicable to all levels of the evolutionary hierarchy.
The proposed model "unifies" a large variety of concepts and applies a coherent terminology to fields, which have at first sight nothing in common. For the observed life symptoms, theoretical predictions can be compared with past and future empirical observations .
What is most important is the model's inference power: from the observations of a set of units at a given moment of life-time (a snapshot of the system), one can predict the average behavior of a single unit over its entire life-time .
The system is built of basic energy/information transformation processors that are born and run to death in an irreversible way (*birth and death processors*)

The model is characterized by energy input, upgraded as well as degraded energy outputs and a limited internal transformation capacity (see fig. 1 below as an example). In addition to the traditional energy flows our model is based on two very simple postulates

1) the internal accumulation potential is finite and irreversibly filled up to a threshold value.
2) the internal accumulation level regulates the internal accumulation rate through positive feedback

(autocatalytic behavior of internal accumulation) .
This very simple model results in important statistical features concerning the behavior of a single processor
over its life-time and the statistical behavior of a population of similar processors.

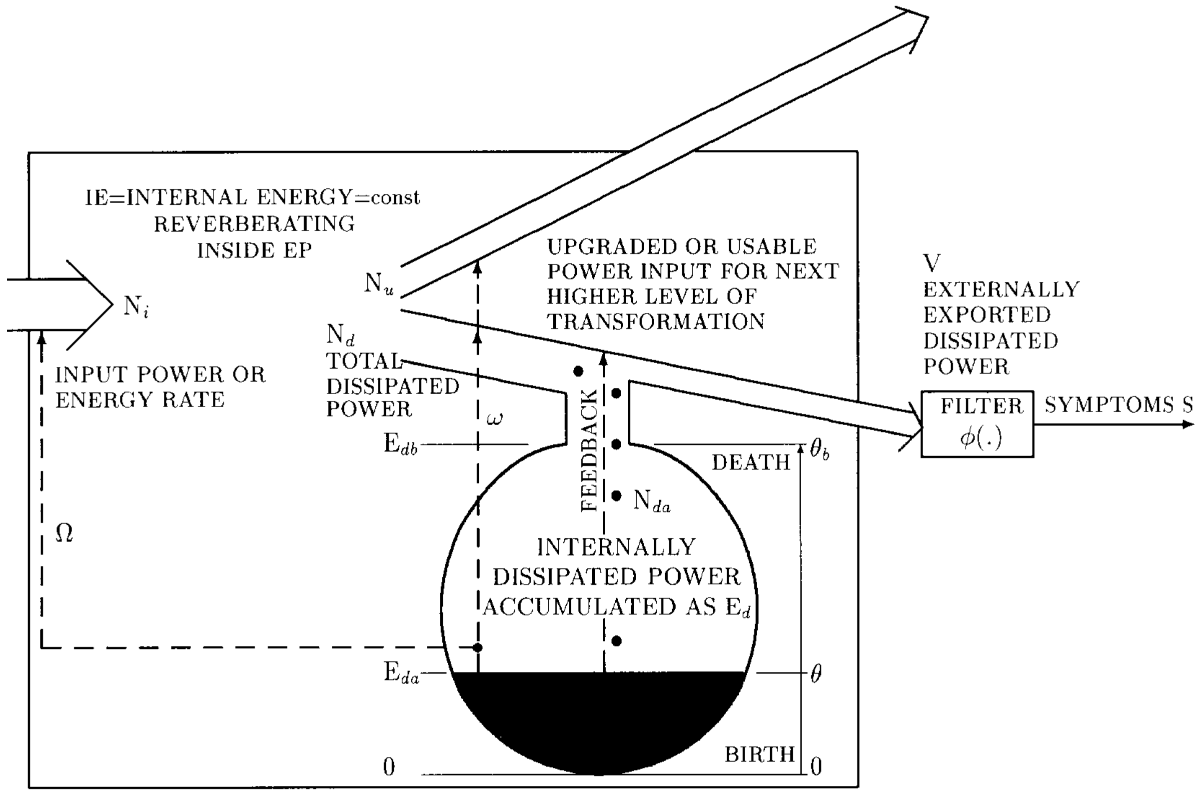
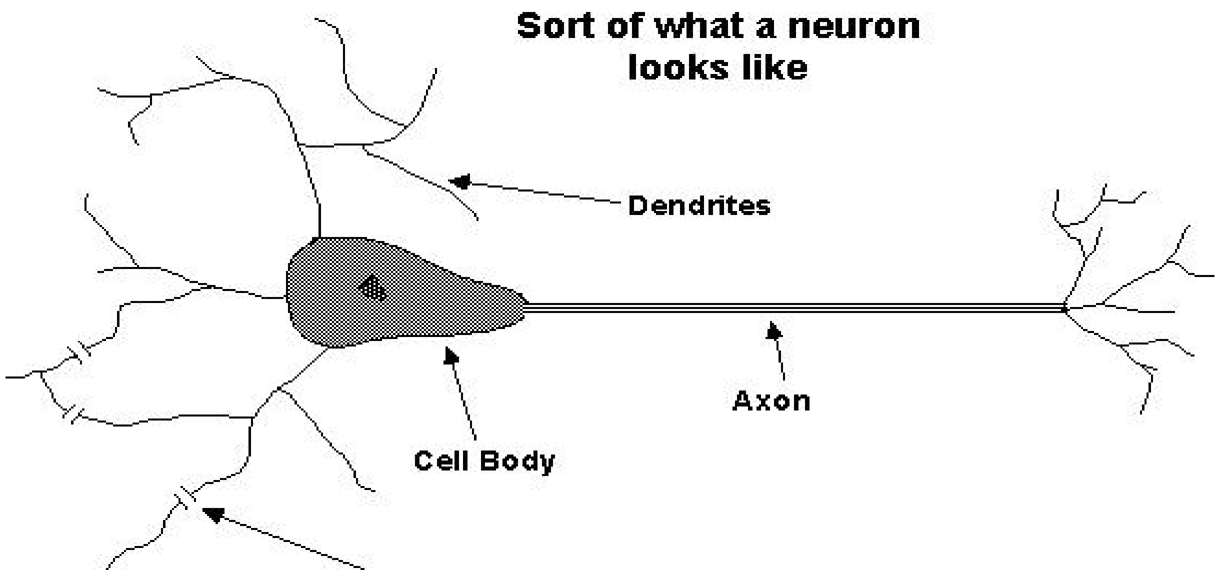
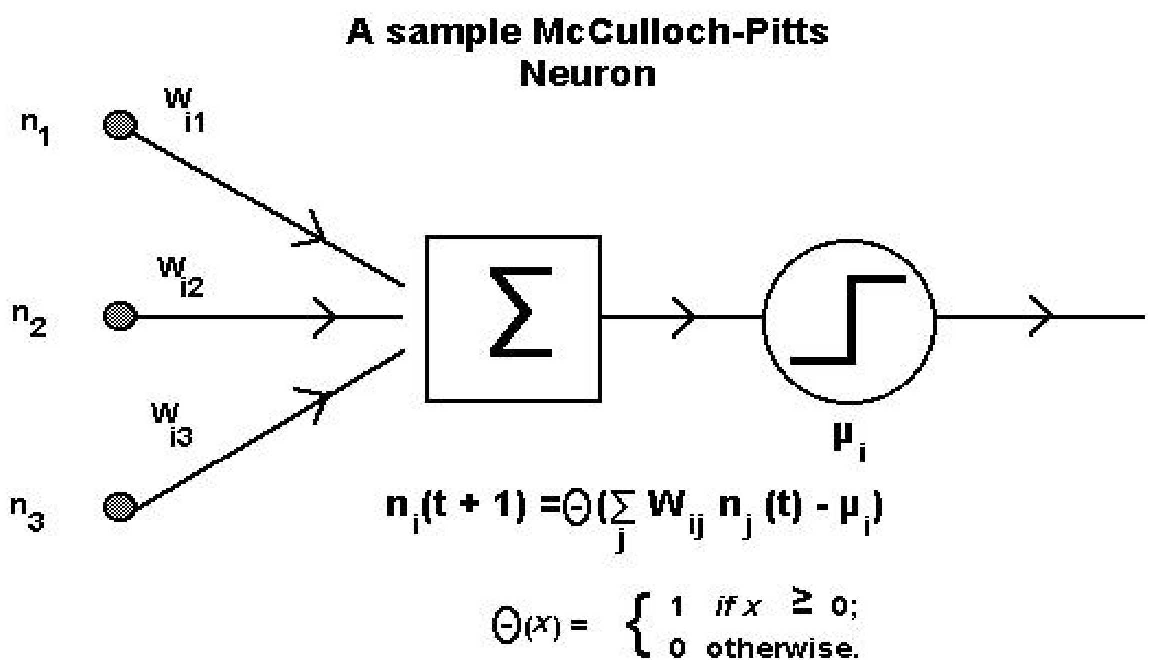


Figure . basic birth and death processor, the interaction unit of a self-organized system.
A part of the input energy rate N_i is irreversibly accumulated in an internal reservoir E_d from birth to death of the processor. Note the feedback of the internally accumulated energy and the output flow of dissipated downgraded energy observed as Symptom S. When the internal reservoir is full at E_{db}, the system brakes down (natural death).

Artificial Neuron equivalent to birth and death processor

An artificial neuron is a mathematical function conceived as a crude model, or abstraction of biological neurons. Artificial neurons are the constitutive units in an artificial neural network. Depending on the specific model used, it can receive different names, such as semi-linear unit, Nv neuron, binary neuron, linear threshold function or McCulloch-Pitts neuron. The artificial neuron receives one or more inputs (representing the one or more dendrites) and sums them to produce an output (synapse). Usually the sums of each node are weighted, and the sum is passed through a non-linear function known as an activation function or transfer function. The transfer functions usually have a sigmoid shape.





Networks of Birth and Death processors and Artificial Neural Networks

A self-similar network of Birth and Death processors, Energy transformation

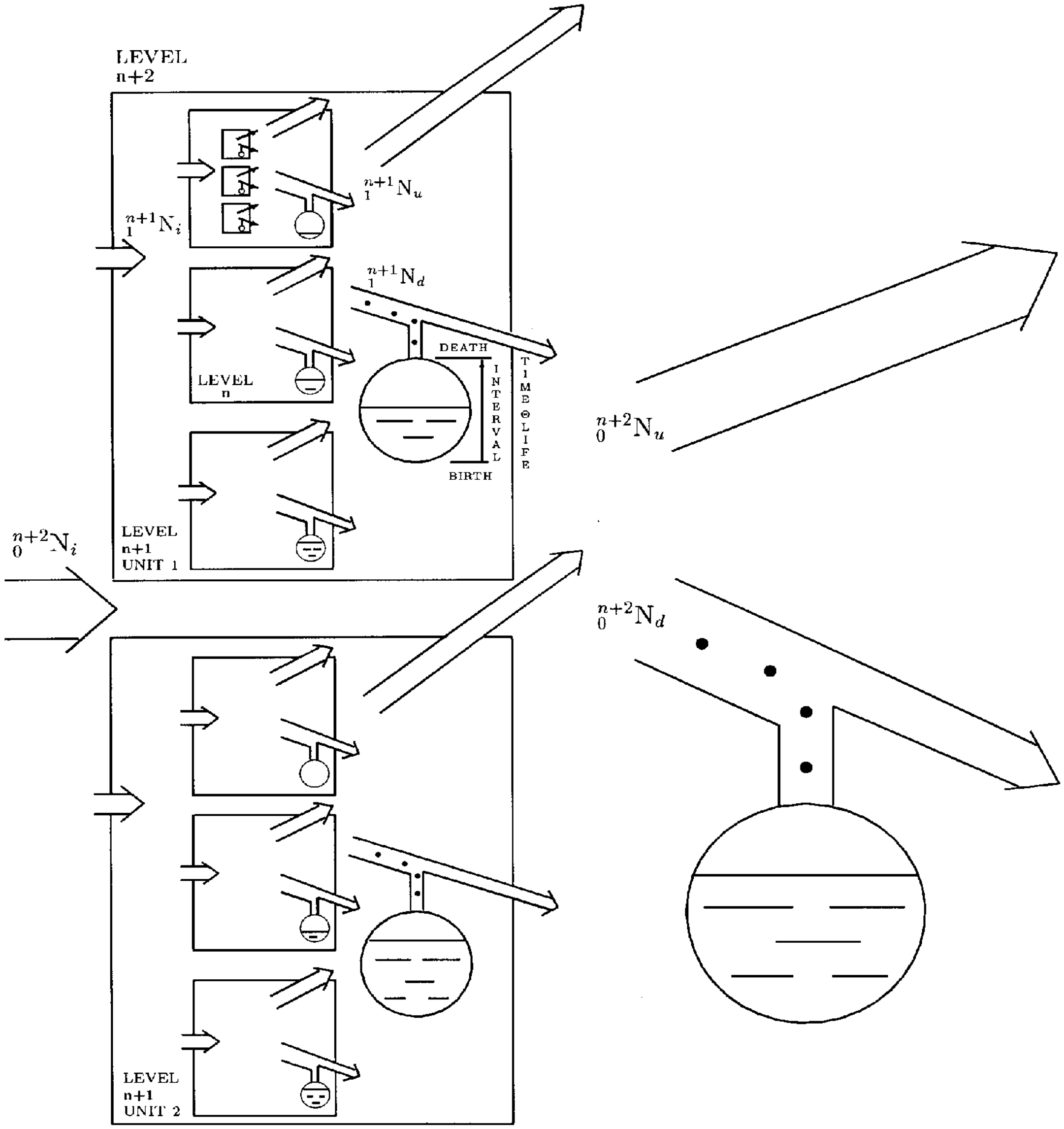


Figure. The self-similar 'fractal' hierarchy of energy transformation processors.

You can zoom in or zoom out of the self-similar hierarchy of energy transformation processors (Birth and Death processors). For a unit on any level of the hierarchy you will find the structure of a basic Birth and Death processor upgrading energy to the next level of the hierarchy, downgrading energy and internally accumulating downgraded energy until a threshold, it's death. Every unit can be considered as a binary threshold automation with two possible states : 0 processing or alive and 1 dead.

A self-similar network of Artificial Neural Networks, Information transformation

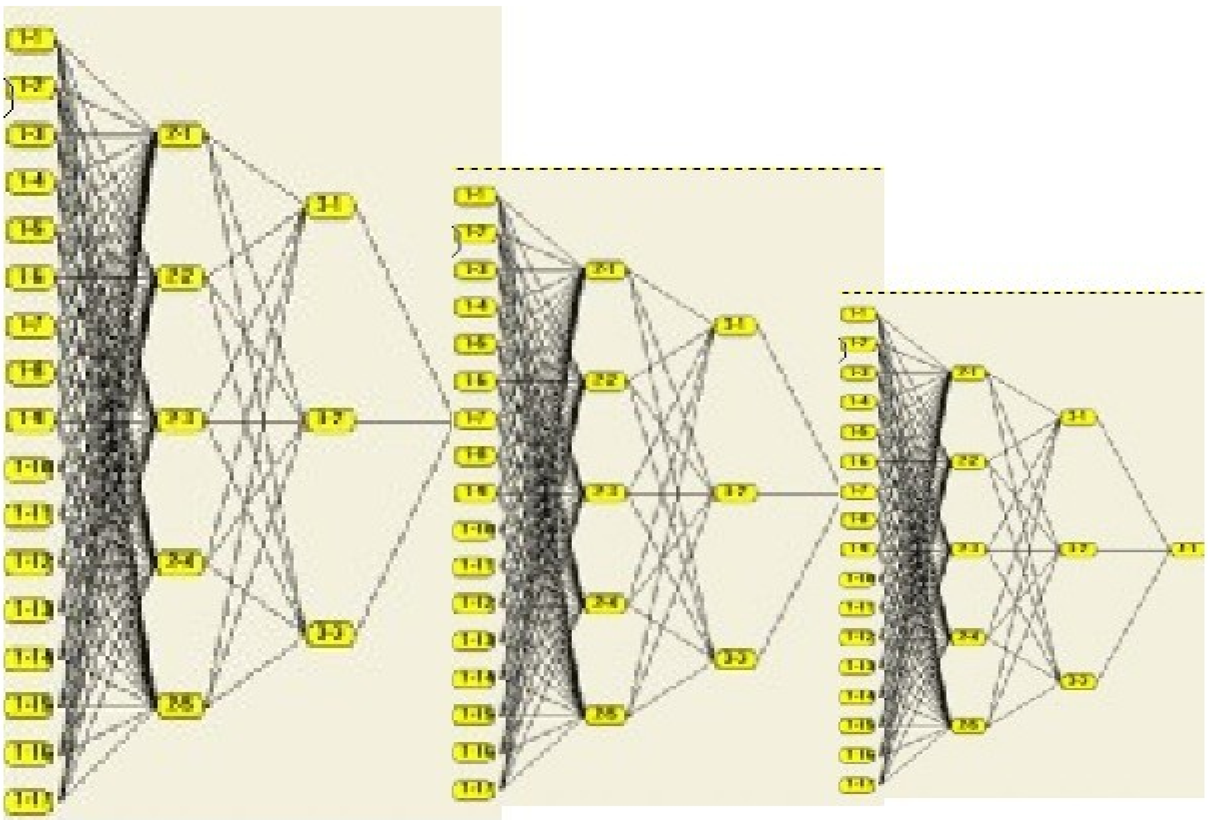


Figure. The fractal hierarchy of Neural Networks which maps the hierarchy of Energy transformation processors.

You can zoom in or zoom out of the self-similar hierarchy of information transformation processors (artificial neurons). For a unit on any level of the hierarchy you will find the structure of a basic neuron accumulating weighted incoming information pulses up to a threshold, when it fires a pulse to the next hierarchical level. Every unit can be considered as a binary threshold automaton with two possible states : 0 accumulating and 1 firing. This maps the binary threshold automaton of a Birth and Death processor.

Multilayer Perceptron, a robust universal mapper

This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a sigmoid function as an activation function.

The *universal approximation theorem* for neural networks states that every continuous function that maps intervals of real numbers to some output interval of real numbers can be approximated arbitrarily closely by a multi-layer perceptron with just one hidden layer.

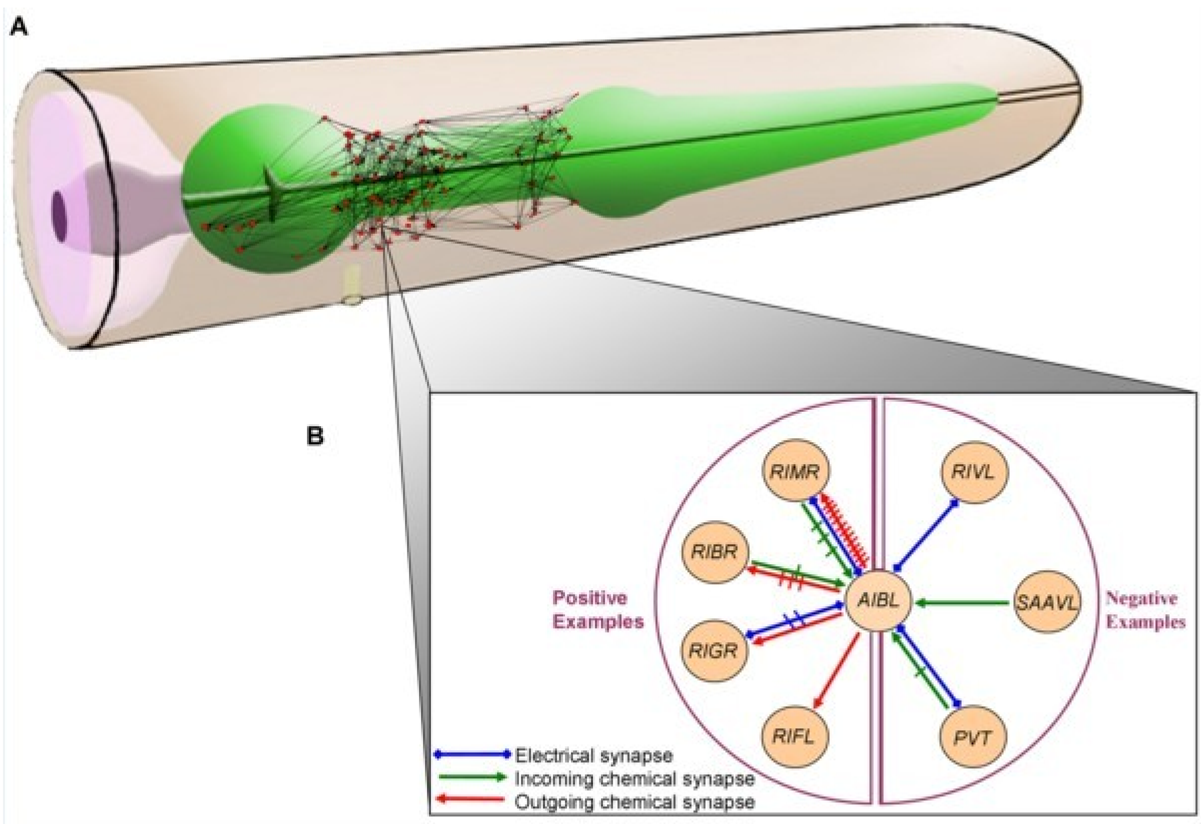
Multi-layer networks use a variety of learning techniques, the most popular being *back-propagation*. Here, the output values are compared with the correct answer to compute the value of some predefined error-function. By various techniques, the error is then fed back through the network. Using this information, the algorithm **adjusts the weights** of each connection in order to reduce the value of the error function by some small amount. After repeating this process for a sufficiently large number of training cycles, the network will usually converge to some state where the error of the calculations is small. In this case, one would say that the network has *learned* a certain target function. To adjust weights properly, one applies a general method for non-linear optimization that is called gradient descent. For this, the derivative of the error function with respect to the network weights is calculated, and the weights are then changed such that the error decreases (thus going downhill on the surface of the error function). For this reason, back-propagation can only be applied on networks with differentiable activation functions.

In general, the problem of teaching a network to perform well, even on samples that were not used as training samples, is a quite subtle issue that requires additional techniques. This is especially important for cases where only very limited numbers of training samples are available. The danger is that the network overfits the training data and fails to capture the true statistical process generating the data. Today there are practical solutions that make back-propagation in multi-layer perceptrons the solution of choice for many machine learning tasks.

Note that a major feature of an Artificial Neural Network of the perceptron type is its **robustness**. You can eliminate several or many nodes of the network, the remaining nodes will perform similar and show similar results. This feature called 'degeneracy' for biological networks can also be found within technical self-organized networks like the Internet. You can take out several nodes, but the routing will perform as prior to the incident.

Barabási [Barabási, 2003] already points out the major features of self-organized networks. “In reality, the market is nothing but a directed network. Companies, firms, corporations, financial institutions, governments, and all potential economic players are the nodes. Links quantify various interactions between these institutions, involving purchases and sales, joint research and marketing projects, and so forth. The ***weight*** of the links captures the value of the transaction, and the direction points from the provider to the receiver. The structure and evolution of this ***weighted*** and ***directed*** network determine the outcome of all macroeconomic processes.” (bold faces are ours) That is exactly the description of an artificial neural network of the perceptron type, a weighted and directed network.

Natural Neural Networks (C.elegans)



Trophic Web and Features of multilayer Perceptron (case study of lake Constance)

Box 4.1 The pelagic food web of Upper Lake Constance

Lake Constance is a large (500 km^2) and deep (z_{max} = 254 m) perialpine lake in central Europe, which has been intensively studied throughout the twentieth century. The lake consists of the more shallow Lower Lake Constance, and the deep Upper Lake Constance (Figure 4.1(a)). Due to its deep slope the latter has a truly pelagic zone, which seems to be energetically independent from littoral subsidies. Like many other temperate lakes, Lake Constance went through a period of severe eutrophication starting in the 1930s and culminating in the 1960s/1970s (Bäuerle and Gaedke 1998 and references therein). Beginning with the 1980s total phosphorus concentrations declined again. However, the response of the plankton community to oligotrophication was delayed. The pelagic food web of Upper Lake Constance during the oligotrophication period has been analyzed within several years of intensive sampling (Bäuerle and Gaedke 1998). Different food-web approaches, that is, body-mass size distributions (Gaedke 1992, 1993; Gaedke and Straile 1994*b*), binary food webs (Gaedke 1995), and mass-balanced flow networks (Gaedke and Straile 1994*a,b*; Straile 1995; Gaedke et al. 1996; Straile 1998; Gaedke et al. 2002) were applied to a dataset consisting of five, respectively eight (Gaedke et al. 2002) years of almost weekly sampling. A special strength of this dataset is that it encompasses both, the classical food chain as well as the microbial food web. For carbon flow models the pelagic food web was aggregated into eight different compartments (Figure 4.1(b)), of which five can be assigned to the "classical food chain," that is, phytoplankton, rotifers, herbivorous crustaceans, carnivorous crustaceans, and fish, and three to the microbial loop, that is, bacteria, heterotrophic nanoflagellates, and ciliates (Figure 4.1(b)). In addition flows between these eight compartments and the detritus/DOC (dissolved organic carbon) pool were considered (exsudation of phytoplankton, egestion and excretion of consumers, DOC uptake by bacteria). To analyze seasonal changes in carbon flows data were subdivided into up to 10 seasonal time intervals per year lasting between 14 and 102 days. For all seasonal time intervals mass-balanced carbon and phosphorous flows were established and further processed with the techniques of network analyses (Ulanowicz 1986). The results shown here are based on 44 different mass-balanced food-web diagrams for different seasonal time intervals from the study years 1987 to 1991 (Straile 1995, 1998).

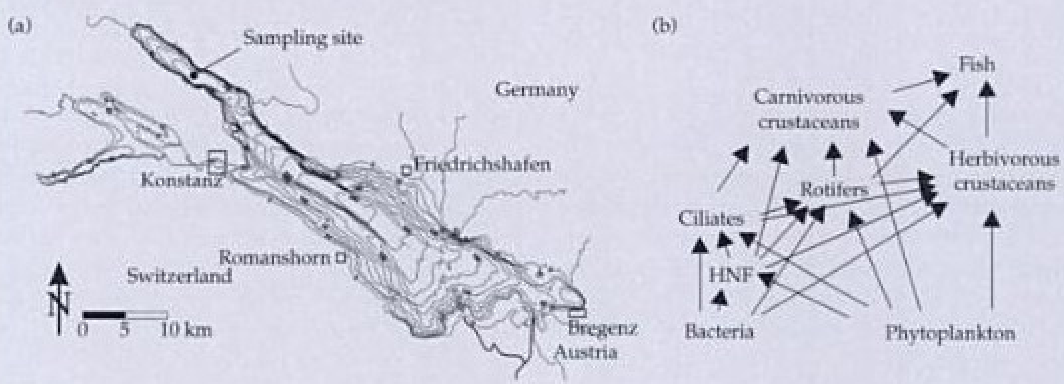
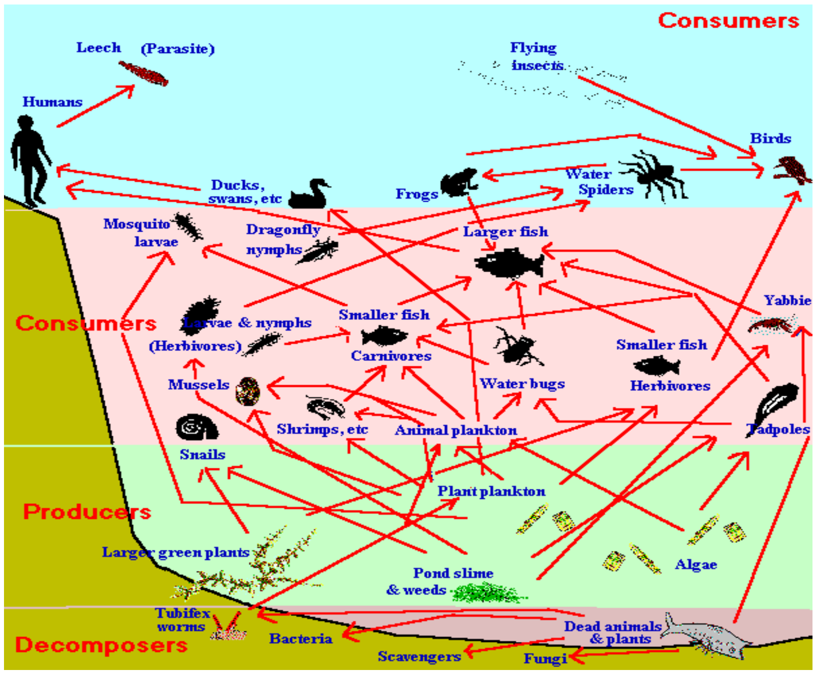


Figure 4.1 (a) Map of Lake Constance, and (b) aggregation of the Lake Constance pelagic food web into eight trophic guilds. Cannibalistic food-web interactions were considered for ciliates, rotifers, and carnivorous crustaceans, but are not shown here. Also not shown are the flows between these compartments and the detritus/DOC pool.



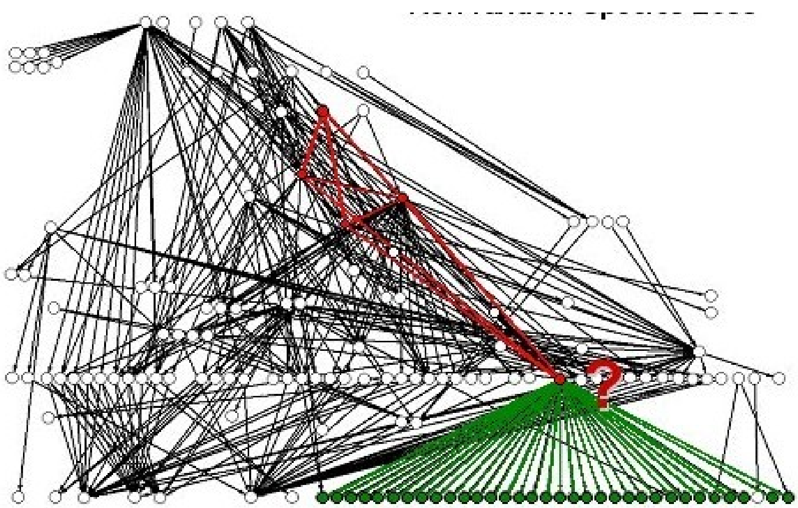


Figure. biomass-size distribution of aquatic ecosystems (trophic web or food web) Winiwarter and Vidondo modeled the ecosystem evolution of the lake Constance by a neural network of the feed forward type with back-propagation (multilayer perceptron).

To map the graph of the food web on an artificial neural networked we used a multilayer perceptron introducing an additional input layer, the time series of daily energy input from the sun (corrected for wind) day d, d-1, d-2 etc for one consecutive year. We used a single additional output layer, the coefficient alpha, slope of the observed biomass size distribution on day d.
The coefficient alpha shows strong seasonal variations which repeat each other year after year.
The model was trained with seven years of empirical data – biomass size distributions characterized by the slope alpha - yielding an excellent correlation coefficient. Forecasts with the model compared with subsequently observed data showed a high degree of coincidence between the model output and observed data.
This shows that a single parameter, the time series of daily energy input can describe the entire biodiversity of the system and its **seasonal evolution over time**.
What is remarkable is that the slope of the biomass size distribution is independent of the specific species being part of the size distribution. If one species decreases or disappears during a season or from the overall system it is replaced by other species and the slope of the distribution acts as an attractor which drives the temporal short and long term evolution of the system.
Note that the introduction of water purification stations did not alter biodiversity of the system but it did not alter its overall dynamics.
The novelty of this approach consists of mapping a complex trophic web on a simplified Artificial Neural Network (Perceptron) which allows to better understand the robustness and dynamics of the complex networks.
It is the topology of the trophic web and the weights of its interaction links which make up its memory and the consecutive daily runs – the learning process of the web – which allows its robustness.

Following the evolution of complex network theory we observe according to Barabási the stages of:

- simple random graph (Erdős and Rényi) a static viewpoint
- small world networks (Watts and Strogatz) explaining the famous six degrees between random chaos and complete order
- scale-free networks (Barabási)
  introducing dynamics explaining the existence of hubs and the observation of power laws in terms of network growth and preferential attachment.
  and finally our contribution
- artificial neural networks (Winiwarter)
  explaining the memory, learning and intelligence (robustness) of complex weighted and directed self-organized networks.

Future evolution: is the singularity near?

“Who will be man's successor? To which the answer is: We are ourselves creating our own successors. Man will become to the machine what the horse and the dog are to man; the conclusion being that machines are, or are becoming, animate.”

Samuel Butler, 1863 letter “Darwin among the machines?”

Increase in complexity, the first law of genesis (Winiwarter)

There seems to be a general agreement among scientists that during the evolution from the big bang to the world wide web the complexity of the observed systems increases from atoms over molecules, unicellular and pluricellular organisms to biological neural networks etc.

However this general statement of increasing complexity is difficult, if not impossible to express in quantitative terms. So far there exists no overall agreement on a measure of complexity.

In a speculative paper [Winiwarter 1983a] we tentatively defined a quantitative measure of complexity which is measurable at least in the realm of nuclear physics. The observation of this complexity during the nuclear evolution in a massive star (building up more and more heavy elements from Hydrogen over Carbon and Oxygen to Uranium) and the natural radioactive decay of heavy elements suggests a regularity, the first law of genesis:

The complexity of a self-organized system of matter can only increase or remain constant.

To extend this law from the nuclear realm to bio-chemistry, biology and sociology is a speculation, difficult to prove, since there are no empirical measures of the “binding energy” or synergy betweeen interaction units in these fields. However we observe, that with the emergence of a new level the binding energy decreases often by orders of magnitude. Strong nuclear bonds, weak nuclear bonds, chemical bonds, bonds of the genetic network, social bonds, links between sites of the World Wide Web...

PZM power laws, the second law of genesis (Winiwarter)

As illustrated in the chapter on observed Pareto-Zipf-Mandelbrot (PZM) distributions we observe regularities of the PZM type for virtually all levels of the evolutionary hierarchy. In 1983 when most of the observations like the World Wide Web did not exist we postulated on a speculative basis a general law, the second law of genesis:

Any self-organized system reveals Pareto-Zipf regularity for its statistical structure.

Barabási pointed out that every time we observe a power law the underlying network shows a specific topology, a hierarchical organization in local modules linked to global hubs.

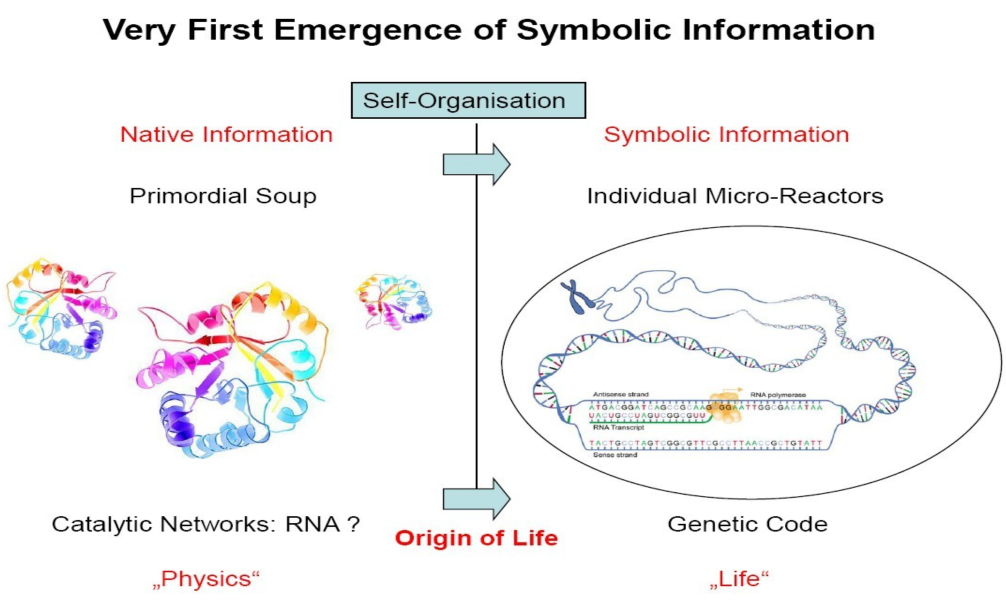
We propose the hypothesis that PZM regularities are not only characteristic for the topology of complex self-organized network but also for common processes within the network like bottom up feed forward of information and top down back-propagatio of information, which can be modeled by Artificial Neural Networks of the multilayer perceptron type.

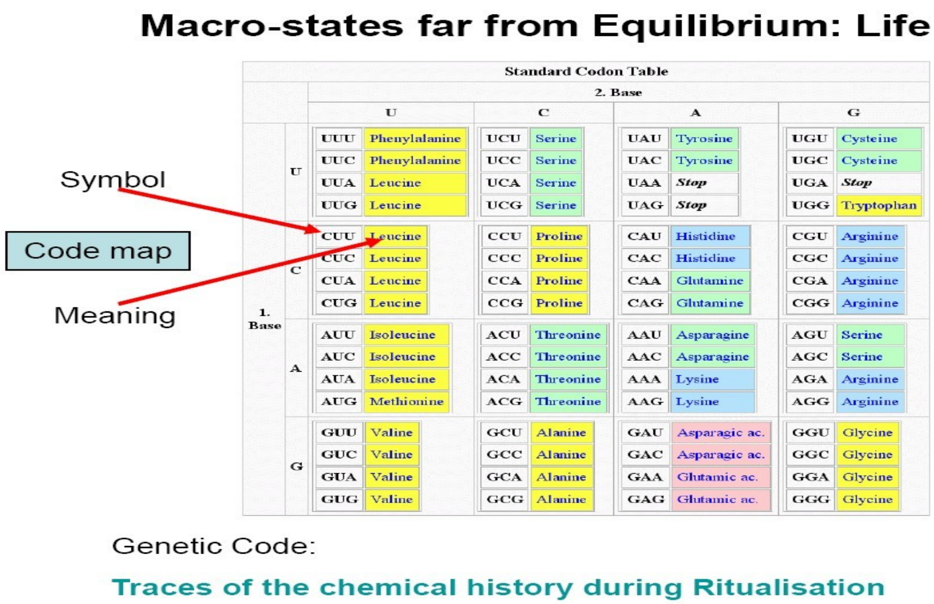
The literature on the observation of PZM regularities of complex graphs ranges from power grid systems, natural neural networks, protein interaction maps, metabolic pathways, ecological networks, electronic circuits, Internet topology, scientific collaborations to lexical networks.

Ritualization : the Self-Organization process of symbolic information

In our evolutionary hierarchy, the Russian dolls of communication networks, each new level emerges within the prior level. The network of cellular energy communication emerges within the network of chemical bonds communication. The genetic communication network of DNA emerges within the network of RNA communication. The network of the central nervous system (neural network) emerges within the network of the genetic code.

What is remarkable is that every emergence of a new level seems like a symmetry brake in the process of evolution where the new emerging level maps in a symbolic way the communication networks from which it emerges. This is actually the case for multilayer Artificial Neural Networks, where each higher level maps the respective lower level.

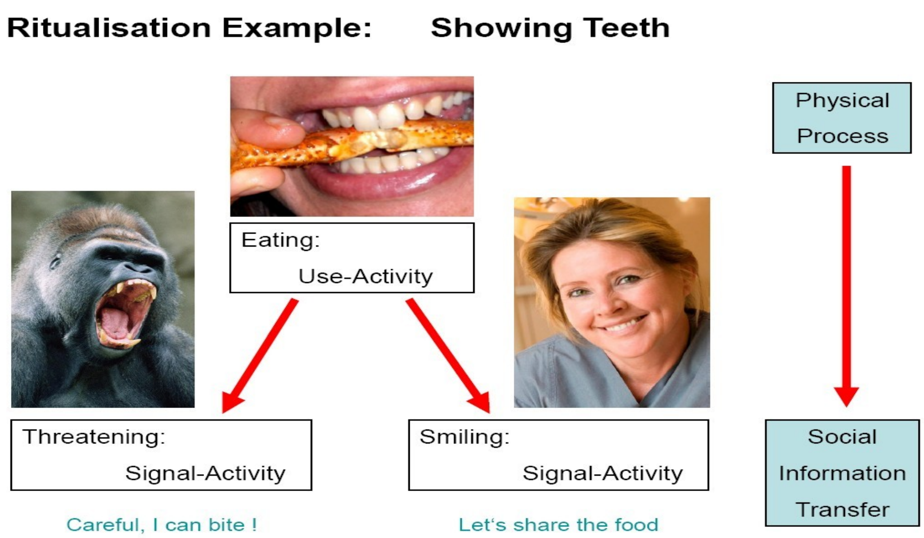




Standard Codon Table

1. Base	2. Base: U		C		A		G	
U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine
	UUC	Phenylalanine	UCC	Serine	UAC	Tyrosine	UGC	Cysteine
	UUA	Leucine	UCA	Serine	UAA	*Stop*	UGA	*Stop*
	UUG	Leucine	UCG	Serine	UAG	*Stop*	UGG	Tryptophan
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine
	CUC	Leucine	CCC	Proline	CAC	Histidine	CGC	Arginine
	CUA	Leucine	CCA	Proline	CAA	Glutamine	CGA	Arginine
	CUG	Leucine	CCG	Proline	CAG	Glutamine	CGG	Arginine
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine
	AUC	Isoleucine	ACC	Threonine	AAC	Asparagine	AGC	Serine
	AUA	Isoleucine	ACA	Threonine	AAA	Lysine	AGA	Arginine
	AUG	Methionine	ACG	Threonine	AAG	Lysine	AGG	Arginine
G	GUU	Valine	GCU	Alanine	GAU	Asparagic ac.	GGU	Glycine
	GUC	Valine	GCC	Alanine	GAC	Asparagic ac.	GGC	Glycine
	GUA	Valine	GCA	Alanine	GAA	Glutamic ac.	GGA	Glycine
	GUG	Valine	GCG	Alanine	GAG	Glutamic ac.	GGG	Glycine

Feistel calls this process of the emergence of symbols “ritualization” in analogy to the emergence of behavior in ethology.



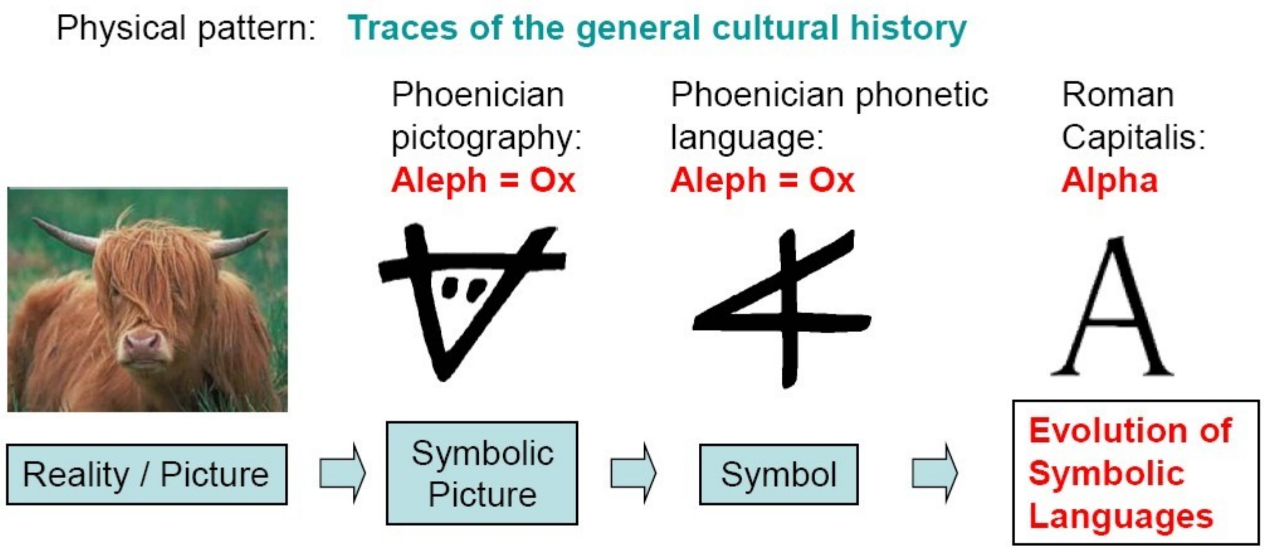
The symbol “showing teeth” stands for the communication “Let's share food”. It is only during the process of communication that the symbol acquires meaning. The symbol itself without the context is meaningless.

This phase transition takes place within the Ecosystem (central nervous system communication network) with the emergence of a semiotic communication network in social communities.

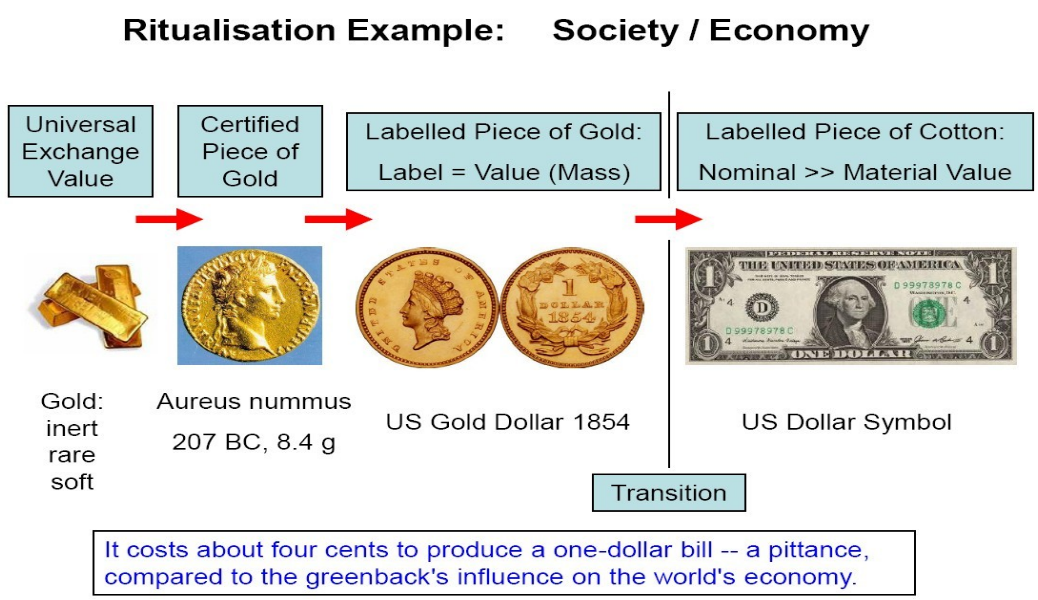
Likewise we observe the process of Ritualization during the emergence of our modern alphabet.

Within the ritual verbal symbolic communication network of culture emerges the level of mechanical tools communication network which gives rise to a written verbal communication network.

Note that the symbols of letters acquire meaning only within the context of words. The attentive reader will remember, that word frequency distributions of all languages and all time reveal a PZM regularity called Zipf's law.



If we continue to the next levels in our evolutionary hierarchy we observe the emergence of formal symbolic communication networks for economic exchange. The emerging symbols range from a piece of gold over simple certified coins to face value coins and finally to the highly abstract dollar bill which lost its link to gold value only recently.



The attentive reader will recall that the statistical structure of personal fortunes and of firm sizes expressed in monetary values reveal PZM regularities.

Extending the operator hierarchy of Gerard Jagers from the biological to the economic realm we can define the following "operators":

- **Cell** = multi-atom unit with the exchange of chemical compounds
  The closure is defined as the cell membrane.
- **Memon** = multi-cellular unit with a hardwired neural network with the exchange of perceptions
  The closure is defined by the physical organs of perception.
- **Oikos** = multi-memon unit with a hard structured home (greek 'oikos' means fireplace or home) with the exchange of basic goods like food and fuel for the fireplace
  The closure is defined by the physical walls of the home.
- **Market** = multi-oikos unit with a physical market place with the exchange of physical goods
  The closure is defined by the physical frontiers of the market .
- **Stock market** = multi-market with a virtual market place of the stock market where stocks and bonds are traded
  The closure is defined by virtual frontiers essentially the currency of the stock-market.

The singularity is near (Kurzweil)

Different sources of evidence and theoretical arguments indicate the technological innovation shares some basic traits with the patterns displayed by biological novelty. The rise and fall of technological creations also resembles the origination and extinction patterns observable in some groups of organisms and Jacques Monod actually suggested that the evolution of technology is sometimes closer to Darwinian selection than biology itself.

Foreword to The Intelligent Universe of James Gardner

by Ray Kurzweil

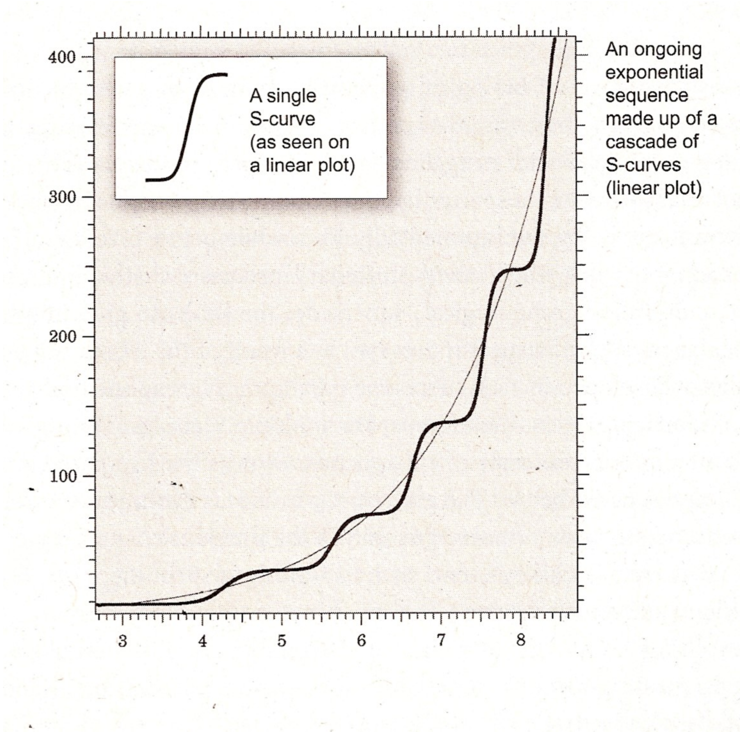
"The explosive nature of exponential growth means it may only take a quarter of a millennium to go from sending messages on horseback to saturating the matter and energy in our solar system with sublimely intelligent processes. The ongoing expansion of our future superintelligence will then require moving out into the rest of the universe, where we may engineer new universes. A new book by James Gardner tells that story.

Consider that the price-performance of computation has grown at a superexponential rate for over a century. The doubling time(of computes per dollar) was three years in 1900 and two years in the middle of the 20th century; and priceperformance is now doubling each year. This progression has been remarkably smooth and predictable through five paradigms of computing substrate: electromechanical calculators, relay-based computers, vacuum tubes, transistors, and now several decades of Moore's Law (which is based on shrinking the size of key features on a flat integrated circuit). The sixth paradigm—three-dimensional molecular computing—is already beginning to work and is waiting in the wings. We see similar smooth exponential progressions in every other aspect of information technology, a phenomenon I call the law of accelerating returns."

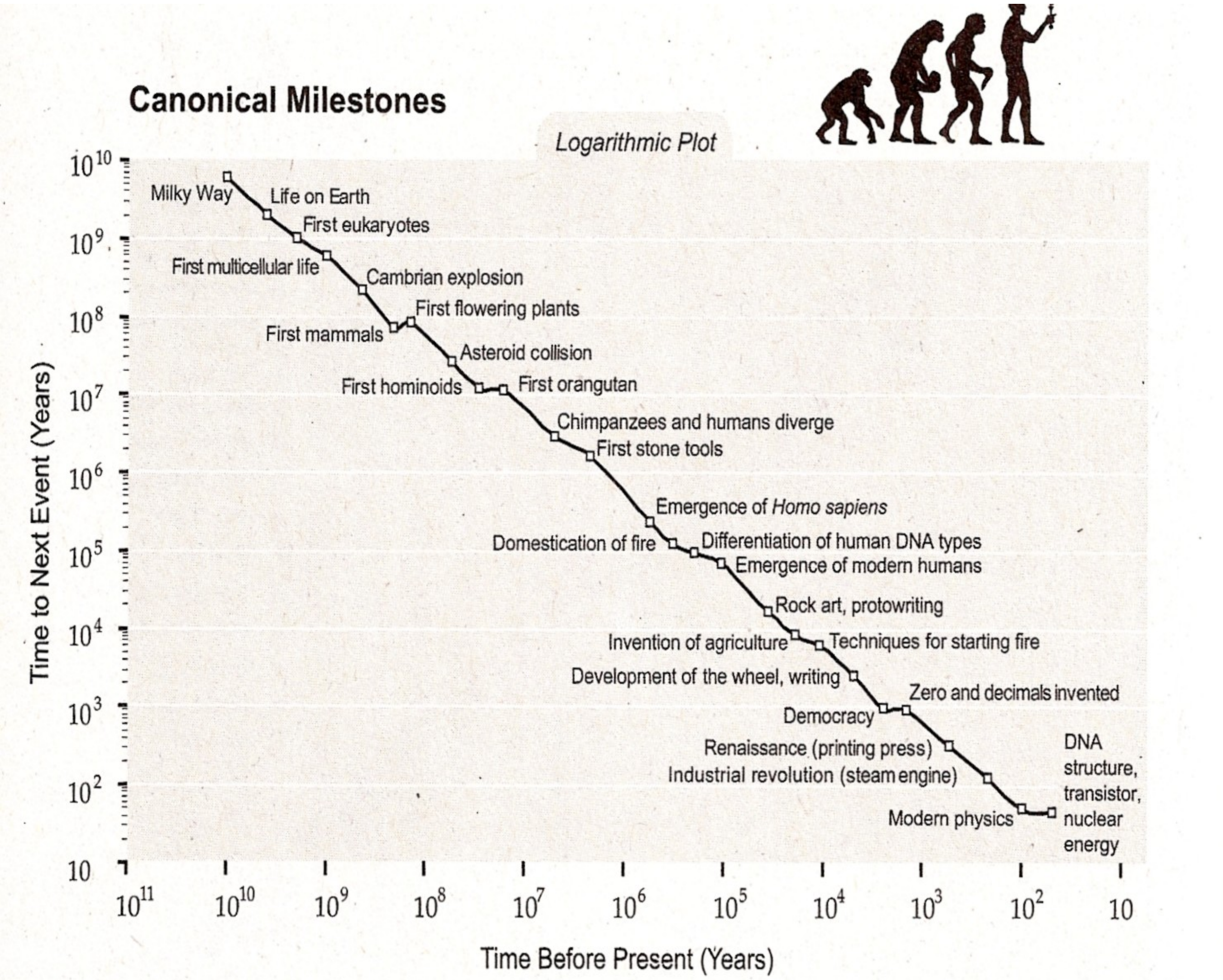
According to Ray Kurzweil the evolution of technology follows the law of accelerating returns.
Each paradigm develops in three stages:

1. Slow growth (the early phase of exponential growth)
2. Rapid growth (the late, explosive phase of exponential growth) as seen in the S-curve figure below.
3. A leveling off as the particular paradigm matures

The S-curve illustration shows how an ongoing exponential trend can be composed of a cascade of S-curves.



This exponential growth is illustrated by the graph on milestones of evolution from the origin of life to todays technology.
The data are compiled from thirteen different sources and clearly show the exponential acceleration of the emergence of innovation during the process of evolution. This yields a straight line in log-log coordinates.

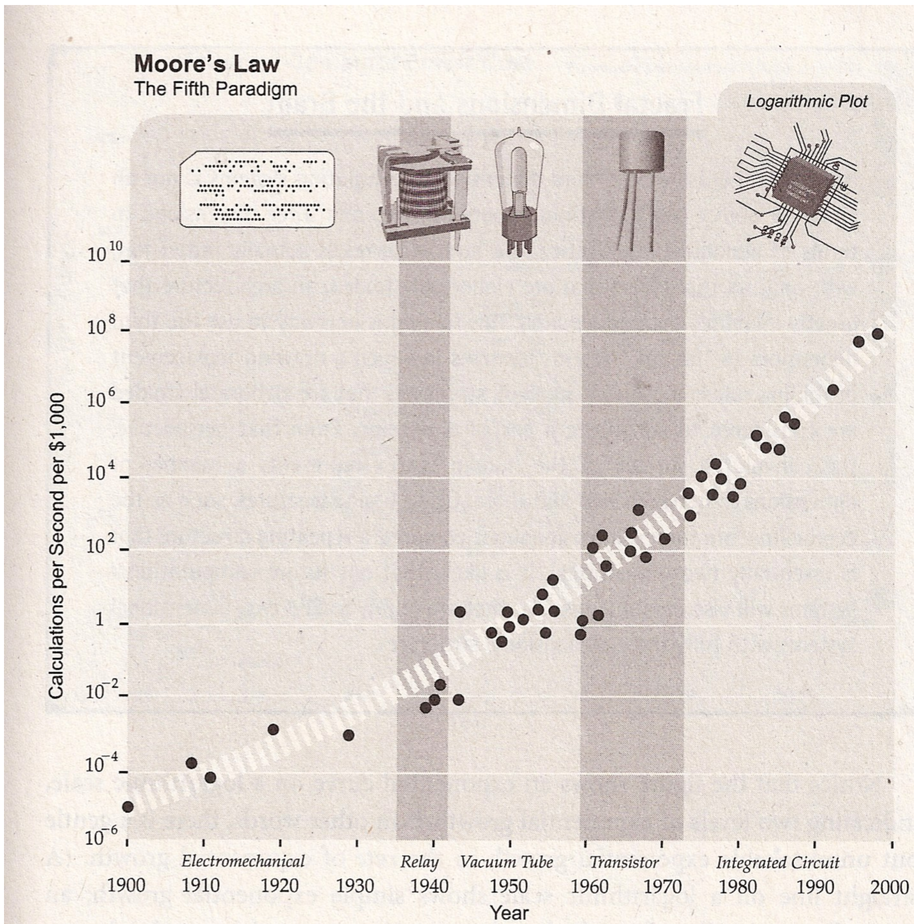


Canonical milestones based on clusters of events from thirteen lists.

Note that several milestones of the Kurzweil graph coincide with levels of our evolutionary Hierarchy in chapter one. This is not surprising, since our order criterion for the hierarchy was the first observation of the emergence of a new level. According to our hierarchy the next evolutionary level should emerge within the Internet and we named it the Web Agent level. Independent software modules communicate within the web directly with each other without human Intervention. This is already the case in many applications of Amazon and Google.

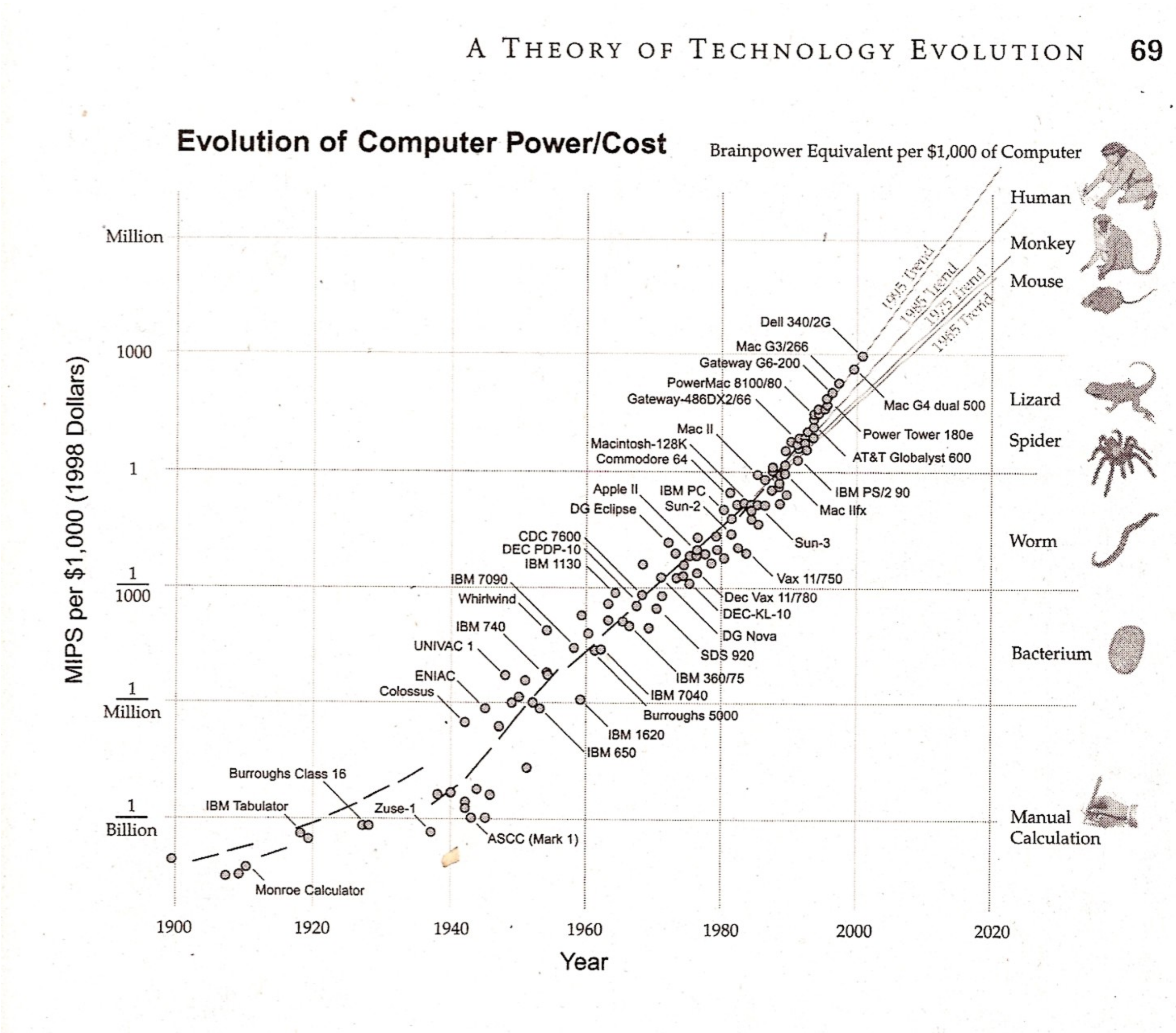
According to the operator hierarchy these modules aggregate to form finally self-replicating units and for science fiction writers the sexual reproduction of Web Agent modules is near.

In the field of computing technology this regularity of exponential increase in calculations per seconds is shown in the following graph and commonly known as Moor's Law.
Note the trend is followed despite the drastic change in technology from simple Hollerith punched cards over electromagnetic relays, vacuum tubes, transistors up to integrated circuits.
The graph also shows a decrease in doubling time (higher than simple exponential or superexponential growth).



The five paradigms of exponential growth of computing: **Each time one paradigm has run out of steam, another has picked up the pace.**

The following graph shows this acceleration of exponential growth in terms of computing cost (MIPS per $1.000) over time. The more we advance in time the steeper the slope of the extrapolated line and eventually this process should lead to a singularity of near infinite growth in time.
The brain power of a single human should be reached in the near future at least by 2020 and the time where total artificial computing power will exceed the total brain power of mankind is not far.



Concerning the future emerging technologies which make the exponential growth possible we add another citation of Ray Kurzweil.

Fractal Dimensions and the Brain

Note that the use of the third dimension in computing systems is not an either-or choice but a continuum between two and three dimensions. In terms of biological intelligence, the human cortex is actually rather flat, with only six thin layers that are elaborately folded, an architecture that greatly increases the surface area. This folding is one way to use the third dimension. In "fractal" systems (systems in which a drawing replacement or folding rule is iteratively applied), structures that are elaborately folded are considered to constitute a partial dimension. From that perspective, the convoluted surface of the human cortex represents a number of dimensions in between two and three. Other brain structures, such as the cerebellum, are three-dimensional but comprise a repeating structure that is essentially two-dimensional. It is likely that our future computational systems will also combine systems that are highly folded two-dimensional systems with fully three-dimensional structures.

A fractal hierarchy of computer structure, a fractal hierarchy of the human brain, a fractal hierarchy of Artificial Neural Networks all seem to point into the same direction in a convergent way.

From a theoretical point of view recent progress in multi / infinite dimensional coding theory, complex valued and multidimensional Neural Networks [Murthy, 2008] point the direction of future research.

Conclusions

Self-similar structure and processes of self-organized systems

1 . All observable evolutionary systems can be described as hierarchical Energy&Information transformation webs .
2. The graph of the web is directed. On each level, energy transformation processors feed upgraded energy into higher levels and feed downgraded energy back to lower levels . This holds for natural evolutionary systems (astrophysical, geochemical, ecosystems ...) but also for artificial or man-made evolutionary systems (city systems,economic systems ...).
3. Energy transformation processors in all observable evolutionary systems show birth and death processes . We presented a simple model for an energy transformation processor, which is general enough to be applicable to any type of energy transformation. A Birth and Death processor is characterized by a limited transformation capacity in which dissipation energy is irreversibly internally accumulated until a threshold capacity is reached .
4. Birth and Death processors can formally be described from two points of view - from an energy-transformation point of view, a Birth and Death processor continuously transforms and accumulates energy up to a threshold (breakdown at death or replication at birth).
- from an information-transformation point of view however, as observed by the metasystem of the processor, a Birth and Death processor is equivalent to a binary threshold automaton or formal neuron with the two possible states : e.g. in the energy processing hierarchy the two states are 0-state (operation or silence) or 1-state (firing at death) .
5. Birth and Death processors organized in an hierarchical transformation system (trophic web) are as a consequence formally equivalent to a neural network of the feed forward type. Lower level processors feed into higher level processors. Hence all the features of neural networks, like memory, adaptation/learning and optimization can be looked at in an analog way in trophic webs of energy/information transformation processors.
6. The error-distribution of a single neuron trained with back-propagation according to gradient descent follows a Pareto-Zipf distribution .
7. Pareto-Zipf distributions are stable under addition, hence we should observe error-distributions of this type also for massively parallel processor networks (like trophic webs) trained with back-propagation.
8. Long tailed distributions or power laws of the Pareto-Zipf-Mandelbrot type, called generalized life-symptoms, are empirically observed for symptoms of virtually all known types of Energy / Information transformation systems .

Self-organized systems have memory (network topology and weights of links)

If we consider that self-organized systems are complex modular scalefree networks of interaction units of the Small World type , they can be mapped on multilayer Perceptrons. Artificial Neural Networks of this type, called multilayer feed forward networks with back-propagatio of errors reveal memory.

The memory lies in the specific **topology** of the network (individual neurons) and in the weights of each **interaction link**. Thus the complex network topology viewed within the ANN paradigm can explain the enigma of how information is stored in the system over many processing cycles.

It's the global field generated by all processors that "drives" the process of evolution based on energy optimization (maximum entropy production) specific to the level of evolution.
GUT, gravitation, strong nuclear, weak nuclear, electo-magnetic, chemical, geothermal, wind, water, fire, genetic code, words, written codes, computer codes ...

Self-organized systems are learning (Hebb's rule, engrams)

Learning (adaptation) of the network takes place in the form of back-propagation . The downgraded energy of each processor is re-cycled influencing the processor parameters in the next processing cycle. Training of the web functions according to gradient descent through positive (or negative) feed-back from higher level processors to lower level processors in both senses of the word (energetic feed-back and cybernetic feedback) .

1. ***ritualization*** **(repetitive use of pathways, Hebb's rule), hardwires the networks information flow, like timetables hardwire a railroad, or air transportation network.**

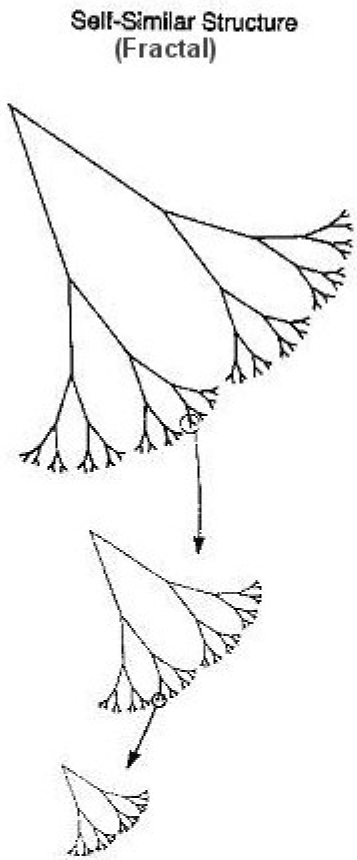
Figure. Hebb's rule "*cells that fire together, wire together*"

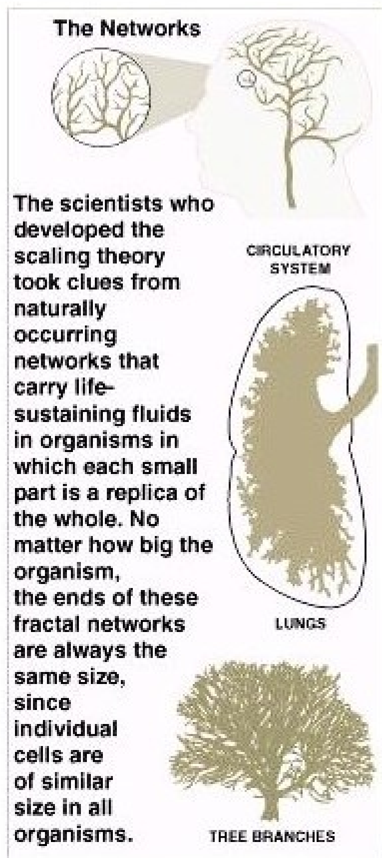
Hebbian theory has been the primary basis for the conventional view that when analyzed from a holistic level, **engrams are neuronal nets or** neural networks.

Gordon Allport

"If the inputs to a system cause the same pattern of activity to occur repeatedly, the set of active elements constituting that pattern will become increasingly strongly interassociated. That is, each element will tend to turn on every other element and (with negative weights) to turn off the elements that do not form part of the pattern. To put it another way, the pattern as a whole will become 'auto-associated'. We may call a learned (auto-associated) pattern an ***engram***."

Image dans son **contexte original**, sur la page universe-review.ca/R10-35-metabolic.htm.





__Self-organized systems are intelligent (maximum or minimum objective function)__

Most theoretical attempts to explain the evolution of specific network topologies are based on extremum principles, which means that the self-organized network strives to achieve a maximum or minimum of a function which characterizes the global system. The principle of maximum entropy production could be a good candidate to apply in most cases of real networks.

The above approach may be the basis for a general theory explaining self-organization, self-learning and evolution in nature.

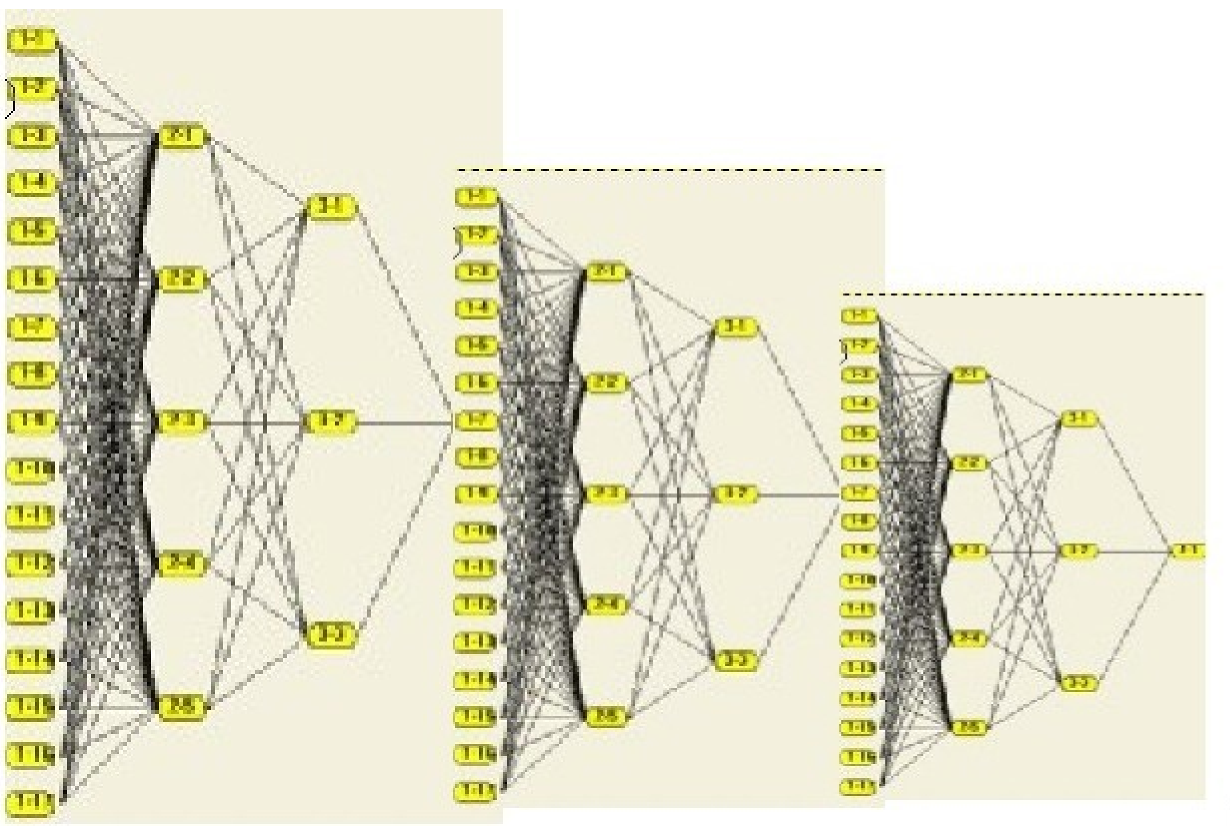


Figure. The universe: a self-similar hierarchy of multilayer Perceptrons, from stars to the World Wide Web.

Bibliography

The top hundred books of the bordalier institute library, which have contributed to the writing of this book, ranked by number of citations :

> 10.000 citations : "The fractal revolution"

The **structure** of **scientific revolutions** - all 22 versions »
TS **Kuhn** - 1962 - philosophyfaculty.ucsd.edu
Page 1. The **Structure** of **Scientific Revolutions** THOMAS S. KUHN (1962) (Excerpt)
VIII. The Response to Crisis Let us then assume that **...**
cited by 19803 -

The **Fractal Geometry** of **Nature** - all 6 versions »
BB **Mandelbrot**, JA Wheeler - American Journal of Physics, 1983 - link.aip.org
The **Fractal Geometry** of **Nature**. [American Journal of Physics 51, 286 (1983)].
Benoit B. Mandelbrot Author, John A. Wheeler. Keywords. **...**
cited by 12419 -

1.000 - 10.000 citations : "The network revolution"

The **Logic** of **Scientific Discovery** - all 6 versions »
KR **Popper** - 2002 - books.google.com
Page 1. The **Logic** of **Scientific Discovery** Karl Popper Routledge Classics R
Routledge Taylor&FrancisCroup , Page 2. Page 3. Page 4. Page 5. **...**
cited by 5973 -

Introduction to the **Theory** of **Neural Computation** - all 11 versions »
J **Hertz**, A **Krogh**, RG **Palmer** - 1991 - books.google.com
Page 1. **INTRODUCTION** TO THE **THEORY** OF **NEURAL COMPUTATION** JobnHer»z AndersKrvgh Richard
G. Palmer Page 2. **INTRODUCTION** TO THE **THEORY** OF **NEURAL COMPUTATION** Page 3 **...**
cited by 4668 -

Tractatus Logico-Philosophicus - all 25 versions »
L **Wittgenstein** - 1981 - books.google.com
... Page 2. Page 3. International Library of Psychology Philosophy and Scientific Method
Tractatus Logico-Philosophicus This One D6DH-RBW-GZPY Page 4. Page 5. **...**
cited by 3934 -

Graph Theory - all 6 versions »
F **Harary** - 1999 - books.google.com
Page 1. **GRAPH THEORY** Frank Harary Page 2. FRANK **...** MARK TWAIN There are several reasons for the acceleration of interest in **graph theory**. It **...**
cited by 3842 -

An **introduction** to **cybernetics**
WR **Ashby** - 1956 - orton.catie.ac.cr
... Signatura : 681.3. ASH. 1966. Autor : Ashby, W. Ross. Título: An **introduction** to **cybernetics**. P.imprenta : John Wiley. New York. Science Editions. (US) 1966. **...**
cited by 2765 -

Cybernetics Or Control and Communication in the Animal and the Machine - all 3 versions »
N **Wiener** - 1961 - books.google.com
Page 1. **Cybernetics**: or Control and Communication in the Animal and the Machine Norbert Wiener Page 2. Page 3. Page 4. Page 5. **CYBERNETICS ...**
cited by 1908 -

Small Worlds: The Dynamics of Networks Between Order and Randomness - all 13 versions »
DJ **Watts** - 1999 - books.google.com
3 n> <—<• O3 GO <—t- C D- K-- CD 05 n O 3 x_ **Small Worlds** r^ *< Duncan J. Watts Page 2. **Small Worlds** Page 3. PRINCETON STUDIES IN COMPLEXITY ED ITORS **...**
cited by 1771 -

Simulating Social Networks: A Review of Three Books
D Watts, AL Barabási, MA Cambridge, P Publishing, … - jasss.soc.surrey.ac.uk
... DOROGOVTSEV SN and JFF Mendes 2000. **Evolution** of reference networks with aging. **...** The large-scale organisation of metabolic networks. **Nature**, 407:651-654. **...**

Gödel, Escher, Bach: An Eternal Golden Braid - all 6 versions »
DR **Hofstadter** - 1999 - books.google.com
... mi **GODEL**, **ESCHER**, **BACH**: ^77 Eternal Golden Braid **...** Hofstadter, Douglas R. 1945- **Godel**, **Escher**, **Bach**. Bibliography: p. 746 Includes index. 1. Metamathematics. **...**
cited by 1522 -

Autopoiesis and **Cognition**: The Realization of the Living - all 3 versions »
HR **Maturana**, FJ **Varela** - 1980 - books.google.com
... AUTOPOIESIS AND **COGNITION** This One C63J-AH9-6WT5 **...** Library of Congress Cataloging in Publication Data Maturana, [lumberto R. 1928- **Autopoiesis** and **cognition**. **...**
cited by 1451 -

Scaling: why is animal size so important?
K **Schmidt-Nielsen** - 1984 - orton.catie.ac.cr
... Signatura: 596 SCHs. Autor: **Schmidt-Nielsen**, K. Título: **Scaling** : why is animal size so important?. P. imprenta: Cambridge. Cambridge University Press. 1984. **...**
Cited by 1421

Linked: The New Science of Networks
AL Barabási, RE Crandall - American Journal of Physics, 2003 - link.aip.org
... With his book **Linked**, Albert-László **Barabási** has "…a simple aim; to get you to think networks" (p. 7). Beginning with a walk through the history of **...**
Cited by 1412 -

Synergetics. an introduction
H **Haken** - Springer Series in **Synergetics**, 1983 - adsabs.harvard.edu
Title: **Synergetics**. an introduction. Authors: Haken, Hermann. Publication: Springer Series in **Synergetics**, Berlin: Springer, 1983, 3rd ed. **...**
cited by 1372 -

General System Theory: Foundations, Development, Applications
L **von Bertalanffy** - 1969 - George Braziller
cited by 1330 -

How **nature works**: the science of self-organized criticality - all 6 versions »
P **Bak** - 1996 - soc.surrey.ac.uk
JASSS logo. How **Nature Works**: The Science of Self-Organised Criticality. Per Bak New York, NY: Copernicus Press 1996 Cloth: ISBN 0-387-94791-4 Order this book. **...**
cited by 1291 - -

Design for a **Brain**
WR **Ashby** - 1954 - Chapman and Hall
cited by 1133 -

Hierarchy: Perspectives for Ecological Complexity - all 3 versions »
TFH **Allen**, TB **Starr** - 1988 - lavoisier.fr
... Hierarchy: Perspectives for Ecological Complexity. Auteur(s) : ALLEN T., STARR B. Date de parution: 10-1982 Langue : ANGLAIS Etat **...**
cited by 1107 - -

500 - 1.000 citations "the EVOLUTION of Networks"

" The **ghost** in the **machine**".
A **Koestler** - Psychiatr Commun, 1968 - ncbi.nlm.nih.gov
Psychiatr Commun. 1968;10(2):45-8. "The **ghost** in the **machine**". Koestler A. MeSH Terms: Psychological Theory*. PMID: 5735264 [PubMed - indexed for MEDLINE]
cited by 960 -

Living systems. - all 4 versions »
JG **Miller** - Curr Mod Biol, 1972 - ncbi.nlm.nih.gov
Curr Mod Biol. 1972 Jan;4(2):55-256. **Living systems**. Miller JG. MeSH Terms:

Biochemistry; Biology*/classification; Cell Physiology; Cybernetics*; DNA/metabolism **...**
cited by 888 -

Principles of **biological autonomy**
FJ **Varela** - 1979 - North Holland, New York
cited by 836 -

Evolution of **Networks**: From Biological Nets to the Internet and WWW - all 2 versions »
SN **Dorogovtsev**, JFF **Mendes** - 2003 - books.google.com
Page 1. **Evolution** of **Networks** From 8IologIcal Nets to the lnternet and WWW SN Dorogovtsev l lFF Mendes **...** Page 2. **Evolution** of **Networks** Th J. s On e P1TY-97N-JK67 **...**
cited by 764 -

Six Degrees: The Science of a Connected Age - all 3 versions »
DJ **Watts** - 2003 - books.google.com
... Page 3. PRAISE FOR DUNCAN WATTS'S **SIX DEGREES** "Watts smoothly combines a historical survey of the field with real- world examples, often drawn from his personal **...**
cited by 755 -

Theory of **hierarchical**, multilevel, systems
MD Mesarovic, **D Macko**, **Y Takahara** - **1970** - slac.stanford.edu
... Title: **Theory** of **hierarchical**, multilevel, systems. Author(s): **MD Mesarovic**, **D**. **Macko**, **Y Takahara**. Date: **1970**. Publisher: Academic Press. Pages: 294. **...**
Cited by 656 -

The **Handbook** of **Brain Theory** and **Neural Networks** - ►**mit.edu** [HTML]
MA Arbib - 2003 - books.google.com
... tor.* second edition Page 2. Page 3. Page 4. Page 5. **The Handbook** of **Brain Theory** and **Neural Networks** Th±s One 8F2T-3Y5-5QNK Page 6. Page 7. **...**
Cited by 647 -

Self-organization in **nonequilibrium systems**: from dissipative structures to order through …
G **Nicolis**, I **Prigogine** - 1977 - slac.stanford.edu
... Title: **Self** - **organization** in **non** - **equilibrium systems** : from dissipative structures to order through fluctuations. Author(s): G. Nicolis and I. Prigogine. **...**
cited by 605 -

The **self-organizing universe**: scientific and human implications of the emerging paradigm of …
E **Jantsch** - 1980 - Pergamon
cited by 629 -

On **Human Communication**: A Review, a Survey, and a Criticism.
C **Cherry** - 1966 - eric.ed.gov
ERIC logo. ED053625 - On **Human Communication**: A Review, a Survey, and a Criticism. ERIC Home. **...** Title: On **Human Communication**: A Review, a Survey, and a Criticism. **...**
cited by 618 -

The **Hypercycle** - all 5 versions »
M **Eigen**, P **Schuster** - Naturwissenschaften, 1978 - Springer
Page 1. **The Hypercycle** A Principle of Natural Self-Organization Part
B" **The** Abstract **Hypercycle** Manfred Eigen Max-Planck **...**
cited by 609 -

100 - 500 citations : "Evolving hierarchical systems"

What is **Life**
E **Schrödinger** - The Physical Aspect of the Living Cell, 1944
cited by 440 -

The **Long Tail** - all 15 versions »
C **Anderson** - Wired Magazine, 2004 - natpe.org
... ANATOMY OF THE **LONG TAIL ...** The appetite for Rhapsody's more obscure tunes
(charted below in yellow) makes up the so-called **Long Tail**. **...**
cited by 402 -

Sync: The Emerging Science of Spontaneous Order - all 3 versions »
S **Strogatz** - 2003 - books.google.com
... eurious. Bu1 if **SYNC** is, well, ou1 of syne wi1h global news, if's eeriaialy ia
1une wifh 1he seieniifie world." **...** raaeher. **SYNC** is a grea1 read." **...**
cited by 395 -

Evolving Hierarchical Systems: Their Structure and Representation
SN **Salthe** - 1985 - books.google.com
... Evolving Hierarchical Systems Their Structure and Representation Stanley N. Salthe
A bold effort to find a functional and transactional framework for synthetic **...**
cited by 358 -

Development and **Evolution**: **Complexity** and **Change** in **Biology**
SN Salthe - **1993** - books.google.com
... Development and **Evolution Complexity** and **Change** in **Biology** Stanley N. **Salthe** A Bradford **...**
Second printing, 1996 © **1993** Massachusetts Institute of Technology **...**
cited by 223

Nexus: Small Worlds and the Groundbreaking Science of Networks - all 5 versions »
M **Buchanan** - 2003 - books.google.com
... Nexus reveals the new science of connection **... Nexus** : small worlds and the
groundbreaking science of networks /Mark Buchanan.— 1st ed. p. cm. **...**
cited by 223 -

Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and …
AL Barabási - 2003 - Plume Books
Cité 218 fois -

Libraries of the **Future**. - all 7 versions »
JCR **Licklider** - 1965 - eric.ed.gov
ERIC logo. ED036106 - **Libraries** of the **Future**. ERIC Home. **...** ERIC #: ED036106.
Title: **Libraries** of the **Future**. Authors: Licklider, JCR. **...**
cited by 205 - -

Hierarchy theory: the challenge of complex systems
HH **Pattee** - 1973 - George Braziller, New York
cited by 192 -

Handbook of **graphs** and **networks** - all 7 versions »
S **Bornholdt**, HG **Schuster** - 2003 - doi.wiley.com
... Handbook of **Graphs** and **Networks** From the Genome to the Internet **Handbook** of **Graphs** and **Networks**: From the Genome to the Internet. **...**
cited by 187 -

The **Structure** and **Dynamics** of **Networks**
MEJ **Newman**, AL Barabási, DJ **Watts** - 2006 - books.google.com
The **Structure** and **Dynamics** of x3 **NETWORKS** \\ K"«» v • '>.< V" i:. "' A• DUNCAN J. WAITS .;-'. **...** The **Structure** and **Dynamics** of **Networks ...**
cited by 182 -

Scaling in **Biology**
JH **Brown**, GB **West** - 2000 - books.google.com
... Page 2. **Scaling** in **Biology** Thi s On e TR1D-CT8-1355 Page 3. **...** Library of Congress Cataloging-m-Publicaiion Data **Scaling** in **biology**/editors. **...**
cited by 177 -

Complexity and the **Function** of **Mind** in **Nature**
P **Godfrey-Smith** - 1996 - books.google.com
... these discussions a vague plan to investigate the "**function** of **mind** in **nature**," and the relations between intelligence and environmental **complexity**, took on a **...**
cited by 158 -

Thinking in **Complexity**: The Complex Dynamics of Matter, Mind, and Mankind - all 4 versions »
K **Mainzer** - 1997 - books.google.com
... Page 2. Klaus Mainzer **Thinking** in **Complexity** Th1 6 One **... Thinking** in **complexity** : the computational dynamics of matter, mind, and mankind / Klaus Mainzer. **...**
cited by 144 -

The **Singularity** Is **Near**: When Humans Transcend Biology
R **Kurzweil** - 2005 - books.google.com
... Page 9. **RAY KURZWEIL** The **Singularity** Is **Near** WHEN HUMANS TRANSCEND BIOLOGY VIKING **...**
The **singularity** is **near** : when humans transcend biology / **Ray Kurzweil**, p. cm. **...**
Cité 137 fois - - Les 2 versions

An **Introduction** to the **Modeling** of **Neural Networks** - all 4 versions »
P **Peretto** - 1992 - books.google.com
... GENERAL EDITOR: C. Godreche AN **INTRODUCTION** TO THE **MODELING** OF **NEURAL NETWORKS** Page 4. **...** PERETTO: An **Introduction** to the **Modeling** of **Neural Networks** Page 5. **...**
cited by 132 -

Life Itself
F **Crick** - 1981 - Simon & Schuster
cited by 128 -

L'organisation biologique et la **théorie** de **l'information**
H **Atlan** - 1972 - Hermann
cited by 115 -

Cosmic Evolution: The Rise of Complexity in Nature - all 3 versions »
E **Chaisson** - 2001 - books.google.com
... Page 2. **COSMIC EVOLUTION** ^*^&&£• ffi&&\ Copyright © 2001 The President and Fellows of Harvard College **...** Page 4. **COSMIC EVOLUTION** THERISEOFCOMPLEXITY IN NATURE **...**
cited by 106 -

15 - 100 citations : "The languages of Nature"

Note that the books with the fewest number of citations are the most closely related to the topic of this book

---

The **synergism hypothesis**: a theory of progressive evolution
PA **Corning** - 1983 - McGraw-Hill Companies
cited by 98 -

The **Laws** of the **Web**: Patterns in the Ecology of Information - all 4 versions »
BA **Huberman** - 2001 - books.google.com
... OF INFORMATION BERNARDO A. HUBERMAN Page 2. . : Page 3. Page 4. Page 5. The **Laws** of the **Web** This One DZK7-OSE-R3DX Page 6. Page 7. The **...**
cited by 75 -

Up from **Eden**: A Transpersonal View of Human Evolution
K **Wilber** - 1996 - books.google.com
... **UP** FROM **EDEN** A TRANSPERSONAL VIEW OF HUMAN EVOLUTION Th±e One BRUE-1RW-7WC3 **...** Heading Toward Omega, and Life at Death "**Up** from **Eden** is a colossal achievement. **...**
cited by 71 -

Le hasard et la nécessite, essai sur la philosophie naturelle de la biologie moderne
J **Monod** - 1970 - orton.catie.ac.cr
... Autor: Monod, Jacques. Título: Le **hasard** et la **nécessite**, essai sur la philosophie

naturelle de la biologie moderne. P.imprenta: Du Seuil. Paris. FR. 1970. **...**
cited by 63 - -

The **Food Web** of a **Tropical Rain Forest** - all 3 versions »
DP **Reagan**, RB **Waide** - 1996 - books.google.com
Page 1. The **Food Web** of a **Tropical Rain Forest** Edited by Douglas P. Reagan and Robert B. Waide Page 2. **...** Page 5. The **Food Web** of a **Tropical Rain Forest** Page 6. **...**
cited by 60 -

Hierarchically Organized Systems in **Theory** and **Practice**
PA **Weiss**, HK Buechner - 1971 - Hafner Pub. Co.
cited by 59

Fifth Language: Learning a Living in the Computer Age
R **Logan** - 1996 - Stoddart
cited by 57 -

From **Complexity** to **Creativity**: Explorations in Evolutionary, Autopoietic, and Cognitive Dynamics - all 9 versions »
B **Goertzel** - 1997 - books.google.com
FROM **COMPLEXITY** TO **CREATIVITY** Explorations in Evolutionary, Autopoietic, and Cognitive Dynamics BEN GOERTZEL IFSR International Series on Systems Science and **...**
cited by 56 -

La **dynamique des villes**
D **Pumain** - 1982 - Economica, Paris
cited by 54 -

Laws of the Game: How the Principles of Nature Govern Chance
M **Eigen**, R **Winkler** - 1993 - books.google.com
... p. cm.—(Princeton science library) Translation of: **Das Spiel**. Originally published: 1st American ed. New York: Knopf: Distributed by Random House, 1981 . **...**
cited by 49 -

La **relativité dans tous ses états**: au-delà de l'espace-temps
L **Nottale** - 1998 - Hachette, [Paris]
cited by 47 -

The predator-prey model: do we live in a Volterra world
M **Peschel**, W **Mende** - Akademie-Verlag, Berlin, 1986
cited by 47 -

The **Biological Universe**
SJ **Dick** - The **Biological Universe**, by **Steven J**. **Dick**, 1996 - adsabs.harvard.edu
... Publication: The **Biological Universe**, by **Steven J**. **Dick**, pp. 596. ISBN 0521343267. Cambridge, UK: Cambridge University Press, June 1996. **...**
Cité 39 fois - - Les 2 versions

Patience dans l'azur: l'evolution cosmique - all 2 versions »
H **Reeves** - Paris: Seuil, c1981., 1981 - adsabs.harvard.edu
Title: **Patience dans l'azur** : l'evolution cosmique. Authors: Reeves, Hubert.
Publication: Paris : Seuil, c1981. Publication Date: 00/1981. Origin: LOC. **...**
cited by 35 -

The **Sixth Language**-Learning a Living in the Internet Age - all 2 versions »
RK **LOGAN** - Canadian Journal of Communication-, 2001 - cjc-online.ca
Paul Heyer, . (2001, January 1). The **Sixth Language**: Learning a Living in the Internet
Age. **...** The **Sixth Language**: Learning a Living in the Internet Age. **...**
cited by 30 - -

Evolution's Arrow: The Direction of **Evolution** and the Future of Humanity - ►surrey.ac.uk **[HTML]**
J **Stewart** - 2000 - books.google.com
... National Library of Australia Cataloguing-in-publication data: **Stewart**, **John**, 1952
Nov. 27- **Evolution's Arrow** : the direction of **evolution** and the future of **...**
Cité 26 fois - - Les 4 versions

The **nature** of **life**
JH **Postlethwait**, JL Hopson - 1992 - McGraw-Hill Companies
cited by 21 -

< 15 citations : "The Extended Mind" The Emergence of Language, the Human Mind, and Culture

---

Les arbres de l'évolution
L **Nottale**, J **Chaline**, P **Grou** - Paris: Hachette, 2000
cited by 13 -

ALGOL-manual der ALCOR-Gruppe
R Baumann - 1969 - R. Oldenbourg Munchen
cited by 12 -

Biocosm - all 2 versions »
JN **Gardner** - … Life is the Architect of the Universe (http://www. **biocosm**. … - **biocosm**.org
... Heaven **BIOCOSM** The New Scientific Theory of Evolution: **...** Publishing The Selfish
Biocosm Hypothesis The Selfish **Biocosm** hypothesis asserts that: **...**
cited by 11 -

Lonely Planets
D **Grinspoon** - setileague.org
... Page 2 Book Review: **Lonely Planets ...** To this category belongs David

Grinspoon's **Lonely Planets**: The Natural Philosophy of Alien Life. **...**
cited by 11 -

Cybernetics, theory and applications
R **Trappl** - 1983 - Hemisphere Pub. Corp Washington
cited by 9 -

L'héritage de la **liberté**: de l'animalité à l'humanitude
A **Jacquard** - 1986 - Seuil
cited by 7 -

L'univers et la **lumière**
L **Nottale** - … Bibliothèque Scientifique. Paris: Flammarion. Livro para …, 1994
Cité 6 fois

Les **arbres** de **l'évolution**: univers, vie, sociétés
L **Nottale**, J **Chaline**, P **Grou** - 2000 - Hachette littératures
Cité 5 fois

Le Secret du Monde
J **Kepler** - Les Belles Lettres, Paris, 1984
cited by 5 -

Complex adaptive systems: an introduction to computational models of social life
JH **Miller**, J Howard, SE **Page** - 2007 - Princeton, New Jersey: Princeton University Press
cited by 4 -

La **biologie** de **l'esprit**
R **Chauvin** - 1990 - Editions du Rocher, Monaco
cited by 4 -

L'homme fou
A **Bourguignon** - 1994 - Presses universitaires de France, Paris
cited by 2 -

Multi/Infinite Dimensional **Neural Networks**, Multi/Infinite Dimensional Logic Theory - all 2 versions »
GR **Murthy** - International Journal of **Neural** Systems, 2005 – worldscinet.com

Multidimensional Neural Networks Unified Theory
GR **Murthy** - New Age International Publishers, 1980

http://www.newagepublishers.com/servlet/nagetbiblio?bno=001548
... Keywords: **Multi-dimensional neural networks**; logic functions; tensors; logic gates; logic circuits. **...** 2. Mathematical Model of **Multi-Dimensional Neural Networks ...**
cited by 3 -

The **Intelligent Universe**: AI, ET, and the Emerging Mind of the Cosmos
J **Gardner** - 2007 - books.google.com
... THE **INTELLIGENT UNIVERSE** EDITED AND TYPESET BY ASTRID DERIDDER Cover design by Dutton and Sherman Printed in the USA by Book-mart Press **...**
cited by 2 -

The **Extended Mind**: The Emergence of Language, the Human **Mind**, and Culture
RK **Logan** - 2007 - books.google.com
Page 1. THE **EXTENDED MIND** The Emergence of Language, the Human **Mind**, and Culture ROBERT K. LOGAN Page 2. THE **EXTENDED MIND**: THE EMERGENCE **...**

Driving Forces in **Physical**, **Biological** and **Socio-economic Phenomena**: A Network Science Investigation … - all 2 versions »
BM **Roehner** - 2007 - books.google.com
... Page 2. **DRIVING FORCES** IN **PHYSICAL**, **BIOLOGICAL** AND **SOCIO-ECONOMIC PHENOMENA** ANetwork Science Investigation of Social Bonds and Interactions **...**
cited by 1 -

The Sleepwalkers (Hutchinson, London 1968) 624 pp. A. Koestler
A **Koestler** - **Die Nachtwandler**-Das Bild des Universums im Wandel der Zeit
cited by 1 -

List of publications by Peter Winiwarter relevant to this book :

Roehner, B., & Winiwarter, P. (1985). AGGREGATION OF INDEPENDENT PARETIAN RANDOM VARIABLES [***probabilit***y]. Advances in Applied Probabilities, 17, 465-469.http://www.bordalierinstitute.com/probability.pdf
statistical stability of Pareto distributions

Winiwarter, P. (1983a). THE GENESIS MODEL PART I: COMPLEXITY, A MEASURE FOR THE EVOLUTION OF SELF-ORGANISED SYSTEMS OF MATTER [***GenesisModel 1***]. Speculations in Science and Technology, 6(1), 11-20.http://www.bordalierinstitute.com/genesisModel1.pdf
a quantitative measure of complexity, complexity can only increase or remain constant.

Winiwarter, P. (1983b). THE GENESIS MODEL - PART II: FREQUENCY DISTRIBUTIONS OF ELEMENTS IN SELFORGANIZED SYSTEMS [***GenesisModel 2***]. Speculations in Science and Technology, 6(2). 103-112.http://www.bordalierinstitute.com/genesisModel2.pdf

Winiwarter, P. (1984). HOMEOSTATIC TENDENCIES OF FIRM-SIZE DISTRIBUTIONS AND THE EVOLUTION ECONOMIC SYSTEMS [***firmsize***]. In AFCET (Ed.), Cybernetics and Systems (pp. 999-1005). 6th International Congress of Systems WOGSadfasdf, 2. http://www.bordalierinstitute.com/firmsize.pdf

Winiwarter, P. (1986). AUTOGNOSIS: THE THEORY OF HIERARCHICAL SELF-IMAGE BUILDING SYSTEMS [***Autognosis***]. In Society for General Systems Research (Ed.), Mental Images, Values and Reality (Vol. A, pp. A21-A31). 30TH ANNUAL MEETING. http://www.bordalierinstitute.com/autognosis.pdf

Winiwarter, P. (1986). CONCEPTS OF SELF-ORGANIZATION () SELFORGANIZATION OF CONCEPTS [***Selforg***]. In Society for General Systems Research (Ed.), Mental Images, Values and Reality (Vol. D, pp. D62-D76). 30TH ANNUAL MEETING. http://www.bordalierinstitute.com/self-organization.pdf

Winiwarter, P. (1986). NATURAL NUCLEAR COMPUTERS: A MODEL FOR HUMAN INTELLIGENCE [***Nuclear computers***]. In Society for General Systems Research (Ed.), Mental Images, Values and Reality (Vol. D, pp. D1-D9). 30TH ANNUAL MEETING. http://www.bordalierinstitute.com/starcomputations.pdf

Winiwarter, P. (1999). A periodic System of System concepts [***Conceptual framework***]. Semiotica, 125(1/3), 47-62. http://www.bordalierinstitute.com/semiotica.pdf

Winiwarter, P., & Cempel, C. (1992). Life Symptoms: The Behaviour of Open Systems with limited Dissipation Capacity [***Life symptoms***]. Systems Research, 9(4), 9-34.http://www.bordalierinstitute.com/lifesymptoms.pdf
energy transformation and dissipation capacity

Winiwarter, P., & Cempel, C. From Trophic Webs and Hierarchical Energy Transformation Systems to Multilayer Neural Networks [***GenesisModel 3***]. (1996). . In PPSN IV. Berlin:http://www.bordalierinstitute.com/neuralnetworks.pdf
equivalence of a network of binary threashold "birth and death processors" and a "multilayer perceptron" artificial neural network

Winiwarter,P., & Cempel, C. Evolutionary Hiearchies and Energy Processing in Nature [***evolutioniary hierachies***]. (1996)In Austrian Society for Cybernetic Studies. Trappl, R. (Ed.), Vol.2, p 26-31.
http://www.bordalierinstitute.com/evolutionaryhierarchies.pdf
coevolution of macroscopic and microscopic hierarchies

About the author, from Wikipedia, the free encyclopedia

Peter Winiwarter (born October 4, 1945) is an Austrian-born French resident scientist. Since more than 25 years he is Director of the Bordalier Institute in France.

He introduced in 1983 the *first law of genesis* stating that the complexity of self-organized systems can only increase and the *second law of genesis* stating that unit-size distributions of self-organized systems follow a Pareto-Zipf-Mandelbrot PZM law. He introduced in 1992 the equivalence concept of trophic webs and artificial neural networks. At hand of numerous examples he shows the widespread empirical evidence of PZM regularities in natural, technological and social systems, from stars to the World Wide Web.

Contents

Birth and education

Winiwarter was born to an Austrian father and a German mother in the community of St.Valentin in Lower Austria close to the city of Linz. His father, Friedrich, was a surgeon and director of the local hospital, while his mother, Ida, assisted his father as aid. He attended high school at the 'humanistic gymnasium' in Linz, specializing in humanities and classical languages (8years of Latin and 6 years of ancient Greek). At the age of 17 he graduated and won an American Field Service Scholarship to the United States where he spent one year in Southern California. Between 1964 and 1970, he studied physics , mathematics and philosophy at the University of Vienna; during that time, he began doing research in general systems theory.

In 1970 he earned a Ph.D. In Nuclear Physics under the direction of P.Hille in the field of Spin Distribution of Nuclear Level Density.

In 1974 he received an M.B.A. from INSEAD, Fontainebleau, France, the European Institute for Business Administration, one of the highest rated business schools of the world, which has now a second campus in Singapour.

Academic career

In 1970 Winiwarter received a fellowship by the Austrian Ministry of Research and Education at the Nuclear Data compilation Center of the OECD NEA nuclear data bank at Saclay France, where he worked on the first worldwide Nuclear Data Base in the field of IT.

In 1971 he received a fellowship of the United Nations to work under the direction of Nobel prize winner Y.M. Franck as a Senior scientist at the, Joint Institute for Nuclear Research. at Dubna USSR, where he lived as the only western scientist during one year.

Subsequently he worked as a consultant for the International Atomic Energy Agency. in Vienna, Austria.

In 1974 he received an M.B.A. from one of the world's top businesses schools INSEAD in Fontainebleau, France leaving his Nuclear physics career behind and concentrating on the management of complex computer systems in the field of business.

In subsequent years he worked for an American Think tank Arthur D. Little (ADL), Cambridge, Massachusetts with assignments in North-Africa and the most important countries in Europe.

Since 1983 Winiwarter joins his primary interest in Systems Theory and Hierarchy Theory at the Bordalier Institute with part-time assignments as consultant to earn the necessary funds to do transdisciplinary research, not financed by the academic institutions.

It is this financial independence which allowed him to do unconventional research in transdisciplinary systems theory.

As of 2001, Winiwarter is a French citizen, and a permanent resident of France.

Research and achievements

Winiwarter has been a major contributor to the development of complex systems theory, together with several other scientists from physics, mathematics, and computer science (Ilya Prigogine, Manfred Eigen). His biggest role has been the introduction of the *periodic system of systems concept* . Among the topics in systems theory Winiwarter has studied the Behaviour of Open Systems with Limited Energy Dissipation Capacity and the mapping of trophic webs to Artificial Neural Networks.

Awards

American Field Service (AFS) International Scholarship

Austrian Ministry for Research and Education fellowship

United Nations Research fellowship

Selected publications see list above

Acknowledgments

First of all I would like to thank Professor Czeslaw Cempel from the Poznan Institute of Technology in Poland for the more than 20 years of intense collaboration. We have published many papers together and the discussions with him were always full of helpful advise.

I thank also the contributors to more technical sections of this book. Stan Salthey, from Bingham University, the 'pope' of hierarchy theory contributed with a special section on the difference between compositional and subsumptional hierarchy, not easy to grasp at first sight and often confused in the literature.

Thanks to Gerard Jagers, whom I got to know at the first conference on evolution and development of the universe. [www.evodevouniverse.com]. Gerard contributed with a section on operator hierarchy.

The early works of my research have been largely influenced by my colleagues of the International Society for Systems Science, Pierre Auger – now French academician – and Bertrand Roehner from the Laboratory of Theoretical Physics of the university Paris VII, where I held a position of visiting professor for one year.

Finally I thank Nobel prize winner Manfred Eigen (the father of the hypercycle) for heaving read several of my manuscripts and who has encouraged me to pursue the difficult task of transdisciplinary researcher.

Back cover *Neural Network Nature*

*"What I like about the book is the employment of the idea of network evolution as a pattern that permeates all orders of magnitude, implying the universe is networked. I also find the ideas of network and ecology to be the dominant images and metaphors of our time which seem to converge and synthesize in this **unified theory of networks**."*

Richard Thomas, Beal Institute

Recursionism

In the philosophy of Subhash Kak **recursionism** refers to the idea that replicated forms and self-similar forms are common in the physical world, and that this has some mystical significance. Kak describes recursionism as follows:

> *Patterns repeat across space, time, scale and fields. Recursion is an expression of the fundamental laws of nature, and it is to be seen both at the physical and the abstract levels as also across relational entities. Recursionism provides a way of knowing since it helps us to find meaning by a shift in perspective and by abstraction.*

The idea of recursionism also occurs in Hindu Vedanta philosophy, where it is seen most prominently in the Upanishads. There are recursionist strands in the works of Fichte, Schopenhauer,Nietzsche.