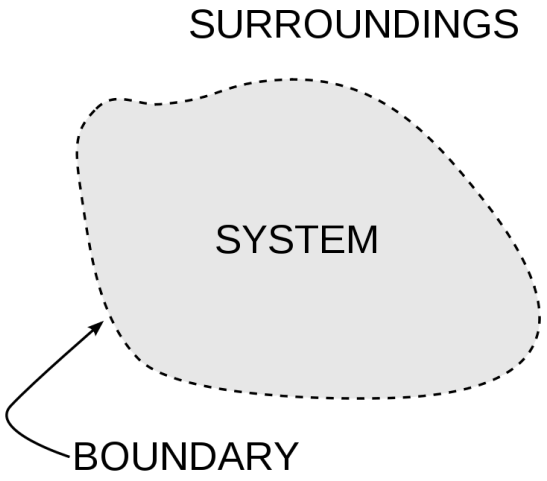
Complex Systems Theory and Biodynamics

Complexity, Emergent Systems and Complex Biological Systems

Complex Systems Theory

System

System (from Latin *systēma*, in turn from Greek *σύστημα* systēma) is a set of interacting or interdependent entities, real or abstract, forming an integrated whole.



A schematic representation of a closed system and its boundary

The concept of an **'integrated whole**' can also be stated in terms of a system embodying a set of relationships which are differentiated from relationships of the set to other elements, and from relationships between an element of the set and elements not a part of the relational regime.

The scientific research field which is engaged in the study of the general properties of systems include systems theory, systems science, systemics and systems engineering. They investigate the abstract properties of the matter and organization, searching concepts and principles which are independent of the specific domain, substance, type, or temporal scales of existence.

Most systems share the same common characteristics. These common characteristics include the following

- Systems are abstractions of reality.
- Systems have structure which is defined by its parts and their composition.
- Systems have behavior, which involves inputs, processing and outputs of material, information or energy.
- Systems have interconnectivity, the various parts of a system have functional as well as structural relationships between each other.

The term *system* may also refer to a set of rules that governs behavior or structure.

History

The term *System* has a long history which can be traced back to the Greek language.

In the 19th century the first to develop the concept of a "system" in the natural sciences was the French physicist Nicolas Léonard Sadi Carnot who studied thermodynamics. In 1824 he studied what he called the *working substance* (system), i.e. typically a body of water vapor, in steam engines, in regards to the system's ability to do work when heat is applied to it. The working substance could be put in contact with either a boiler, a cold reservoir (a stream of cold water), or a piston (to which the working body could do work by pushing on it). In 1850, the German physicist Rudolf Clausius generalized this picture to

include the concept of the surroundings and began to use the term "working body" when referring to the system.

One of the pioneers of the general systems theory was the biologist Ludwig von Bertalanffy. In 1945 he introduced *models, principles, and laws that apply to generalized systems or their subclasses, irrespective of their particular kind, the nature of their component elements, and the relation or 'forces' between them.*[1]

Significant development to the concept of a *system* was done by Norbert Wiener and Ross Ashby who pioneered the use of mathematics to study systems [2] [3] .

In the 1980s the term complex adaptive system was coined at the interdisciplinary Santa Fe Institute by John H. Holland, Murray Gell-Mann and others.

System concepts

Environment and boundaries

Systems theory views the world as a complex system of interconnected parts. We scope a system by defining its boundary; this means choosing which entities are inside the system and which are outside - part of the environment. We then make simplified representations (models) of the system in order to understand it and to predict or impact its future behavior. These models may define the structure and/or the behavior of the system.

Natural and man-made systems

There are natural and man-made (designed) systems. Natural systems may not have an apparent objective but their outputs can be interpreted as purposes. Man-made systems are made with purposes that are achieved by the delivery of outputs. Their parts must be related; they must be “designed to work as a coherent entity” - else they would be two or more distinct systems

Open system

An open system usually interacts with some entities in their environment. A closed system is isolated from its environment.

Process and transformation process

A system can also be viewed as a bounded transformation process, that is, a process or collection of processes that transforms inputs into outputs. Inputs are consumed; outputs are produced. The concept of input and output here is very broad. E.g., an output of a passenger ship is the movement of people from departure to destination.

Subsystem

A *subsystem* is a set of elements, which is a system itself, and a part of a larger system.

Types of systems

Evidently, there are many types of systems that can be analyzed both quantitatively and qualitatively. For example, with an analysis of urban systems dynamics, [A.W. Steiss] [4] defines five intersecting systems, including the physical subsystem and behavioral system. For sociological models influenced by systems theory, where Kenneth D. Bailey [5] defines systems in terms of conceptual, concrete and abstract systems; either isolated, closed, or open, Walter F. Buckley [6] defines social systems in sociology in terms of mechanical,

organic, and process models. Bela H. Banathy [7] cautions that with any inquiry into a system that understanding the type of system is crucial and defines Natural and Designed systems.

In offering these more global definitions, the author maintains that it is important not to confuse one for the other. The theorist explains that natural systems include sub-atomic systems, living systems, the solar system, the galactic system and the Universe. Designed systems are our creations, our physical structures, hybrid systems which include natural and designed systems, and our conceptual knowledge. The human element of organization and activities are emphasized with their relevant abstract systems and representations. A key consideration in making distinctions among various types of systems is to determine how much freedom the system has to select purpose, goals, methods, tools, etc. and how widely is the freedom to select distributed (or concentrated) in the system.

George J. Klir [8] maintains that no "classification is complete and perfect for all purposes," and defines systems in terms of abstract, real, and conceptual physical systems, bounded and unbounded systems, discrete to continuous, pulse to hybrid systems, et cetera. The interaction between systems and their environments are categorized in terms of absolutely closed systems, relatively closed, and open systems. The case of an absolutely closed system is a rare, special case. Important distinctions have also been made between hard and soft systems.[9] Hard systems are associated with areas such as systems engineering, operations research and quantitative systems analysis. Soft systems are commonly associated with concepts developed by Peter Checkland through Soft Systems Methodology (SSM) involving methods such as action research and emphasizing participatory designs. Where hard systems might be identified as more "scientific," the distinction between them is actually often hard to define.

Cultural system

A cultural system may be defined as the interaction of different elements of culture. While a cultural system is quite different from a social system, sometimes both systems together are referred to as the sociocultural system. A major concern in the social sciences is the problem of order. One way that social order has been theorized is according to the degree of integration of cultural and social factors.

Economic system

An economic system is a mechanism (social institution) which deals with the production, distribution and consumption of goods and services in a particular society. The economic system is composed of people, institutions and their relationships to resources, such as the convention of property. It addresses the problems of economics, like the allocation and scarcity of resources.

Biological system

Application of the system concept

Systems modeling is generally a basic principle in engineering and in social sciences. The system is the representation of the entities under concern. Hence inclusion to or exclusion from system context is dependent of the intention of the modeler.

No model of a system will include all features of the real system of concern, and no model of a system must include all entities belonging to a real system of concern.

Systems in information and computer science

In computer science and information science, **system** could also be a method or an algorithm. Again, an example will illustrate: There are systems of counting, as with Roman numerals, and various systems for filing papers, or catalogues, and various library systems, of which the Dewey Decimal System is an example. This still fits with the definition of components which are connected together (in this case in order to facilitate the flow of information).

System can also be used referring to a framework, be it software or hardware, designed to allow software programs to run, see platform.

Systems in engineering and physics

In engineering and physics, a physical system is the portion of the universe that is being studied (of which a thermodynamic system is one major example). Engineering also has the concept of a system that refers to all of the parts and interactions between parts of a complex project. Systems engineering refers to the branch of engineering that studies how this type of system should be planned, designed, implemented, built, and maintained.

Systems in social and cognitive sciences and management research

Social and cognitive sciences recognize systems in human person models and in human societies. They include human brain functions and human mental processes as well as normative ethics systems and social/cultural behavioral patterns.

In management science, operations research and organizational development (OD), human organizations are viewed as **systems** (conceptual systems) of interacting components such as subsystems or system aggregates, which are carriers of numerous complex processes and organizational structures. Organizational development theorist Peter Senge developed the notion of organizations as systems in his book *The Fifth Discipline*.

Systems thinking is a style of thinking/reasoning and problem solving. It starts from the recognition of system properties in a given problem. It can be a leadership competency. Some people can *think globally while acting locally*. Such people consider the potential consequences of their decisions on other parts of larger systems. This is also a basis of systemic coaching in psychology.

Organizational theorists such as Margaret Wheatley have also described the workings of organizational systems in new metaphoric contexts, such as quantum physics, chaos theory, and the self-organization of systems.

Systems applied to strategic thinking

In 1988, military strategist, John A. Warden III introduced his Five Ring System model in his book, *The Air Campaign* contending that any complex system could be broken down into five concentric rings. Each ring--Leadership, Processes, Infrastructure, Population and Action Units--could be used to isolate key elements of any system that needed change. The model was used effectively by Air Force planners in the First Gulf War. [10] , [11] , [12] . In the late 1990's, Warden applied this five ring model to business strategy[13] .

See also

Examples of systems

- Complex system
- Computer system
- List of systems (WikiProject)
- Meta-system
- Solar System
- Systems in human anatomy

Theories about systems

- Chaos theory
- Cybernetics
- Formal system
- Systems ecology
- Systems intelligence
- Systems theory
- World-systems approach

Related topics

- Complexity theory and organizations
- Glossary of systems theory
- Network
- System of systems (engineering)
- Systems art
- Wikipedia Books: System

References

[1] 1945, *Zu einer allgemeinen Systemlehre,* Blätter für deutsche Philosophie, 3/4. (Extract in: Biologia Generalis, 19 (1949), 139-164.

[2] 1948, *Cybernetics: Or the Control and Communication in the Animal and the Machine*. Paris, France: Librairie Hermann & Cie, and Cambridge, MA: MIT Press.Cambridge, MA: MIT Press.

[3] 1956. *An Introduction to Cybernetics (http://pespmc1.vub.ac.be/ASHBBOOK.html)*, Chapman & Hall.

[4] Steiss 1967, p.8-18.

[5] Bailey, 1994.

[6] Buckley, 1967.

[7] Banathy, 1997.

[8] Klir 1969, pp. 69-72

[9] Checkland 1997; Flood 1999.

[10] Warden, John A. III (1988). *The Air Campaign: Planning for Combat*. Washington, D.C.: National Defense University Press. ISBN 9781583481004.

[11] Warden, John A. III (September 1995). "Chapter 4: Air theory for the 21st century" (http://www.airpower.maxwell.af.mil/airchronicles/battle/chp4.html) (in *Air and Space Power Journal*). *Battlefield of the Future: 21st Century Warfare Issues*. United States Air Force. . Retrieved on December 26, 2008.

[12] Warden, John A. III (1995). " Enemy as a System (http://www.airpower.maxwell.af.mil/airchronicles/apj/apj95/spr95_files/warden.htm)". *Airpower Journal* **Spring** (9): 40-55. . Retrieved on 2009-03-25.

[13] Russell, Leland A.; Warden, John A. (2001). *Winning in FastTime: Harness the Competitive Advantage of Prometheus in Business and in Life*. Newport Beach, CA: GEO Group Press. ISBN 0971269718.

Further reading

- Alexander Backlund (2000). "The definition of system". In: *Kybernetes* Vol. 29 nr. 4, pp. 444-451.
- Kenneth D. Bailey (1994). *Sociology and the New Systems Theory: Toward a Theoretical Synthesis*. New York: State of New York Press.
- Bela H. Banathy (1997). "A Taste of Systemics" (http://www.newciv.org/ISSS_Primer/asem04bb.html), ISSS The Primer Project.
- Walter F. Buckley (1967). *Sociology and Modern Systems Theory*, New Jersey: Englewood Cliffs.
- Peter Checkland (1997). *Systems Thinking, Systems Practice*. Chichester: John Wiley & Sons, Ltd.
- Robert L. Flood (1999). *Rethinking the Fifth Discipline: Learning within the unknowable*. London: Routledge.
- George J. Klir (1969). Approach to General Systems Theory, 1969.

External links

- *Definitions of Systems and Models* (http://www.physicalgeography.net/fundamentals/4b.html) by Michael Pidwirny, 1999-2007.
- *Definitionen von "System" (1572-2002)* (http://www.muellerscience.com/SPEZIALITAETEN/System/System_Definitionen.htm) by Roland Müller, 2001-2007 (most in German).

Dynamics

Dynamics (from Greek *δυναμικός - dynamikos* "powerful", from *δύναμις - dynamis* "power") may refer to:

In Physics

- Dynamics (physics), in physics, dynamics refers to time evolution of physical processes
- Analytical dynamics refers to the motion of bodies as induced by external forces
- Relativistic dynamics may refer to a combination of relativistic and quantum concepts
- Molecular dynamics, the study of motion on the molecular level
- Thermodynamics, a branch of physics that studies the relationships between heat and mechanical energy
- Fluid dynamics, the study of fluid flow; includes:
  - Aerodynamics, the study of gases in motion
  - Hydrodynamics, the study of liquids in motion
- In quantum physics, *dynamics* may refer to how forces are quantized, as in quantum electrodynamics or quantum chromodynamics

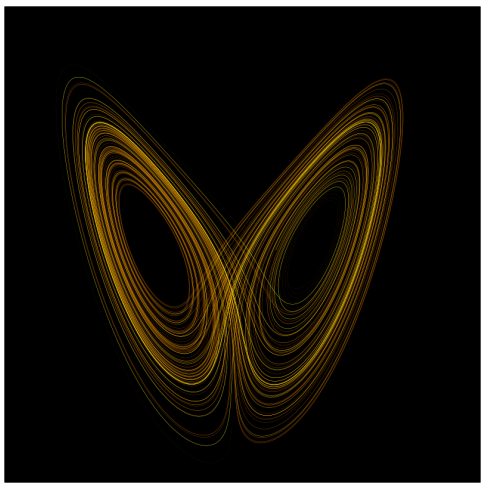
Other

- System dynamics, the study of the behaviour of complex systems
- A Dynamical system in mathematics or complexity
- Dynamics (music), In music, dynamics refers to the softness or loudness of a sound or note. The term is also applied to the written or printed musical notation used to indicate dynamics

- Group dynamics, the study of social group processes
- Psychodynamics, the study of the interrelationship of various parts of the mind, personality, or psyche as they relate to mental, emotional, or motivational forces especially at the subconscious level
- Neurodynamics, an area of research in the brain sciences which places a strong focus upon the spatio-temporal (dynamic) character of neural activity in describing brain function
- Power dynamics, the dynamics of power, used in sociology
- Dynamic programming in computer science and control theory
- Dynamic program analysis, in computer science is a set of methods for analyzing code that is performed with executing programs built from that software on a real or virtual processor
- Microsoft Dynamics is a line of business software owned and developed by Microsoft
- UMass Dynamics is a well-known a cappella group based out of UMass Amherst

Dynamical system

The **dynamical system** concept is a mathematical formalization for any fixed "rule" which describes the time dependence of a point's position in its ambient space. Examples include the mathematical models that describe the swinging of a clock pendulum, the flow of water in a pipe, and the number of fish each spring in a lake.



The Lorenz attractor is an example of a non-linear dynamical system. Studying this system helped give rise to Chaos theory.

At any given time a dynamical system has a *state* given by a set of real numbers (a vector) which can be represented by a point in an appropriate *state space* (a geometrical manifold). Small changes in the state of the system correspond to small changes in the numbers. The *evolution rule* of the dynamical system is a fixed rule that describes what future states follow from the current state. The rule is deterministic: for a given time interval only one future state follows from the current state.

Overview

The concept of a dynamical system has its origins in Newtonian mechanics. There, as in other natural sciences and engineering disciplines, the evolution rule of dynamical systems is given implicitly by a relation that gives the state of the system only a short time into the future. (The relation is either a differential equation, difference equation or other time scale.) To determine the state for all future times requires iterating the relation many times—each advancing time a small step. The iteration procedure is referred to as *solving the system* or *integrating the system*. Once the system can be solved, given an initial point it is possible to determine all its future points, a collection known as a *trajectory* or *orbit*.

Before the advent of fast computing machines, solving a dynamical system required sophisticated mathematical techniques and could only be accomplished for a small class of dynamical systems. Numerical methods executed on computers have simplified the task of determining the orbits of a dynamical system.

For simple dynamical systems, knowing the trajectory is often sufficient, but most dynamical systems are too complicated to be understood in terms of individual trajectories. The difficulties arise because:

- The systems studied may only be known approximately—the parameters of the system may not be known precisely or terms may be missing from the equations. The approximations used bring into question the validity or relevance of numerical solutions. To address these questions several notions of stability have been introduced in the study of dynamical systems, such as Lyapunov stability or structural stability. The stability of the dynamical system implies that there is a class of models or initial conditions for which the trajectories would be equivalent. The operation for comparing orbits to establish their equivalence changes with the different notions of stability.
- The type of trajectory may be more important than one particular trajectory. Some trajectories may be periodic, whereas others may wander through many different states of the system. Applications often require enumerating these classes or maintaining the system within one class. Classifying all possible trajectories has led to the qualitative study of dynamical systems, that is, properties that do not change under coordinate changes. Linear dynamical systems and systems that have two numbers describing a state are examples of dynamical systems where the possible classes of orbits are understood.
- The behavior of trajectories as a function of a parameter may be what is needed for an application. As a parameter is varied, the dynamical systems may have bifurcation points where the qualitative behavior of the dynamical system changes. For example, it may go from having only periodic motions to apparently erratic behavior, as in the transition to turbulence of a fluid.
- The trajectories of the system may appear erratic, as if random. In these cases it may be necessary to compute averages using one very long trajectory or many different trajectories. The averages are well defined for ergodic systems and a more detailed understanding has been worked out for hyperbolic systems. Understanding the probabilistic aspects of dynamical systems has helped establish the foundations of statistical mechanics and of chaos.

It was in the work of Poincaré that these dynamical systems themes developed.

Basic definitions

A dynamical system is a manifold M called the phase (or state) space and a smooth evolution function Φ^t that for any element of $t \in T$, the time, maps a point of the phase space back into the phase space. The notion of smoothness changes with applications and the type of manifold. There are several choices for the set T. When T is taken to be the reals, the dynamical system is called a *flow*; and if T is restricted to the non-negative reals, then the dynamical system is a *semi-flow*. When T is taken to be the integers, it is a *cascade* or a *map*; and the restriction to the non-negative integers is a *semi-cascade*.

Examples

The evolution function Φ^t is often the solution of a *differential equation of motion*

$$\dot{x} = v(x)\,.$$

The equation gives the time derivative, represented by the dot, of a trajectory *x(t)* on the phase space starting at some point x_0. The *vector field v(x)* is a smooth function that at every point of the phase space *M* provides the velocity vector of the dynamical system at that point. (These vectors are not vectors in the phase space *M*, but in the tangent space TM_x of the point *x*.) Given a smooth Φ^t, an autonomous vector field can be derived from it.

There is no need for higher order derivatives in the equation, nor for time dependence in *v(x)* because these can be eliminated by considering systems of higher dimensions. Other types of differential equations can be used to define the evolution rule:

$$G(x, \dot{x}) = 0$$

is an example of an equation that arises from the modeling of mechanical systems with complicated constraints.

The differential equations determining the evolution function Φ^t are often ordinary differential equations: in this case the phase space *M* is a finite dimensional manifold. Many of the concepts in dynamical systems can be extended to infinite-dimensional manifolds—those that are locally Banach spaces—in which case the differential equations are partial differential equations. In the late 20th century the dynamical system perspective to partial differential equations started gaining popularity.

Further examples

- Logistic map
- Double pendulum
- Arnold's cat map
- Horseshoe map
- Baker's map is an example of a chaotic piecewise linear map
- Billiards and outer billiards
- Hénon map
- Lorenz system
- Circle map
- Rössler map
- List of chaotic maps
- Swinging Atwood's machine
- Quadratic map simulation system
- Bouncing ball simulation system

Linear dynamical systems

Linear dynamical systems can be solved in terms of simple functions and the behavior of all orbits classified. In a linear system the phase space is the N-dimensional Euclidean space, so any point in phase space can be represented by a vector with N numbers. The analysis of linear systems is possible because they satisfy a superposition principle: if *u(t)* and *w(t)* satisfy the differential equation for the vector field (but not necessarily the initial condition), then so will *u(t) + w(t)*.

Flows

For a flow, the vector field $\Phi(x)$ is a linear function of the position in the phase space, that is,

$$\phi(x) = Ax + b\,,$$

with A a matrix, b a vector of numbers and x the position vector. The solution to this system can be found by using the superposition principle (linearity). The case $b \neq 0$ with $A = 0$ is just a straight line in the direction of b:

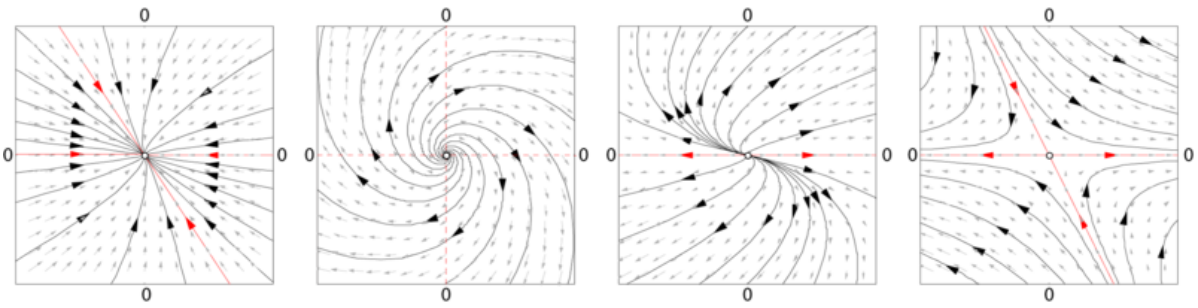
$$\Phi^t(x_1) = x_1 + bt\,.$$

When b is zero and $A \neq 0$ the origin is an equilibrium (or singular) point of the flow, that is, if $x_0 = 0$, then the orbit remains there. For other initial conditions, the equation of motion is given by the exponential of a matrix: for an initial point x_0,

$$\Phi^t(x_0) = e^{tA}x_0\,.$$

When $b = 0$, the eigenvalues of A determine the structure of the phase space. From the eigenvalues and the eigenvectors of A it is possible to determine if an initial point will converge or diverge to the equilibrium point at the origin.

The distance between two different initial conditions in the case $A \neq 0$ will change exponentially in most cases, either converging exponentially fast towards a point, or diverging exponentially fast. Linear systems display sensitive dependence on initial conditions in the case of divergence. For nonlinear systems this is one of the (necessary but not sufficient) conditions for chaotic behavior.



Linear vector fields and a few trajectories.

Maps

A discrete-time, affine dynamical system has the form

$$x_{n+1} = Ax_n + b\,,$$

with A a matrix and b a vector. As in the continuous case, the change of coordinates $x \to x + (1 - A)^{-1}b$ removes the term b from the equation. In the new coordinate system, the origin is a fixed point of the map and the solutions are of the linear system $A^n x_0$. The solutions for the map are no longer curves, but points that hop in the phase space. The orbits are organized in curves, or fibers, which are collections of points that map into themselves under the action of the map.

As in the continuous case, the eigenvalues and eigenvectors of A determine the structure of phase space. For example, if u_1 is an eigenvector of A, with a real eigenvalue smaller than one, then the straight lines given by the points along $\alpha\, u_1$, with $\alpha \in \mathbf{R}$, is an invariant curve of the map. Points in this straight line run into the fixed point.

There are also many other discrete dynamical systems.

Local dynamics

The qualitative properties of dynamical systems do not change under a smooth change of coordinates (this is sometimes taken as a definition of qualitative): a *singular point* of the vector field (a point where $v(x) = 0$) will remain a singular point under smooth transformations; a *periodic orbit* is a loop in phase space and smooth deformations of the phase space cannot alter it being a loop. It is in the neighborhood of singular points and periodic orbits that the structure of a phase space of a dynamical system can be well understood. In the qualitative study of dynamical systems, the approach is to show that there is a change of coordinates (usually unspecified, but computable) that makes the dynamical system as simple as possible.

Rectification

A flow in most small patches of the phase space can be made very simple. If y is a point where the vector field $v(y) \neq 0$, then there is a change of coordinates for a region around y where the vector field becomes a series of parallel vectors of the same magnitude. This is known as the rectification theorem.

The rectification theorem says that away from singular points the dynamics of a point in a small patch is a straight line. The patch can sometimes be enlarged by stitching several patches together, and when this works out in the whole phase space M the dynamical system is *integrable*. In most cases the patch cannot be extended to the entire phase space. There may be singular points in the vector field (where $v(x) = 0$); or the patches may become smaller and smaller as some point is approached. The more subtle reason is a global constraint, where the trajectory starts out in a patch, and after visiting a series of other patches comes back to the original one. If the next time the orbit loops around phase space in a different way, then it is impossible to rectify the vector field in the whole series of patches.

Near periodic orbits

In general, in the neighborhood of a periodic orbit the rectification theorem cannot be used. Poincaré developed an approach that transforms the analysis near a periodic orbit to the analysis of a map. Pick a point x_0 in the orbit γ and consider the points in phase space in that neighborhood that are perpendicular to $v(x_0)$. These points are a Poincaré section $S(\gamma, x_0)$, of the orbit. The flow now defines a map, the Poincaré map $F : S \to S$, for points starting in S and returning to S. Not all these points will take the same amount of time to come back, but the times will be close to the time it takes x_0.

The intersection of the periodic orbit with the Poincaré section is a fixed point of the Poincaré map F. By a translation, the point can be assumed to be at $x = 0$. The Taylor series of the map is $F(x) = J \cdot x + O(x^2)$, so a change of coordinates h can only be expected to simplify F to its linear part

$$h^{-1} \circ F \circ h(x) = J \cdot x \,.$$

This is known as the conjugation equation. Finding conditions for this equation to hold has been one of the major tasks of research in dynamical systems. Poincaré first approached it assuming all functions to be analytic and in the process discovered the non-resonant

condition. If $\lambda_1,\ldots,\lambda_\nu$ are the eigenvalues of J they will be resonant if one eigenvalue is an integer linear combination of two or more of the others. As terms of the form $\lambda_i - \sum$ (multiples of other eigenvalues) occurs in the denominator of the terms for the function h, the non-resonant condition is also known as the small divisor problem.

Conjugation results

The results on the existence of a solution to the conjugation equation depend on the eigenvalues of J and the degree of smoothness required from h. As J does not need to have any special symmetries, its eigenvalues will typically be complex numbers. When the eigenvalues of J are not in the unit circle, the dynamics near the fixed point x_0 of F is called *hyperbolic* and when the eigenvalues are on the unit circle and complex, the dynamics is called *elliptic*.

In the hyperbolic case the Hartman-Grobman theorem gives the conditions for the existence of a continuous function that maps the neighborhood of the fixed point of the map to the linear map $J \cdot x$. The hyperbolic case is also *structurally stable*. Small changes in the vector field will only produce small changes in the Poincaré map and these small changes will reflect in small changes in the position of the eigenvalues of J in the complex plane, implying that the map is still hyperbolic.

The Kolmogorov-Arnold-Moser (KAM) theorem gives the behavior near an elliptic point.

Bifurcation theory

When the evolution map Φ^t (or the vector field it is derived from) depends on a parameter μ, the structure of the phase space will also depend on this parameter. Small changes may produce no qualitative changes in the phase space until a special value μ_0 is reached. At this point the phase space changes qualitatively and the dynamical system is said to have gone through a bifurcation.

Bifurcation theory considers a structure in phase space (typically a fixed point, a periodic orbit, or an invariant torus) and studies its behavior as a function of the parameter μ. At the bifurcation point the structure may change its stability, split into new structures, or merge with other structures. By using Taylor series approximations of the maps and an understanding of the differences that may be eliminated by a change of coordinates, it is possible to catalog the bifurcations of dynamical systems.

The bifurcations of a hyperbolic fixed point x_0 of a system family F_μ can be characterized by the eigenvalues of the first derivative of the system $DF_\mu(x_0)$ computed at the bifurcation point. For a map, the bifurcation will occur when there are eigenvalues of DF_μ on the unit circle. For a flow, it will occur when there are eigenvalues on the imaginary axis. For more information, see the main article on Bifurcation theory.

Some bifurcations can lead to very complicated structures in phase space. For example, the Ruelle-Takens scenario describes how a periodic orbit bifurcates into a torus and the torus into a strange attractor. In another example, Feigenbaum period-doubling describes how a stable periodic orbit goes through a series of period-doubling bifurcations.

Ergodic systems

In many dynamical systems it is possible to choose the coordinates of the system so that the volume (really a ν-dimensional volume) in phase space is invariant. This happens for mechanical systems derived from Newton's laws as long as the coordinates are the position and the momentum and the volume is measured in units of (position) × (momentum). The flow takes points of a subset A into the points $\Phi^{t}(A)$ and invariance of the phase space means that

$$\mathrm{vol}(A) = \mathrm{vol}(\Phi^t(A)) .$$

In the Hamiltonian formalism, given a coordinate it is possible to derive the appropriate (generalized) momentum such that the associated volume is preserved by the flow. The volume is said to be computed by the Liouville measure.

In a Hamiltonian system not all possible configurations of position and momentum can be reached from an initial condition. Because of energy conservation, only the states with the same energy as the initial condition are accessible. The states with the same energy form an energy shell Ω, a sub-manifold of the phase space. The volume of the energy shell, computed using the Liouville measure, is preserved under evolution.

For systems where the volume is preserved by the flow, Poincaré discovered the recurrence theorem: Assume the phase space has a finite Liouville volume and let F be a phase space volume-preserving map and A a subset of the phase space. Then almost every point of A returns to A infinitely often. The Poincaré recurrence theorem was used by Zermelo to object to Boltzmann's derivation of the increase in entropy in a dynamical system of colliding atoms.

One of the questions raised by Boltzmann's work was the possible equality between time averages and space averages, what he called the ergodic hypothesis. The hypothesis states that the length of time a typical trajectory spends in a region A is vol(A)/vol(Ω).

The ergodic hypothesis turned out not to be the essential property needed for the development of statistical mechanics and a series of other ergodic-like properties were introduced to capture the relevant aspects of physical systems. Koopman approached the study of ergodic systems by the use of functional analysis. An observable a is a function that to each point of the phase space associates a number (say instantaneous pressure, or average height). The value of an observable can be computed at another time by using the evolution function φ^{t}. This introduces an operator U^{t}, the transfer operator,

$$(U^t a)(x) = a(\Phi^{-t}(x)) .$$

By studying the spectral properties of the linear operator U it becomes possible to classify the ergodic properties of Φ^{t}. In using the Koopman approach of considering the action of the flow on an observable function, the finite-dimensional nonlinear problem involving Φ^{t} gets mapped into an infinite-dimensional linear problem involving U.

The Liouville measure restricted to the energy surface Ω is the basis for the averages computed in equilibrium statistical mechanics. An average in time along a trajectory is equivalent to an average in space computed with the Boltzmann factor $\exp(-\beta H)$. This idea has been generalized by Sinai, Bowen, and Ruelle (SRB) to a larger class of dynamical systems that includes dissipative systems. SRB measures replace the Boltzmann factor and they are defined on attractors of chaotic systems.

Chaos theory

Simple nonlinear dynamical systems and even piecewise linear systems can exhibit a completely unpredictable behavior, which might seem to be random. (Remember that we are speaking of completely deterministic systems!). This seemingly unpredictable behavior has been called *chaos*. Hyperbolic systems are precisely defined dynamical systems that exhibit the properties ascribed to chaotic systems. In hyperbolic systems the tangent space

perpendicular to a trajectory can be well separated into two parts: one with the points that converge towards the orbit (the *stable manifold*) and another of the points that diverge from the orbit (the *unstable manifold*).

This branch of mathematics deals with the long-term qualitative behavior of dynamical systems. Here, the focus is not on finding precise solutions to the equations defining the dynamical system (which is often hopeless), but rather to answer questions like "Will the system settle down to a steady state in the long term, and if so, what are the possible attractors?" or "Does the long-term behavior of the system depend on its initial condition?"

Note that the chaotic behavior of complicated systems is not the issue. Meteorology has been known for years to involve complicated—even chaotic—behavior. Chaos theory has been so surprising because chaos can be found within almost trivial systems. The logistic map is only a second-degree polynomial; the horseshoe map is piecewise linear.

Geometrical definition

A dynamical system is the tuple $\langle \mathcal{M}, f, \mathcal{T} \rangle$, with $\mathcal{M}$ a manifold (locally a Banach space or Euclidean space), $\mathcal{T}$ the domain for time (non-negative reals, the integers, ...) and f an evolution rule t→f^t (with $t \in \mathcal{T}$) such that f^t is a diffeomorphism of the manifold to itself. So, f is a mapping of the time-domain $\mathcal{T}$ into the space of diffeomorphisms of the manifold to itself. In other terms, f(t) is a diffeomorphism, for every time t in the domain $\mathcal{T}$.

Measure theoretical definition

See main article measure-preserving dynamical system.

A dynamical system may be defined formally, as a measure-preserving transformation of a sigma-algebra, the quadruplet (X, Σ, μ, τ). Here, *X* is a set, and Σ is a sigma-algebra on *X*, so that the pair (X, Σ) is a measurable space. μ is a finite measure on the sigma-algebra, so that the triplet (X, Σ, μ) is a probability space. A map $\tau : X \to X$ is said to be Σ-measurable if and only if, for every $\sigma \in \Sigma$, one has $\tau^{-1}\sigma \in \Sigma$. A map τ is said to **preserve the measure** if and only if, for every $\sigma \in \Sigma$, one has $\mu(\tau^{-1}\sigma) = \mu(\sigma)$. Combining the above, a map τ is said to be a **measure-preserving transformation of *X*** , if it is a map from *X* to itself, it is Σ-measurable, and is measure-preserving. The quadruple (X, Σ, μ, τ), for such a τ, is then defined to be a **dynamical system**.

The map τ embodies the time evolution of the dynamical system. Thus, for discrete dynamical systems the iterates $\tau^n = \tau \circ \tau \circ \ldots \circ \tau$ for integer *n* are studied. For continuous dynamical systems, the map τ is understood to be finite time evolution map and the construction is more complicated.

Examples of dynamical systems

Wikipedia links

- Arnold's cat map
- Baker's map is an example of a chaotic piecewise linear map
- Circle map
- Double pendulum
- Billiards and Outer Billiards
- Henon map
- Horseshoe map
- Irrational rotation
- List of chaotic maps
- Logistic map
- Lorenz system
- Rossler map

External links

- Bouncing Ball [1]
- Mechanical Strings [2]
- Journal of Advanced Research in Dynamical and Control Systems [3]
- Swinging Atwood's Machine (SAM) [4]
- Interactive applet for the Standard and Henon Maps [5] by A. Luhn

See also

- Behavioral modeling
- Dynamical systems theory
- List of dynamical system topics
- Oscillation
- People in systems and control
- Sarkovskii's theorem
- System dynamics
- Systems theory

References

[1] http://www.drchaos.net/drchaos/bb.html
[2] http://www.drchaos.net/drchaos/string_web_page/index.html
[3] http://www.i-asr.org/dynamic.html
[4] http://www.drchaos.net/drchaos/Sam/sam.html
[5] http://complexity.xozzox.de/nonlinmappings.html

Further reading

Works providing a broad coverage:

- Ralph Abraham and Jerrold E. Marsden (1978). *Foundations of mechanics*. Benjamin-Cummings. ISBN 0-8053-0102-X. (available as a reprint: ISBN 0-201-40840-6)

- *Encyclopaedia of Mathematical Sciences* (ISSN 0938-0396) has a sub-series on dynamical systems (http://en.wikipedia.org/wiki/User:XaosBits/EMP) with reviews of current research.
- Anatole Katok and Boris Hasselblatt (1996). *Introduction to the modern theory of dynamical systems*. Cambridge. ISBN 0-521-57557-5.
- Christian Bonatti, Lorenzo J. Díaz, Marcelo Viana (2005). *Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective*. Springer. ISBN 3-540-22066-6.
- Diederich Hinrichsen and Anthony J. Pritchard (2005). *Mathematical Systems Theory I - Modelling, State Space Analysis, Stability and Robustness*. Springer Verlag. ISBN 978-3-540-44125-0.

Introductory texts with a unique perspective:

- V. I. Arnold (1982). *Mathematical methods of classical mechanics*. Springer-Verlag. ISBN 0-387-96890-3.
- Jacob Palis and Wellington de Melo (1982). *Geometric theory of dynamical systems: an introduction*. Springer-Verlag. ISBN 0-387-90668-1.
- David Ruelle (1989). *Elements of Differentiable Dynamics and Bifurcation Theory*. Academic Press. ISBN 0-12-601710-7.
- Tim Bedford, Michael Keane and Caroline Series, *eds.* (1991). *Ergodic theory, symbolic dynamics and hyperbolic spaces*. Oxford University Press. ISBN 0-19-853390-X.
- Ralph H. Abraham and Christopher D. Shaw (1992). *Dynamics—the geometry of behavior, 2nd edition*. Addison-Wesley. ISBN 0-201-56716-4.

Textbooks

- Steven H. Strogatz (1994). *Nonlinear dynamics and chaos: with applications to physics, biology chemistry and engineering*. Addison Wesley. ISBN 0-201-54344-3.
- Kathleen T. Alligood, Tim D. Sauer and James A. Yorke (2000). *Chaos. An introduction to dynamical systems*. Springer Verlag. ISBN 0-387-94677-2.
- Morris W. Hirsch, Stephen Smale and Robert Devaney (2003). *Differential Equations, dynamical systems, and an introduction to chaos*. Academic Press. ISBN 0-12-349703-5.

Popularizations:

- Florin Diacu and Philip Holmes (1996). *Celestial Encounters*. Princeton. ISBN 0-691-02743-9.
- James Gleick (1988). *Chaos: Making a New Science*. Penguin. ISBN 0-14-009250-1.
- Ivar Ekeland (1990). *Mathematics and the Unexpected (Paperback)*. University Of Chicago Press. ISBN 0-226-19990-8.
- Ian Stewart (1997). *Does God Play Dice? The New Mathematics of Chaos*. Penguin. ISBN 0140256024.

External links

- A collection of dynamic and non-linear system models and demo applets (http://vlab.infotech.monash.edu.au/simulations/non-linear/) (in Monash University's Virtual Lab)
- Arxiv preprint server (http://www.arxiv.org/list/math.DS/recent) has daily submissions of (non-refereed) manuscripts in dynamical systems.
- DSWeb (http://www.dynamicalsystems.org/) provides up-to-date information on dynamical systems and its applications.
- Encyclopedia of dynamical systems (http://www.scholarpedia.org/article/Encyclopedia_of_Dynamical_Systems) A part of Scholarpedia — peer reviewed and written by invited experts.
- Nonlinear Dynamics (http://www.egwald.ca/nonlineardynamics/index.php). Models of bifurcation and chaos by Elmer G. Wiens
- Oliver Knill (http://www.dynamical-systems.org) has a series of examples of dynamical systems with explanations and interactive controls.
- Sci.Nonlinear FAQ 2.0 (Sept 2003) (http://amath.colorado.edu/faculty/jdm/faq-Contents.html) provides definitions, explanations and resources related to nonlinear science

Online books or lecture notes:

- Geometrical theory of dynamical systems (http://arxiv.org/pdf/math.HO/0111177). Nils Berglund's lecture notes for a course at ETH at the advanced undergraduate level.
- Dynamical systems (http://www.ams.org/online_bks/coll9/). George D. Birkhoff's 1927 book already takes a modern approach to dynamical systems.
- Chaos: classical and quantum (http://chaosbook.org). An introduction to dynamical systems from the periodic orbit point of view.
- Modeling Dynamic Systems (http://www.embedded.com/2000/0008/0008feat2.htm). An introduction to the development of mathematical models of dynamic systems.
- Learning Dynamical Systems (http://www.cs.brown.edu/research/ai/dynamics/tutorial/home.html). Tutorial on learning dynamical systems.
- Ordinary Differential Equations and Dynamical Systems (http://www.mat.univie.ac.at/~gerald/ftp/book-ode/). Lecture notes by Gerald Teschl

Research groups:

- Dynamical Systems Group Groningen (http://www.math.rug.nl/~broer/), IWI, University of Groningen.
- Chaos @ UMD (http://www-chaos.umd.edu/). Concentrates on the applications of dynamical systems.
- Dynamical Systems (http://www.math.sunysb.edu/dynamics/), SUNY Stony Brook. Lists of conferences, researchers, and some open problems.
- Center for Dynamics and Geometry (http://www.math.psu.edu/dynsys/), Penn State.
- Control and Dynamical Systems (http://www.cds.caltech.edu/), Caltech.
- Laboratory of Nonlinear Systems (http://lanoswww.epfl.ch/), Ecole Polytechnique Fédérale de Lausanne (EPFL).
- Center for Dynamical Systems (http://www.math.uni-bremen.de/ids.html/), University of Bremen
- Systems Analysis, Modelling and Prediction Group (http://www.eng.ox.ac.uk/samp/), University of Oxford

- Non-Linear Dynamics Group (http://sd.ist.utl.pt/), Instituto Superior Técnico, Technical University of Lisbon
- Dynamical Systems (http://www.impa.br/), IMPA, Instituto Nacional de Matemática Pura e Aplicada.
- Nonlinear Dynamics Workgroup (http://ndw.cs.cas.cz/), Institute of Computer Science, Czech Academy of Sciences.

Simulation software based on Dynamical Systems approach:

- FyDiK (http://fydik.kitnarf.cz/)

Complex system

This article describes complex system as a type of system. For other meanings, see complex systems.

A **complex system** is a system composed of interconnected parts that as a whole exhibit one or more properties (behavior among the possible properties) not obvious from the properties of the individual parts.

A system's complexity may be of one of two forms: disorganized complexity and organized complexity.[1] In essence, disorganized complexity is a matter of a very large number of parts, and organized complexity is a matter of the subject system (quite possibly with only a limited number of parts) exhibiting emergent properties.

Examples of complex systems include ant colonies, human economies and social structures, climate, nervous systems, cells and living things, including human beings, as well as modern energy or telecommunication infrastructures. Indeed, many systems of interest to humans are complex systems.

Complex systems are studied by many areas of natural science, mathematics, and social science. Fields that specialize in the interdisciplinary study of complex systems include systems theory, complexity theory, systems ecology, and cybernetics.

Overview

A complex system is any system featuring a large number of interacting components, whose aggregate activity is non-linear and typically exhibits self-organization under selective pressures.[2] Now the term *complex systems* has multiple meaning:

- A specific kind of systems, that are complex
- A field of science studying these systems, *see further complex systems*
- A paradigm, that complex systems have to be studied with non-linear dynamics, *see further complexity*

Various informal descriptions of complex systems have been put forward, and these may give some insight into their properties. A special edition of *Science* about complex systems [3] highlighted several of these:

- A complex system is a highly structured system, which shows structure with variations (N. Goldenfeld and Kadanoff)
- A complex system is one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve (Whitesides

and Ismagilov)
- A complex system is one that by design or function or both is difficult to understand and verify (Weng, Bhalla and Iyengar)
- A complex system is one in which there are multiple interactions between many different components (D. Rind)
- Complex systems are systems in process that constantly evolve and unfold over time (W. Brian Arthur).

History

Although one can argue that humans have been studying complex systems for thousands of years, the modern scientific study of complex systems is relatively young when compared to areas of science such as physics and chemistry. The history of the scientific study of these systems follows several different strands.

In the area of mathematics, arguably the largest contribution to the study of complex systems was the discovery of chaos in deterministic systems, a feature of certain dynamical systems that is strongly related to nonlinearity.[4] The study of neural networks was also integral in advancing the mathematics needed to study complex systems.

The notion of self-organizing systems is tied up to work in nonequilibrium thermodynamics, including that pioneered by chemist and Nobel laureate Ilya Prigogine in his study of dissipative structures.

Types of complex systems

A commonly accepted taxonomy of complex systems does not exist yet, but most characteristic are the following.

Chaotic systems

For a dynamical system to be classified as chaotic, most scientists will agree that it must have the following properties:

1. it must be sensitive to initial conditions,
2. it must be topologically mixing, and
3. its periodic orbits must be dense

Sensitivity to initial conditions means that each point in such a system is arbitrarily closely approximated by other points with significantly different future trajectories. Thus, an arbitrarily small perturbation of the current trajectory may lead to significantly different future behavior.

Complex adaptive systems

Complex adaptive systems (CAS) are special cases of complex systems. They are complex in that they are diverse and made up of multiple interconnected elements and adaptive in that they have the capacity to change and learn from experience. Examples of complex adaptive systems include the stock market, social insect and ant colonies, the biosphere and the ecosystem, the brain and the immune system, the cell and the developing embryo, manufacturing businesses and any human social group-based endeavor in a cultural and social system such as political parties or communities.

Nonlinear system

A nonlinear system is one whose behavior can't be expressed as a sum of the behaviors of its parts (or of their multiples). In technical terms, the behavior of nonlinear systems is not subject to the principle of superposition. Linear systems are subject to superposition.

Topics on complex systems

Features of complex systems

Complex systems may have the following features:

Difficult to determine boundaries

It can be difficult to determine the boundaries of a complex system. The decision is ultimately made by the observer.

Complex systems may be open

Complex systems are usually open systems — that is, they exist in a thermodynamic gradient and dissipate energy. In other words, complex systems are frequently far from energetic equilibrium: but despite this flux, there may be pattern stability, see synergetics.

Complex systems may have a memory

The history of a complex system may be important. Because complex systems are dynamical systems they change over time, and prior states may have an influence on present states. More formally, complex systems often exhibit hysteresis.

Complex systems may be nested

The components of a complex system may themselves be complex systems. For example, an economy is made up of organisations, which are made up of people, which are made up of cells - all of which are complex systems.

Dynamic network of multiplicity

As well as coupling rules, the dynamic network of a complex system is important. Small-world or scale-free networks which have many local interactions and a smaller number of inter-area connections are often employed. Natural complex systems often exhibit such topologies. In the human cortex for example, we see dense local connectivity and a few very long axon projections between regions inside the cortex and to other brain regions.

May produce emergent phenomena

Complex systems may exhibit behaviors that are emergent, which is to say that while the results may be deterministic, they may have properties that can only be studied at a higher level. For example, the termites in a mound have physiology, biochemistry and biological development that are at one level of analysis, but their social behavior and mound building is a property that emerges from the collection of termites and needs to be analysed at a different level.

Relationships are non-linear

In practical terms, this means a small perturbation may cause a large effect (see butterfly effect), a proportional effect, or even no effect at all. In linear systems, effect is *always* directly proportional to cause. See nonlinearity.

Relationships contain feedback loops

Both negative (damping) and positive (amplifying) feedback are often found in complex systems. The effects of an element's behaviour are fed back to in such a way that the element itself is altered.

See also

- Agent based model
- Complex (Disambiguation)
- Complexity (disambiguation)
- Dissipative system
- System equivalence
- Systems theory

References

[1] Weaver, Warren (1948), " Science and Complexity (http://www.ceptualinstitute.com/genre/weaver/weaver-1947b.htm)", *American Scientist* **36**: 536 (Retrieved on 2007–11–21.),
[2] Luis M. Rocha, 1999.
[3] Science (http://www.sciencemag.org/content/vol284/issue5411/) Vol. 284. No. 5411 (1999)]
[4] History of Complex Systems (http://www.irit.fr/COSI/training/complexity-tutorial/history-of-complex-systems.htm)

Further reading

- Rocha, Luis M., *BITS: Computer and Communications News*. Computing, Information, and Communications Division. Los Alamos National Laboratory. November 1999.

External links

Articles/General Information

- Complex systems (http://www.scholarpedia.org/article/Complex_Systems) in scholarpedia.
- (European) Complex Systems Society (http://cssociety.org)
- (Australian) Complex systems research network. (http://www.complexsystems.net.au/)
- Complex Systems Modeling (http://informatics.indiana.edu/rocha/complex/csm.html) based on Luis M. Rocha, 1999.

Complexity

In general usage, **complexity** tends to be used to characterize something with many parts in intricate arrangement. In science there are at this time a number of approaches to characterizing complexity, many of which are reflected in this article. Seth Lloyd of M.I.T. writes that he once gave a presentation which set out 32 definitions of complexity.[1]

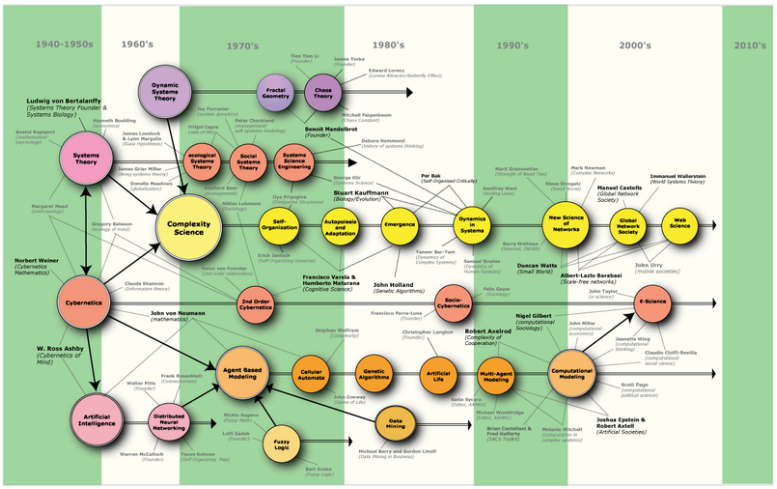
Definitions are often tied to the concept of a 'system' – a set of parts or elements which have relationships among them differentiated from relationships with other elements outside the relational regime. Many definitions tend to postulate or assume that complexity expresses a condition of numerous elements in a system and numerous forms of relationships among the elements. At the same time, what is complex and what is simple is relative and changes with time.

Some definitions key on the question of the probability of encountering a given condition of a system once characteristics of the system are specified. Warren Weaver has posited that the complexity of a particular system is the degree of difficulty in predicting the properties of the system if the properties of the system's parts are given. In Weaver's view, complexity comes in two forms: disorganized complexity, and organized complexity. [2] Weaver's paper has influenced contemporary thinking about complexity. [3]

The approaches which embody concepts of systems, multiple elements, multiple relational regimes, and state spaces might be summarized as implying that complexity arises from the number of distinguishable relational regimes (and their associated state spaces) in a defined system.

Some definitions relate to the algorithmic basis for the expression of a complex phenomenon or model or mathematical expression, as is later set out herein.

Disorganized complexity vs. organized complexity



'Map of Complexity Science. *HERE FOR WEB VERSION OF MAP [4] **The web version of this map provides internet links to all the leading scholars and areas of research in complexity science.**

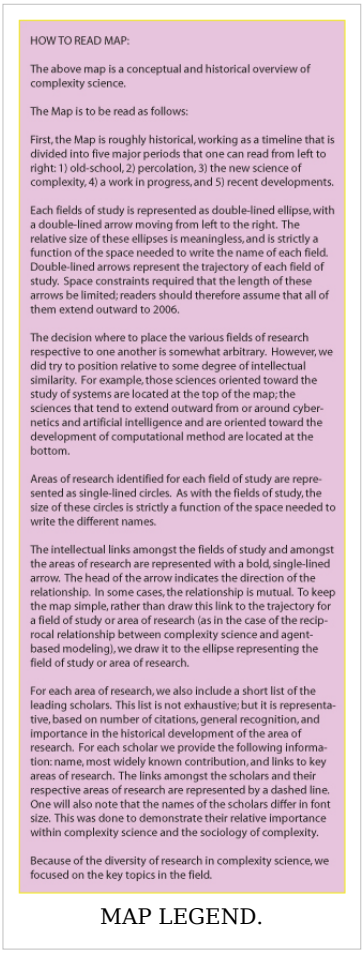
One of the problems in addressing complexity issues has been distinguishing conceptually between the large number of variances in relationships extant in random collections, and the sometimes large, but smaller, number of relationships between elements in systems where constraints (related to correlation of otherwise independent elements) simultaneously reduce the

variations from element independence and create distinguishable regimes of more-uniform, or correlated, relationships, or interactions.

Weaver perceived and addressed this problem, in at least a preliminary way, in drawing a distinction between 'disorganized complexity' and 'organized complexity'.

In Weaver's view, disorganized complexity results from the particular system having a very large number of parts, say millions of parts, or many more. Though the interactions of the parts in a 'disorganized complexity' situation can be seen as largely random, the properties of the system as a whole can be understood by using probability and statistical methods.

A prime example of disorganized complexity is a gas in a container, with the gas molecules as the parts. Some would suggest that a system of disorganized complexity may be compared, for example, with the (relative) simplicity of the planetary orbits – the latter can be known by applying Newton’s laws of motion, though this example involved highly correlated events.



MAP LEGEND.

Organized complexity, in Weaver's view, resides in nothing else than the non-random, or correlated, interaction between the parts. These non-random, or correlated, relationships create a differentiated structure which can, as a system, interact with other systems. The coordinated system manifests properties not carried by, or dictated by, individual parts. The organized aspect of this form of complexity vis a vis other systems than the subject system can be said to "emerge," without any “guiding hand.”

The number of parts does not have to be very large for a particular system to have emergent properties. A system of organized complexity may be understood in its properties (behavior among the properties) through modeling and simulation, particularly modeling and simulation with computers. An example of organized complexity is a city neighborhood as a living mechanism, with the neighborhood people among the system’s parts. [5]

Sources and factors of complexity

The source of disorganized complexity is the large number of parts in the system of interest, and the lack of correlation between elements in the system.

There is no consensus at present on general rules regarding the sources of organized complexity, though the lack of randomness implies correlations between elements. See e.g. Robert Ulanowicz's treatment of ecosystems. [6] Consistent with prior statements here, the number of parts (and types of parts) in the system and the number of relations between the parts would have to be non-trivial – however, there is no general rule to separate “trivial” from “non-trivial.

Complexity of an object or system is a relative property. For instance, for many functions (problems), such a computational complexity as time of computation is smaller when multitape Turing machines are used than when Turing machines with one tape are used. Random Access Machines allow one to even more decrease time complexity (Greenlaw and Hoover 1998: 226), while inductive Turing machines can decrease even the complexity class of a function, language or set (Burgin 2005). This shows that tools of activity can be an important factor of complexity.

Specific meanings of complexity

In several scientific fields, "complexity" has a specific meaning :

- In computational complexity theory, the amounts of resources required for the execution of algorithms is studied. The most popular types of computational complexity are the time complexity of a problem equal to the number of steps that it takes to solve an instance of the problem as a function of the size of the input (usually measured in bits), using the most efficient algorithm, and the space complexity of a problem equal to the volume of the memory used by the algorithm (e.g., cells of the tape) that it takes to solve an instance of the problem as a function of the size of the input (usually measured in bits), using the most efficient algorithm. This allows to classify computational problems by complexity class (such as P, NP ...). An axiomatic approach to computational complexity was developed by Manuel Blum. It allows one to deduce many properties of concrete computational complexity measures, such as time complexity or space complexity, from properties of axiomatically defined measures.
- In algorithmic information theory, the *Kolmogorov complexity* (also called *descriptive complexity*, *algorithmic complexity* or *algorithmic entropy*) of a string is the length of the shortest binary program which outputs that string. Different kinds of Kolmogorov complexity are studied: the uniform complexity, prefix complexity, monotone complexity, time-bounded Kolmogorov complexity, and space-bounded Kolmogorov complexity. An axiomatic approach to Kolmogorov complexity based on Blum axioms (Blum 1967) was introduced by Mark Burgin in the paper presented for publication by Andrey Kolmogorov (Burgin 1982). The axiomatic approach encompasses other approaches to Kolmogorov complexity. It is possible to treat different kinds of Kolmogorov complexity as particular cases of axiomatically defined generalized Kolmogorov complexity. Instead, of proving similar theorems, such as the basic invariance theorem, for each particular measure, it is possible to easily deduce all such results from one corresponding theorem proved in the axiomatic setting. This is a general advantage of the axiomatic approach in mathematics. The axiomatic approach to Kolmogorov complexity was further developed in the book (Burgin 2005) and applied to software metrics (Burgin and Debnath, 2003; Debnath and Burgin, 2003).
- In information processing, complexity is a measure of the total number of properties transmitted by an object and detected by an observer. Such a collection of properties is often referred to as a state.
- In physical systems, complexity is a measure of the probability of the state vector of the system. This should not be confused with entropy; it is a distinct mathematical measure, one in which two distinct states are never conflated and considered equal, as is done for the notion of entropy statistical mechanics.

- In mathematics, Krohn-Rhodes complexity is an important topic in the study of finite semigroups and automata.

There are different specific forms of complexity:

- In the sense of how complicated a problem is from the perspective of the person trying to solve it, limits of complexity are measured using a term from cognitive psychology, namely the hrair limit.
- Unruly complexity denotes situations that do not have clearly defined boundaries, coherent internal dynamics, or simply mediated relations with their external context, as coined by Peter Taylor.
- Complex adaptive system denotes systems which have some or all of the following attributes [7]
  - The number of parts (and types of parts) in the system and the number of relations between the parts is non-trivial – however, there is no general rule to separate “trivial” from “non-trivial;”
  - The system has memory or includes feedback;
  - The system can adapt itself according to its history or feedback;
  - The relations between the system and its environment are non-trivial or non-linear;
  - The system can be influenced by, or can adapt itself to, its environment; and
  - The system is highly sensitive to initial conditions.

Study of complexity

Complexity has always been a part of our environment, and therefore many scientific fields have dealt with complex systems and phenomena. Indeed, some would say that only what is somehow complex – what displays variation without being random – is worthy of interest.

The use of the term complex is often confused with the term complicated. In today’s systems, this is the difference between myriad connecting "stovepipes" and effective "integrated" solutions. [8] This means that complex is the opposite of independent, while complicated is the opposite of simple.

While this has led some fields to come up with specific definitions of complexity, there is a more recent movement to regroup observations from different fields to study complexity in itself, whether it appears in anthills, human brains, or stock markets. One such interndisciplinary group of fields is relational order theories.

Complexity topics

Complex behaviour

The behaviour of a complex system is often said to be due to emergence and self-organization. Chaos theory has investigated the sensitivity of systems to variations in initial conditions as one cause of complex behaviour.

Complex mechanisms

Recent developments around artificial life, evolutionary computation and genetic algorithms have led to an increasing emphasis on complexity and complex adaptive systems.

Complex simulations

In social science, the study on the emergence of macro-properties from the micro-properties, also known as macro-micro view in sociology. The topic is commonly recognized as social complexity that is often related to the use of computer simulation in social science, i.e.: computational sociology.

Complex systems

Systems theory has long been concerned with the study of complex systems (In recent times, *complexity theory* and *complex systems* have also been used as names of the field). These systems can be biological, economic, technological, etc. Recently, complexity is a natural domain of interest of the real world socio-cognitive systems and emerging systemics research. Complex systems tend to be high-dimensional, non-linear and hard to model. In specific circumstances they may exhibit low dimensional behaviour.

Complexity in data

In information theory, algorithmic information theory is concerned with the complexity of strings of data.

Complex strings are harder to compress. While intuition tells us that this may depend on the codec used to compress a string (a codec could be theoretically created in any arbitrary language, including one in which the very small command "X" could cause the computer to output a very complicated string like '18995316'"), any two Turing-complete languages can be implemented in each other, meaning that the length of two encodings in different languages will vary by at most the length of the "translation" language - which will end up being negligible for sufficiently large data strings.

These algorithmic measures of complexity tend to assign high values to random noise. However, those studying complex systems would not consider randomness as complexity.

Information entropy is also sometimes used in information theory as indicative of complexity.

Applications of complexity

Computational complexity theory is the study of the complexity of problems - that is, the difficulty of solving them. Problems can be classified by complexity class according to the time it takes for an algorithm - usually a computer program - to solve them as a function of the problem size. Some problems are difficult to solve, while others are easy. For example, some difficult problems need algorithms that take an exponential amount of time in terms of the size of the problem to solve. Take the travelling salesman problem, for example. It can be solved in time $O(n^2 2^n)$ (where n is the size of the network to visit - let's say the number of cities the travelling salesman must visit exactly once). As the size of the network of cities grows, the time needed to find the route grows (more than) exponentially.

Even though a problem may be computationally solvable in principle, in actual practice it may not be that simple. These problems might require large amounts of time or an inordinate amount of space. Computational complexity may be approached from many different aspects. Computational complexity can be investigated on the basis of time, memory or other resources used to solve the problem. Time and space are two of the most important and popular considerations when problems of complexity are analyzed.

There exist a certain class of problems that although they are solvable in principle they require so much time or space that it is not practical to attempt to solve them. These problems are called intractable.

There is another form of complexity called hierarchical complexity. It is orthogonal to the forms of complexity discussed so far, which are called horizontal complexity

See also

- Chaos theory
- Command and Control Research Program
- Complexity theory (disambiguation page)
- Cyclomatic complexity
- Evolution of complexity
- Game complexity
- Holism in science
- Interconnectedness
- Model of hierarchical complexity
- Occam's razor
- Process architecture
- Programming Complexity
- Sociology and complexity science
- Systems theory
- Variety (cybernetics)

References

[1] Lloyd, Seth (2006). *Programming the Universe*. Knopf. ISBN 978-1400033867..
[2] Weaver, Warren (1948), " Science and Complexity (http://www.ceptualinstitute.com/genre/weaver/weaver-1947b.htm)", *American Scientist* **36**: 536 (Retrieved on 2007–11–21.),
[3] Johnson, Steven (2001). *Emergence: the connected lives of ants, brains, cities, and software*. New York: Scribner. pp. p.46. ISBN 0-684-86875-X..
[4] http://www.art-sciencefactory.com/complexity-map_feb09.html'''CLICK
[5] Jacobs, Jane (1961). *The Death and Life of Great American Cities*. New York: Random House.
[6] Ulanowicz, Robert, "Ecology, the Ascendant Perspective", Columbia, 1997
[7] Johnson, Neil F. (2007). *Two's Company, Three is Complexity: A simple guide to the science of all sciences*. Oxford: Oneworld. ISBN 978-1-85168-488-5.
[8] (Lissack and Roos, 2000)

Further reading

- Lewin, Roger (1992). *Complexity: Life at the Edge of Chaos*. New York: Macmillan Publishing Co. ISBN 9780025704855.
- Waldrop, M. Mitchell (1992). *Complexity: The Emerging Science at the Edge of Order and Chaos*. New York: Simon & Schuster. ISBN 9780671767891.
- Czerwinski, Tom; David Alberts (1997). *Complexity, Global Politics, and National Security (http://www.dodccrp.org/files/Alberts_Complexity_Global.pdf)*. National Defense University. ISBN 9781579060466.
- Czerwinski, Tom (1998). *Coping with the Bounds: Speculations on Nonlinearity in Military Affairs (http://www.dodccrp.org/files/Czerwinski_Coping.pdf)*. CCRP. ISBN 9781414503158 (from Pavilion Press, 2004).

- Lissack, Michael R.; Johan Roos (2000). *The Next Common Sense, The e-Manager's Guide to Mastering Complexity*. Intercultural Press. ISBN 9781857882353.
- Solé, R. V.; B. C. Goodwin (2002). *Signs of Life: How Complexity Pervades Biology*. Basic Books. ISBN 9780465019281.
- Moffat, James (2003). *Complexity Theory and Network Centric Warfare (http://www.dodccrp.org/files/Moffat_Complexity.pdf)*. CCRP. ISBN 9781893723115.
- Smith, Edward (2006). *Complexity, Networking, and Effects Based Approaches to Operations (http://www.dodccrp.org/files/Smith_Complexity.pdf)*. CCRP. ISBN 9781893723184.
- Heylighen, Francis (2008), " Complexity and Self-Organization (http://pespmc1.vub.ac.be/Papers/ELIS-Complexity.pdf)", in Bates, Marcia J.; Maack, Mary Niles, *Encyclopedia of Library and Information Sciences*, CRC, ISBN 9780849397127
- Greenlaw, N. and Hoover, H.J. *Fundamentals of the Theory of Computation*, Morgan Kauffman Publishers, San Francisco, 1998
- Blum, M. (1967) On the Size of Machines, *Information and Control*, v. 11, pp. 257-265
- Burgin, M. (1982) Generalized Kolmogorov complexity and duality in theory of computations, Notices of the Russian Academy of Sciences, v.25, No. 3, pp.19-23
- Mark Burgin (2005), *Super-recursive algorithms*, Monographs in computer science, Springer.
- Burgin, M. and Debnath, N. Hardship of Program Utilization and User-Friendly Software, in *Proceedings of the International Conference "Computer Applications in Industry and Engineering"*, Las Vegas, Nevada, 2003, pp. 314-317
- Debnath, N.C. and Burgin, M., (2003) Software Metrics from the Algorithmic Perspective, in *Proceedings of the ISCA 18th International Conference "Computers and their Applications"*, Honolulu, Hawaii, pp. 279-282
- Meyers, R.A., (2009) "Encyclopedia of Complexity and Systems Science", ISBN 978-0-387-75888-6

External links

- Quantifying Complexity Theory (http://www.calresco.org/lucas/quantify.htm) - classification of complex systems
- Complexity Measures (http://cscs.umich.edu/~crshalizi/notebooks/complexity-measures.html) - an article about the abundance of not-that-useful complexity measures.
- UC Four Campus Complexity Videoconferences (http://eclectic.ss.uci.edu/~drwhite/center/cac.html) - Human Sciences and Complexity
- Complexity Digest (http://www.comdig.com) - networking the complexity community
- The Santa Fe Institute (http://www.santafe.edu/) - engages in research in complexity related topics

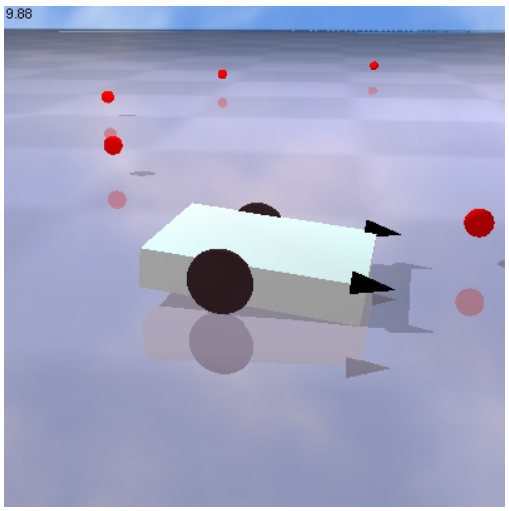
Complex systems

This article describes the new science of complexity, which treats complex systems as field of science. For other meanings, see complex system. For Complex Systems journal, see Complex Systems (journal)

Complex systems is a scientific field which studies the common properties of systems that are considered fundamentally complex. Such systems may exist in nature, society, science and other many fields. It is also called *complex systems theory*, *complexity science*, *study of complex systems*, *sciences of complexity*, *non-equilibrium physics*, and *historical physics*. The key problems of such systems are difficulties with their formal modeling and simulation. From such perspective, in different research contexts complex systems are defined on the base of their different attributes. At present, the consensus related to one universal definition of *complex system* does not exist yet.

Overview

The study of complex systems is bringing a new approach to the many scientific questions that are a poor fit for the usual mechanistic view of reality present in science [1] . *Complex systems* is therefore often used as a broad term encompassing a research approach to problems in many diverse disciplines including anthropology, artificial life, chemistry, computer science, economics, evolutionary computation, earthquake prediction, meteorology, molecular biology, neuroscience, physics, psychology and sociology.



A Braitenberg simulation, programmed in breve, an artificial life simulator.

In these endeavors, scientists often seek simple non-linear coupling rules which *lead to* complex phenomena (rather than describe - see above), but this need not be the case. Human societies (and probably human brains) are complex systems in which neither the components nor the couplings are simple. Nevertheless, they exhibit many of the hallmarks of complex systems. It is worth remarking that non-linearity is not a necessary feature of complex systems modeling: macro-analyses that concern unstable equilibrium and evolution processes of certain biological/social/economic systems can usefully be carried out also by sets of linear equations, which do nevertheless entail reciprocal dependence between variable parameters.

Traditionally, engineering has striven to keep its systems linear, because that makes them simpler to build and to predict. However, many physical systems (for example lasers) are inherently "complex systems" in terms of the definition above, and engineering practice must now include elements of complex systems research.

Information theory applies well to the complex adaptive systems, CAS, through the concepts of object oriented design, as well as through formalized concepts of organization and disorder that can be associated with any systems evolution process.

History

Complex Systems is a new approach to science that studies how relationships between parts give rise to the collective behaviors of a system and how the system interacts and forms relationships with its environment.

The earliest precursor to modern complex systems theory can be found in the classical political economy of the Scottish Enlightenment, later developed by the Austrian school of economics, which says that order in market systems is spontaneous (or emergent) in that it is the result of human action, but not the execution of any human design.[2] [3]

Upon this the Austrian school developed from the 19th to the early 20th century the economic calculation problem, along with the concept of dispersed knowledge, which were to fuel debates against the then-dominant Keynesian economics. This debate would notably lead economists, politicians and other parties to explore the question of computational complexity.

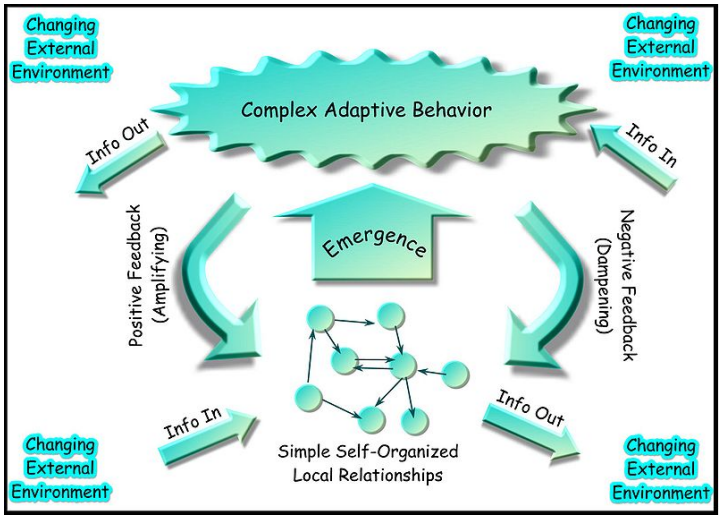
A pioneer in the field, and inspired by Karl Popper's and Warren Weaver's works, Nobel prize economist and philosopher Friedrich Hayek dedicated much of his work, from early to the late 20th century, to the study of complex phenomena,[4] not constraining his work to human economies but to other fields such as psychology,[5] biology and cybernetics.

Further Steven Strogatz from *Sync* stated that "every decade or so, a grandiose theory comes along, bearing similar aspirations and often brandishing an ominous-sounding C-name. In the 1960s it was cybernetics. In the '70s it was catastrophe theory. Then came chaos theory in the '80s and complexity theory in the '90s."

Topics in the complex systems study

Complexity and modeling

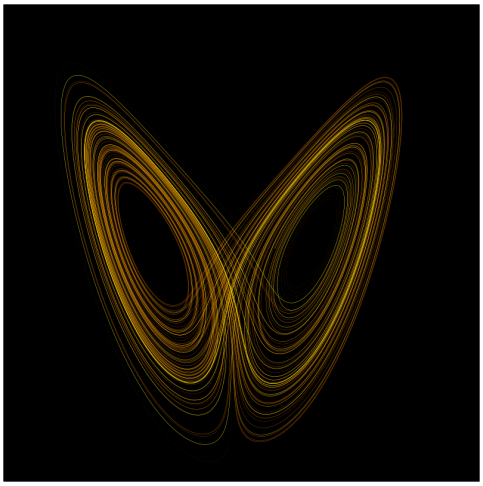
One of Hayek's main contributions to early complexity theory is his distinction between the human capacity to predict the behaviour of simple systems and its capacity to predict the behaviour of complex systems through modeling. He believed that economics and the sciences of complex phenomena in general, which in his view included biology, psychology, and so on, could not be modeled after the sciences that deal with essentially simple phenomena like physics.[6] Hayek would notably explain that complex phenomena, through modeling, can only allow pattern predictions, compared with the precise predictions that can be made out of non-complex phenomena.[7]



A way of modelling a Complex Adaptive System

Complexity and chaos theory

Complexity theory is rooted in Chaos theory, which in turn has its origins more than a century ago in the work of the French mathematician Henri Poincaré. Chaos is sometimes viewed as extremely complicated information, rather than as an absence of order.[8] The point is that chaos remains deterministic. With perfect knowledge of the initial conditions and of the context of an action, the course of this action can be predicted in chaos theory. As argued by Ilya Prigogine,[9] Complexity is non-deterministic, and gives no way whatsoever to predict the future. The emergence of complexity theory shows a domain between deterministic order and randomness which is complex.[10] This is referred as the 'edge of chaos'.[11]



A plot of the Lorenz attractor

When one analyses complex systems, sensitivity to initial conditions, for example, is not an issue as important as within the chaos theory in which it prevails. As stated by Colander,[12] the study of complexity is the opposite of the study of chaos. Complexity is about how a huge number of extremely complicated and dynamic set of relationships can generate some simple behavioural patterns, whereas chaotic behaviour, in the sense of deterministic chaos, is the result of a relatively small number of non-linear interactions.[10]

Therefore, the main difference between Chaotic systems and complex systems is their history.[13] Chaotic systems don’t rely on their history as complex ones do. Chaotic behaviour pushes a system in equilibrium into chaotic order, which means, in other words, out of what we traditionally define as 'order'. On the other hand, complex systems evolve far from equilibrium at the edge of chaos. They evolve at a critical state built up by a history of irreversible and unexpected events. In a sense chaotic systems can be regarded as a subset of complex systems distinguished precisely by this absence of historical dependence. Many real complex systems are, in practice and over long but finite time periods, robust. However, they do possess the potential for radical qualitative change of kind whilst retaining systemic integrity. Metamorphosis serves as perhaps more than a metaphor for such transformations.

Research centers, conferences, and journals

Institutes and research centers

- New England Complex Systems Institute
- Santa Fe Institute
- Center for Social Dynamics & Complexity (CSDC) at Arizona State University [14]
- Southampton Institute for Complex Systems Simulation [15]
- Center for the Study of Complex Systems at the University of Michigan [16]
- Center for Complex Systems and Brain Sciences at Florida Atlantic University [17]

Journals

- Complex Systems journal
- Interdisciplinary Description of Complex Systems journal

See also

- Cognitive Science
- Complex adaptive system
- Complexity
- Complexity economics
- Dynamical system
- Dynamical systems theory
- Emergence
- Enterprise systems engineering
- Generative sciences
- Multi-agent system
- Nonlinearity
- Pattern oriented modeling
- Process architecture
- Systems theory
- Systems theory in anthropology
- Self organization
- Sociology and complexity science

References

[1] http://www.narberthpa.com/Bale/lsbale_dop/cybernet.htm Bale, L.S. 1995, *Gregory Bateson, Cybernetics and the Social/Behavioral Sciences*
[2] Ferguson, Adam (1767). *An Essay on the History of Civil Society (http://oll.libertyfund.org/index.php?option=com_staticxt&staticfile=show.php?title=1428&Itemid=28)*. London: T. Cadell. art Third, Section II, p. 205. .
[3] Friedrich Hayek, *The Results of Human Action but Not of Human Design*, in *New Studies in Philosophy, Politics, Economics*, Chicago: University of Chicago Press, (1978), pp. 96-105.
[4] Bruce J. Caldwell, Popper and Hayek: Who influenced whom? (http://www.unites.uqam.ca/philo/pdf/Caldwell_2003-01.pdf), Karl Popper 2002 Centenary Congress, 2002.
[5] Friedrich von Hayek, *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, The University of Chicago Press, 1952.
[6] Reason Magazine - The Road from Serfdom (http://www.reason.com/news/show/33304.html)
[7] Friedrich August von Hayek - Prize Lecture (http://nobelprize.org/nobel_prizes/economics/laureates/1974/hayek-lecture.html)
[8] Hayles, N. K. (1991). *Chaos Bound: Orderly Disorder in Contemporary Literature and Science*. Cornell University Press, Ithaca, NY.
[9] Prigogine, I. (1997). *The End of Certainty*, The Free Press, New York.
[10] Cilliers, P. (1998). *Complexity and Postmodernism: Understanding Complex Systems*, Routledge, London.
[11] Per Bak (1996). *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, New York, USA.
[12] Colander, D. (2000). *The Complexity Vision and the Teaching of Economics*, E. Elgar, Northampton, MA.
[13] Buchanan, M.(2000). *Ubiquity : Why catastrophes happen*, three river press, New-York.
[14] http://csdc.asu.edu/
[15] http://www.icss.soton.ac.uk
[16] http://www.cscs.umich.edu/
[17] http://www.ccs.fau.edu/

Further reading

- L.A.N. Amaral and J.M. Ottino, *Complex networks — augmenting the framework for the study of complex system* (http://amaral.northwestern.edu/Publications/Papers/Amaral-2004-Eur.Phys.J.B-38-147.pdf), 2004.
- Murray Gell-Mann, *Let's Call It Plectics* (http://www.santafe.edu/~mgm/plectics.html), 1995/96.
- Nigel Goldenfeld and Leo P. Kadanoff, *Simple Lessons from Complexity* (http://guava.physics.uiuc.edu/~nigel/articles/complexity.html), 1999
- A. Gogolin, A. Nersesyan and A. Tsvelik, *Theory of strongly correlated systems* (http://www.cmth.bnl.gov/~tsvelik/theory.html), Cambridge University Press, 1999.

- Kelly, K. (1995). *Out of Control* (http://www.kk.org/outofcontrol/contents.php), Perseus Books Group.
- Graeme Donald Snooks, "A general theory of complex living systems: Exploring the demand side of dynamics", *Complexity*, vol. 13, no. 6, July/August 2008.
- Sorin Solomon and Eran Shir, *Complexity; a science at 30* (http://www.europhysicsnews.org/index.php?option=article&access=standard&Itemid=129&url=/articles/epn/abs/2003/02/epn03204/epn03204.html), 2003.

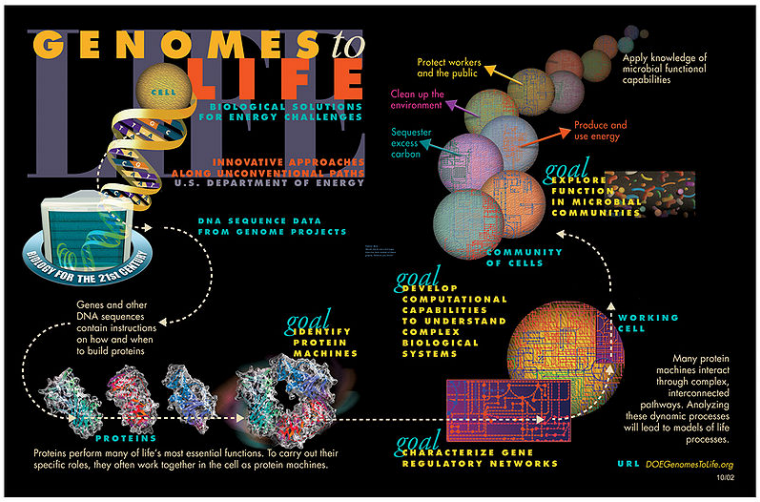
External links

- The Open Agent-Based Modeling Consortium (http://www.openabm.org)
- Complexity Science Focus (http://www.complexity.ecs.soton.ac.uk/)

Complex Systems Biology

Systems biology

Systems biology is a biology-based inter-disciplinary study field that focuses on the systematic study of complex interactions in biological systems, thus using a new perspective (holism instead of reduction) to study them. Particularly from year 2000 onwards, the term is used widely in the biosciences, and in a variety of contexts. Because the scientific method has been used primarily toward reductionism, one of the goals of systems biology is to discover new emergent properties that may arise from the systemic view used by this discipline in order to understand better the entirety of processes that happen in a biological system.



Example of systems biology research.

Overview

Systems biology can be considered from a number of different aspects:

- Some sources discuss systems biology as a **field of study**, particularly, the study of the interactions between the components of *biological systems*, and how these interactions give rise to the function and behavior of that system (for example, the enzymes and metabolites in a metabolic pathway).[1] [2]
- Other sources consider systems biology as a **paradigm**, usually defined in antithesis to the so-called reductionist paradigm, although fully consistent with the scientific method. The distinction between the two paradigms is referred to in these quotations:

  "*The reductionist approach has successfully identified most of the components and many of the interactions but, unfortunately, offers no convincing concepts or methods to understand how system properties emerge...the pluralism of causes and effects in biological networks is better addressed by observing, through quantitative measures, multiple components simultaneously and by rigorous data integration with mathematical models*" Science[3]

  "*Systems biology...is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different....It means changing our philosophy, in the full sense of the term*" Denis Noble[4]

- Still other sources view systems biology in terms of the **operational protocols used for performing research**, namely a cycle composed of theory, analytic or computational modelling to propose specific testable hypotheses about a biological system, experimental validation, and then using the newly acquired quantitative description of cells or cell processes to refine the computational model or theory.[5] [6] Since the objective is a model of the interactions in a system, the experimental techniques that most suit systems biology are those that are system-wide and attempt to be as complete as possible. Therefore, transcriptomics, metabolomics, proteomics and high-throughput techniques are used to collect quantitative data for the construction and validation of models.
- Engineers consider systems biology as the application of dynamical systems theory to molecular biology.
- Finally, some sources see it as a **socioscientific phenomenon** defined by the strategy of pursuing integration of complex data about the interactions in biological systems from diverse experimental sources using interdisciplinary tools and personnel.

This variety of viewpoints is illustrative of the fact that systems biology refers to a cluster of peripherally overlapping concepts rather than a single well-delineated field. However the term has widespread currency and popularity as of 2007, with chairs and institutes of systems biology proliferating worldwide (Such as the Institute for Systems Biology).

History

Systems biology finds its roots in:

- the quantitative modelling of enzyme kinetics, a discipline that flourished between 1900 and 1970,
- the simulations developed to study neurophysiology, and
- control theory and cybernetics.

One of the theorists who can be seen as a precursor of systems biology is Ludwig von Bertalanffy with his general systems theory. One of the first numerical simulations in biology was published in 1952 by the British neurophysiologists and Nobel prize winners Alan Lloyd Hodgkin and Andrew Fielding Huxley, who constructed a mathematical model that explained the action potential propagating along the axon of a neuronal cell.[7] Their model described a cellular function emerging from the interaction between two different molecular components, a potassium and a sodium channels, and can therefore be seen as the beginning of computational systems biology.[8] In 1960, Denis Noble developed the first computer model of the heart pacemaker.[9]

The formal study of systems biology, as a distinct discipline, was launched by systems theorist Mihajlo Mesarovic in 1966 with an international symposium at the Case Institute of Technology in Cleveland, Ohio entitled "Systems Theory and Biology."[10] [11]

The 1960s and 1970s saw the development of several approaches to study complex molecular systems, such as the Metabolic Control Analysis and the biochemical systems theory. The successes of molecular biology throughout the 1980s, coupled with a skepticism toward theoretical biology, that then promised more than it achieved, caused the quantitative modelling of biological processes to become a somewhat minor field.

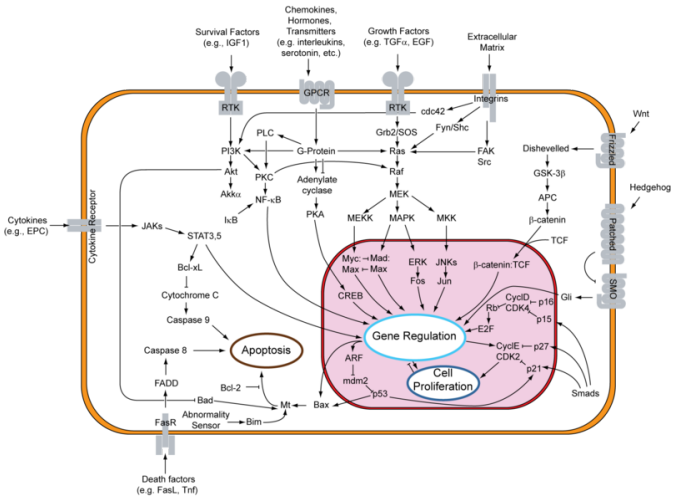
However the birth of functional genomics in the 1990s meant that large quantities of high quality data became available, while the computing power exploded, making more realistic

models possible. In 1997, the group of Masaru Tomita published the first quantitative model of the metabolism of a whole (hypothetical) cell.

Around the year 2000, when Institutes of Systems Biology were established in Seattle and Tokyo, systems biology emerged as a movement in its own right, spurred on by the completion of various genome projects, the large increase in data from the omics (e.g. genomics and proteomics) and the accompanying advances in high-throughput experiments and bioinformatics. Since then, various research institutes dedicated to systems biology have been developed. As of summer 2006, due to a shortage of people in systems biology[12] several doctoral training centres in systems biology have been established in many parts of the world.

Techniques associated with systems biology

According to the interpretation of System Biology as the ability to obtain, integrate and analyze complex data from multiple experimental sources using interdisciplinary tools, some typical technology platforms are:



Overview of signal transduction pathways

- Transcriptomics: whole cell or tissue gene expression measurements by DNA microarrays or serial analysis of gene expression
- Proteomics: complete identification of proteins and protein expression patterns of a cell or tissue through two-dimensional gel electrophoresis and mass spectrometry or multi-dimensional protein identification techniques (advanced HPLC systems coupled with mass spectrometry). Sub disciplines include phosphoproteomics, glycoproteomics and other methods to detect chemically modified proteins.
- Metabolomics: identification and measurement of all small-molecules metabolites within a cell or tissue
- Glycomics: identification of the entirety of all carbohydrates in a cell or tissue.

In addition to the identification and quantification of the above given molecules further techniques analyze the dynamics and interactions within a cell. This includes:

- Interactomics which is used mostly in the context of protein-protein interaction but in theory encompasses interactions between all molecules within a cell,
- Fluxomics, which deals with the dynamic changes of molecules within a cell over time,
- Biomics: systems analysis of the biome.

The investigations are frequently combined with large scale perturbation methods, including gene-based (RNAi, mis-expression of wild type and mutant genes) and chemical approaches using small molecule libraries. Robots and automated sensors enable such large-scale experimentation and data acquisition. These technologies are still emerging and many face problems that the larger the quantity of data produced, the lower the quality. A

wide variety of quantitative scientists (computational biologists, statisticians, mathematicians, computer scientists, engineers, and physicists) are working to improve the quality of these approaches and to create, refine, and retest the models to accurately reflect observations.

The investigations of a single level of biological organization (such as those listed above) are usually referred to as Systematic Systems Biology. Other areas of Systems Biology includes Integrative Systems Biology, which seeks to integrate different types of information to advance the understanding the biological whole, and Dynamic Systems Biology, which aims to uncover how the biological whole changes over time (during evolution, for example, the onset of disease or in response to a perturbation). Functional Genomics may also be considered a sub-field of Systems Biology.

The systems biology approach often involves the development of mechanistic models, such as the reconstruction of dynamic systems from the quantitative properties of their elementary building blocks.[13] [14] For instance, a cellular network can be modelled mathematically using methods coming from chemical kinetics and control theory. Due to the large number of parameters, variables and constraints in cellular networks, numerical and computational techniques are often used. Other aspects of computer science and informatics are also used in systems biology. These include new forms of computational model, such as the use of process calculi to model biological processes, the integration of information from the literature, using techniques of information extraction and text mining, the development of online databases and repositories for sharing data and models (such as BioModels Database), approaches to database integration and software interoperability via loose coupling of software, websites and databases[15] and the development of syntactically and semantically sound ways of representing biological models, such as the Systems Biology Markup Language (SBML).

See also

Related fields

- Complex systems biology
- Complex systems
- Complex systems biology
- Bioinformatics
- Biological network inference
- Biological systems engineering
- Biomedical cybernetics
- Biostatistics
- Theoretical Biophysics
- Relational Biology
- Translational Research
- Computational biology
- Computational systems biology
- Scotobiology
- Synthetic biology
- Systems biology modeling
- Systems ecology
- Systems immunology

Related terms

- Life
- Artificial life
- Gene regulatory network
- Metabolic network modelling
- Living systems theory
- Network Theory of Aging
- Regulome
- Systems Biology Markup Language (SBML)
- SBO
- Viable System Model
- Antireductionism

Systems biologists

- Category:Systems biologists

Lists

- Category:Systems biologists
- List of systems biology conferences
- List of omics topics in biology
- List of publications in systems biology
- List of systems biology research groups

References

[1] Snoep J.L. and Westerhoff H.V.; Alberghina L. and Westerhoff H.V. (Eds.) (2005.). "From isolation to integration, a systems biology approach for building the Silicon Cell". *Systems Biology: Definitions and Perspectives*: p7, Springer-Verlag.

[2] "Systems Biology - the 21st Century Science" (http://www.systemsbiology.org/Intro_to_ISB_and_Systems_Biology/Systems_Biology_--_the_21st_Century_Science). .

[3] Sauer, U. et al. (27 April 2007). "Getting Closer to the Whole Picture". *Science* **316**: 550. doi: 10.1126/science.1142502 (http://dx.doi.org/10.1126/science.1142502). PMID 17463274.

[4] Denis Noble (2006). *The Music of Life: Biology beyond the genome*. Oxford University Press. ISBN 978-0199295739. p21

[5] "Systems Biology: Modelling, Simulation and Experimental Validation" (http://www.bbsrc.ac.uk/science/areas/ebs/themes/main_sysbio.html). .

[6] Kholodenko B.N., Bruggeman F.J., Sauro H.M.; Alberghina L. and Westerhoff H.V.(Eds.) (2005.). "Mechanistic and modular approaches to modeling and inference of cellular regulatory networks". *Systems Biology: Definitions and Perspectives*: p143, Springer-Verlag.

[7] Hodgkin AL, Huxley AF (1952). "A quantitative description of membrane current and its application to conduction and excitation in nerve". *J Physiol* **117**: 500–544. PMID 12991237.

[8] Le Novere (2007). "The long journey to a Systems Biology of neuronal function". *BMC Systems Biology* **1**: 28. doi: 10.1186/1752-0509-1-28 (http://dx.doi.org/10.1186/1752-0509-1-28).

[9] Noble D (1960). "Cardiac action and pacemaker potentials based on the Hodgkin-Huxley equations". *Nature* **188**: 495–497. doi: 10.1038/188495b0 (http://dx.doi.org/10.1038/188495b0). PMID 13729365.

[10] Mesarovic, M. D. (1968). *Systems Theory and Biology*. Springer-Verlag.

[11] " A Means Toward a New Holism (http://www.jstor.org/view/00368075/ap004022/00a00220/0)". *Science* **161** (3836): 34-35. doi: 10.1126/science.161.3836.34 (http://dx.doi.org/10.1126/science.161.3836.34). .

[12] "Working the Systems" (http://sciencecareers.sciencemag.org/career_development/previous_issues/articles/2006_03_03/working_the_systems/(parent)/158). .

[13] Gardner, TS; di Bernardo D, Lorenz D and Collins JJ (4 July 2003). "Inferring genetic networks and identifying compound of action via expression profiling". *Science* **301**: 102-1005. doi: 10.1126/science.1081900 (http://dx.doi.org/10.1126/science.1081900). PMID 12843395.

[14] di Bernardo, D; Thompson MJ, Gardner TS, Chobot SE, Eastwood EL, Wojtovich AP, Elliot SJ, Schaus SE and Collins JJ (March 2005). "Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks". *Nature Biotechnology* **23**: 377-383. doi: 10.1038/nbt1075 (http://dx.doi.org/10.1038/nbt1075). PMID 15765094.

[15] such as Gaggle (http://gaggle.systemsbiology.net), SBW (http://sys-bio.org)), or commercial suits, e.g., MetaCore (http://www.genego.com/metacore.php) and MetaDrug (http://www.genego.com/metadrug.php)

Further reading

Books

- Hiroaki Kitano (editor). *Foundations of Systems Biology.* MIT Press: 2001. ISBN 0-262-11266-3
- CP Fall, E Marland, J Wagner and JJ Tyson (Editors). "Computational Cell Biology." Springer Verlag: 2002 ISBN 0-387-95369-8
- G Bock and JA Goode (eds).*In Silico*" *Simulation of Biological Processes*, Novartis Foundation Symposium 247. John Wiley & Sons: 2002. ISBN 0-470-84480-9
- E Klipp, R Herwig, A Kowald, C Wierling, and H Lehrach. *Systems Biology in Practice.* Wiley-VCH: 2005. ISBN 3-527-31078-9
- L. Alberghina and H. Westerhoff (Editors) – *Systems Biology: Definitions and Perspectives*, Topics in Current Genetics 13, Springer Verlag (2005), ISBN 978-3540229681
- A Kriete, R Eils. *Computational Systems Biology.*, Elsevier - Academic Press: 2005. ISBN 0-12-088786-X

- K. Sneppen and G. Zocchi, (2005) *Physics in Molecular Biology*, Cambridge University Press, ISBN 0-521-84419-3
- D. Noble, *The Music of life. Biology beyond the genome* Oxford University Press (http://www.musicoflife.co.uk/) 2006. ISBN 0199295735, ISBN 978-0199295739
- Z. Szallasi, J. Stelling, and V.Periwal (eds.) System Modeling in Cellular Biology: From Concepts to Nuts and Bolts (Hardcover), MIT Press: 2006, ISBN 0-262-19548-8
- B Palsson, *Systems Biology - Properties of Reconstructed Networks.* Cambridge University Press: 2006. (http://gcrg.ucsd.edu/book/index.html) ISBN 978-0-521-85903-5
- K Kaneko. *Life: An Introduction to Complex Systems Biology.* Springer: 2006. ISBN 3540326669
- U Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits.* CRC Press: 2006. ISBN 1-58488-642-0 - emphasis on Network Biology (For a comparative review of Alon, Kaneko and Palsson see Werner, E. (March 29, 2007). " All systems go (http://www.nature.com/nature/journal/v446/n7135/pdf/446493a.pdf)" (PDF). *Nature* **446**: 493-494. doi: 10.1038/446493a (http://dx.doi.org/10.1038/446493a). http://www.nature.com/nature/journal/v446/n7135/pdf/446493a.pdf.)
- Andriani Daskalaki (editor) "Handbook of Research on Systems Biology Applications in Medicine" Medical Information Science Reference, October 2008 ISBN 978-1-60566-076-9

Journals

- BMC Systems Biology (http://www.biomedcentral.com/bmcsystbiol) - open access journal on systems biology
- Molecular Systems Biology (http://www.nature.com/msb) - open access journal on systems biology
- IET Systems Biology (http://www.ietdl.org/IET-SYB) - not open access journal on systems biology

Articles

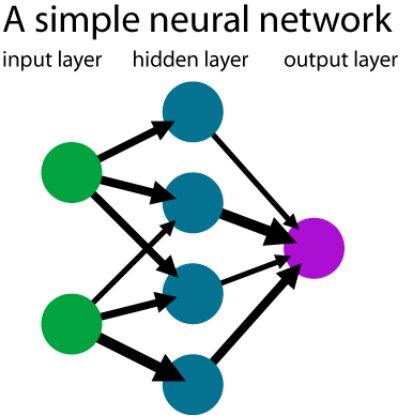
- Binnewies, Tim Terence, Miller, WG, Wang, G. *The complete genome sequence and analysis of the human pathogen Campylobacter lari* (http://www.bio.dtu.dk/English/Publications/1/all.aspx?lg=showcommon&id=231324). Published in journal: Foodborne Pathog Disease (ISSN 1535-3141) , vol: 5, issue: 4, pages: 371-386, 2008, Mary Ann Liebert, Inc. Publishers.
- M. Tomita, Hashimoto K, Takahashi K, Shimizu T, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC, Hutchison CA. E-CELL: Software Environment for Whole Cell Simulation. Genome Inform Ser Workshop Genome Inform. 1997;8:147-155. (http://web.sfc.keio.ac.jp/~mt/mt-lab/publications/Paper/ecell/bioinfo99/btc007_gml.html)
- ScienceMag.org (http://www.sciencemag.org/content/vol295/issue5560/) - Special Issue: Systems Biology, *Science*, Vol 295, No 5560, March 1, 2002
- Marc Vidal and Eileen E. M. Furlong. Nature Reviews Genetics 2004 From OMICS to systems biology (http://www.nature.com/nrg/journal/v5/n10/poster/omics/index.html)
- Marc Facciotti, Richard Bonneau, Leroy Hood and Nitin Baliga. Current Genomics 2004 Systems Biology Experimental Design - Considerations for Building Predictive Gene Regulatory Network Models for Prokaryotic Systems (http://www.ingentaconnect.com/

content/ben/cg/2004/00000005/00000007/art00002)
- Katia Basso, Adam A Margolin, Gustavo Stolovitzky, Ulf Klein, Riccardo Dalla-Favera, Andrea Califano, (2005) "Reverse engineering of regulatory networks in human B cells" (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=15778709&query_hl=7). Nat Genet;37(4):382-90
- Mario Jardon Systems Biology: An Overview (http://www.scq.ubc.ca/?p=253) - a review from the Science Creative Quarterly, 2005
- Johnjoe McFadden, Guardian.co.uk (http://www.guardian.co.uk/life/science/story/0,12996,1477776,00.html) - 'The unselfish gene: The new biology is reasserting the primacy of the whole organism - the individual - over the behaviour of isolated genes', *The Guardian* (May 6, 2005)
- Pharoah, M.C. (online). Looking to systems theory for a reductive explanation of phenomenal experience and evolutionary foundations for higher order thought (http://homepage.ntlworld.com/m.pharoah/) Retrieved Jan, 15 2008.
- WTEC Panel Report on International Research and Development in Systems Biology (http://www.wtec.org/sysbio/welcome.htm) (2005)
- E. Werner, "The Future and Limits of Systems Biology", Science STKE (http://stke.sciencemag.org/content/vol2005/issue278/) 2005, pe16 (2005).
- Francis J. Doyle and Jörg Stelling, "Systems interface biology" (http://www.journals.royalsoc.ac.uk/openurl.asp?genre=article&doi=10.1098/rsif.2006.0143) *J. R. Soc. Interface* Vol 3, No 10 2006
- Kahlem, P. and Birney E. (2006). "Dry work in a wet world: computation in systems biology." Mol Syst Biol 2: 40. (http://www.nature.com/doifinder/10.1038/msb4100080)
- E. Werner, "All systems go" (http://www.nature.com/nature/journal/v446/n7135/pdf/446493a.pdf), "Nature" (http://www.nature.com/nature/journal/v446/n7135/index.html) vol 446, pp 493–494, March 29, 2007. (Review of three books (Alon, Kaneko, and Palsson) on systems biology.)
- Santiago Schnell, Ramon Grima, Philip K. Maini, "Multiscale Modeling in Biology" (http://www.americanscientist.org/template/AssetDetail/assetid/54784), American Scientist, Vol 95, pages 134-142, March-April 2007.
- TS Gardner, D di Bernardo, D Lorenz and JJ Collins. "Inferring genetic networks and identifying compound of action via expression profiling." (http://www.bu.edu/abl/publications.html) Science 301: 102-105 (2003).
- Jeffery C. Way and Pamela A. Silver, Why We Need Systems Biology (http://cs.calstatela.edu/wiki/images/9/9b/Silver.pdf)
- H.S. Wiley, "Systems Biology - Beyond the Buzz." The Scientist (http://www.the-scientist.com/2006/6/1/52/1/). June 2006.]
- Nina Flanagan, "Systems Biology Alters Drug Development." (http://www.genengnews.com/articles/chitem.aspx?aid=2337) Genetic Engineering & Biotechnology News, January 2008

External links

- Systems Biology - BioChemWeb.org (http://www.biochemweb.org/systems.shtml)
- Systems Biology Portal (http://www.systems-biology.org/) - administered by the Systems Biology Institute
- Semantic Systems Biology (http://www.semantic-systems-biology.org)
- SystemsX.ch (http://www.systemsx.ch/) - The Swiss Initiative in Systems Biology
- Systems Biology at the Pacific Northwest National Laboratory (http://www.sysbio.org/)

Neural network



Simplified view of a feedforward artificial neural network

Traditionally, the term **neural network** had been used to refer to 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. Thus the term has two distinct usages:

1. Biological neural networks are made up of real biological neurons that are connected or functionally related in the peripheral nervous system or the central nervous system. In the field of neuroscience, they are often identified as groups of neurons that perform a specific physiological function in laboratory analysis.
2. Artificial neural networks are made up of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons). Artificial neural networks may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. The real, biological nervous system is highly complex and includes some features that may seem superfluous based on an understanding of artificial networks.

This article focuses on the relationship between the two concepts; for detailed coverage of the two different concepts refer to the separate articles: Biological neural network and Artificial neural network.

Overview

In general a biological neural network is composed of a group or groups of chemically connected or functionally associated neurons. A single neuron may be connected to many other neurons and the total number of neurons and connections in a network may be extensive. Connections, called synapses, are usually formed from axons to dendrites, though dendrodendritic microcircuits[1] and other connections are possible. Apart from the electrical signaling, there are other forms of signaling that arise from neurotransmitter diffusion, which have an effect on electrical signaling. As such, neural networks are extremely complex.

Artificial intelligence and cognitive modeling try to simulate some properties of neural networks. While similar in their techniques, the former has the aim of solving particular tasks, while the latter aims to build mathematical models of biological neural systems.

In the artificial intelligence field, artificial neural networks have been applied successfully to speech recognition, image analysis and adaptive control, in order to construct software agents (in computer and video games) or autonomous robots. Most of the currently employed artificial neural networks for artificial intelligence are based on statistical estimation, optimization and control theory.

The cognitive modelling field involves the physical or mathematical modeling of the behaviour of neural systems; ranging from the individual neural level (e.g. modelling the spike response curves of neurons to a stimulus), through the neural cluster level (e.g. modelling the release and effects of dopamine in the basal ganglia) to the complete organism (e.g. behavioural modelling of the organism's response to stimuli).

History of the neural network analogy

The concept of neural networks started in the late-1800s as an effort to describe how the human mind performed. These ideas started being applied to computational models with Turing's B-type machines and the perceptron.

In early 1950s Friedrich Hayek was one of the first to posit the idea of spontaneous order in the brain arising out of decentralized networks of simple units (neurons). In the late 1940s, Donald Hebb made one of the first hypotheses for a mechanism of neural plasticity (i.e. learning), Hebbian learning. Hebbian learning is considered to be a 'typical' unsupervised learning rule and it (and variants of it) was an early model for long term potentiation.

The Perceptron is essentially a linear classifier for classifying data $x \in R^n$ specified by parameters $w \in R^n, b \in R$ and an output function $f = w'x + b$. Its parameters are adapted with an ad-hoc rule similar to stochastic steepest gradient descent. Because the inner product is a linear operator in the input space, the Perceptron can only perfectly classify a set of data for which different classes are linearly separable in the input space, while it often fails completely for non-separable data. While the development of the algorithm initially generated some enthusiasm, partly because of its apparent relation to biological mechanisms, the later discovery of this inadequacy caused such models to be abandoned until the introduction of non-linear models into the field.
The Cognitron (1975) was an early multilayered neural network with a training algorithm. The actual structure of the network and the methods used to set the interconnection weights change from one neural strategy to another, each with its advantages and disadvantages. Networks can propagate information in one direction only, or they can bounce back and forth until self-activation at a node occurs and the network settles on a final state. The ability for bi-directional flow of inputs between neurons/nodes was produced with the Hopfield's network (1982), and specialization of these node layers for specific purposes was introduced through the first hybrid network.

The parallel distributed processing of the mid-1980s became popular under the name connectionism.

The rediscovery of the backpropagation algorithm was probably the main reason behind the repopularisation of neural networks after the publication of "Learning Internal Representations by Error Propagation" in 1986 (Though backpropagation itself dates from

1974). The original network utilised multiple layers of weight-sum units of the type $f = g(w'x + b)$, wh
g was a sigmoid function or logistic function such as used in logistic regression. Training was done by a form of stochastic steepest gradient descent. The employment of the chain rule of differentiation in deriving the appropriate parameter updates results in an algorithm that seems to 'backpropagate errors', hence the nomenclature. However it is essentially a form of gradient descent. Determining the optimal parameters in a model of this type is not trivial, and steepest gradient descent methods cannot be relied upon to give the solution without a good starting point. In recent times, networks with the same architecture as the backpropagation network are referred to as Multi-Layer Perceptrons. This name does not impose any limitations on the type of algorithm used for learning.
The backpropagation network generated much enthusiasm at the time and there was much controversy about whether such learning could be implemented in the brain or not, partly because a mechanism for reverse signalling was not obvious at the time, but most importantly because there was no plausible source for the 'teaching' or 'target' signal.

The brain, neural networks and computers

Neural networks, as used in artificial intelligence, have traditionally been viewed as simplified models of neural processing in the brain, even though the relation between this model and brain biological architecture is debated.

A subject of current research in theoretical neuroscience is the question surrounding the degree of complexity and the properties that individual neural elements should have to reproduce something resembling animal intelligence.

Historically, computers evolved from the von Neumann architecture, which is based on sequential processing and execution of explicit instructions. On the other hand, the origins of neural networks are based on efforts to model information processing in biological systems, which may rely largely on parallel processing as well as implicit instructions based on recognition of patterns of 'sensory' input from external sources. In other words, at its very heart a neural network is a complex statistical processor (as opposed to being tasked to sequentially process and execute).

Neural networks and artificial intelligence

An *artificial neural network* (ANN), also called a *simulated neural network* (SNN) or commonly just *neural network* (NN) is an interconnected group of artificial neurons that uses a mathematical or computational model for information processing based on a connectionistic 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.

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

Background

An artificial neural network involves a network of simple processing elements (artificial neurons) which can exhibit complex global behavior, determined by the connections between the processing elements and element parameters. Artificial neurons were first proposed in 1943 by Warren McCulloch, a neurophysiologist, and Walter Pitts, an MIT logician.[2] One classical type of artificial neural network is the Hopfield net.

In a neural network model simple nodes, which can be called variously "neurons", "neurodes", "Processing Elements" (PE) 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.

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.

The concept of a neural network appears to have first been proposed by Alan Turing in his 1948 paper "Intelligent Machinery".

Applications

The utility of artificial neural network models lies in the fact that they can be used to infer a function from observations and also to use it. 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 signal 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, etc.), sequence recognition (gesture, speech, handwritten text recognition), medical diagnosis, financial applications, data mining (or knowledge discovery in databases, "KDD"), visualization and e-mail spam filtering.

Neural network software

Main article: 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.

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.

Supervised learning

In supervised learning, we are given a set of example pairs $(x, y), x \in X, y \in Y$ and the aim is to find a function f in the allowed class of functions that matches the examples. In other words, we wish to *infer* how the mapping implied by the data and the cost function is related to the mismatch between our mapping and the data.

Unsupervised learning

In unsupervised learning we are given some data x, and a cost function which is to be minimized which can be any function of x and the network's output, f. The cost function is determined by the task formulation. Most applications fall within the domain of estimation problems such as statistical modeling, compression, filtering, blind source separation and clustering.

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 y_t and the environment generates an observation x_t and an instantaneous cost c_t, 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

There are many algorithms for training neural networks; most of them can be viewed as a straightforward application of optimization theory and statistical estimation. They include: Back propagation by gradient descent, Rprop, BFGS, CG etc.

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

Recent developments in this field also saw the use of particle swarm optimization and other swarm intelligence techniques used in the training of neural networks.

Neural networks and neuroscience

Theoretical and computational neuroscience is the field concerned with the theoretical analysis and computational modeling of biological neural systems. Since neural systems are intimately related to cognitive processes and behaviour, the field is closely related to cognitive and behavioural modeling.

The aim of the field is to create models of biological neural systems in order to understand how biological systems work. To gain this understanding, neuroscientists strive to make a link between observed biological processes (data), biologically plausible mechanisms for neural processing and learning (biological neural network models) and theory (statistical learning theory and information theory).

Types of models

Many models are used in the field, each defined at a different level of abstraction and trying to model different aspects of neural systems. They range from models of the short-term behaviour of individual neurons, through models of how the dynamics of neural circuitry arise from interactions between individual neurons, to models of how behaviour can arise from abstract neural modules that represent complete subsystems. These include models of the long-term and short-term plasticity of neural systems and its relation to learning and memory, from the individual neuron to the system level.

Current research

While initially research had been concerned mostly with the electrical characteristics of neurons, a particularly important part of the investigation in recent years has been the exploration of the role of neuromodulators such as dopamine, acetylcholine, and serotonin on behaviour and learning.

Biophysical models, such as BCM theory, have been important in understanding mechanisms for synaptic plasticity, and have had applications in both computer science and neuroscience. Research is ongoing in understanding the computational algorithms used in the brain, with some recent biological evidence for radial basis networks and neural backpropagation as mechanisms for processing data.

Criticism

A common criticism of neural networks, particularly in robotics, is that they require a large diversity of training for real-world operation. Dean Pomerleau, in his research presented in the paper "Knowledge-based Training of Artificial Neural Networks for Autonomous Robot Driving," uses a neural network to train a robotic vehicle to drive on multiple types of roads (single lane, multi-lane, dirt, etc.). A large amount of his research is devoted to (1) extrapolating multiple training scenarios from a single training experience, and (2) preserving past training diversity so that the system does not become overtrained (if, for example, it is presented with a series of right turns – it should not learn to always turn right). These issues are common in neural networks that must decide from amongst a wide variety of responses.

A. K. Dewdney, a former *Scientific American* columnist, wrote in 1997, "Although neural nets do solve a few toy problems, their powers of computation are so limited that I am surprised anyone takes them seriously as a general problem-solving tool." (Dewdney, p.82)

Arguments for Dewdney's position are that to implement large and effective software neural networks, much processing and storage resources need to be committed. While the brain has hardware tailored to the task of processing signals through a graph of neurons, simulating even a most simplified form on Von Neuman technology may compel a NN designer to fill many millions of database rows for its connections - which can lead to abusive RAM and HD necessities. Furthermore, the designer of NN systems will often need to simulate the transmission of signals through many of these connections and their associated neurons - which must often be matched with incredible amounts of CPU processing power and time. While neural networks often yield *effective* programs, they too often do so at the cost of time and money *efficiency*.

Arguments against Dewdney's position are that neural nets have been successfully used to solve many complex and diverse tasks, ranging from autonomously flying aircraft[3] to detecting credit card fraud[4].

Technology writer Roger Bridgman commented on Dewdney's statements about neural nets:

> Neural networks, for instance, are in the dock not only because they have been hyped to high heaven, (what hasn't?) but also because you could create a successful net without understanding how it worked: the bunch of numbers that captures its behaviour would in all probability be "an opaque, unreadable table...valueless as a scientific resource".
>
> In spite of his emphatic declaration that science is not technology, Dewdney seems here to pillory neural nets as bad science when most of those devising them are just trying to be good engineers. An unreadable table that a useful machine could read would still be well worth having.[5]

Some other criticisms came from believers of hybrid models (combining neural networks and symbolic approaches). They advocate the intermix of these two approaches and believe that hybrid models can better capture the mechanisms of the human mind (Sun and Bookman 1994).

See also

- ADALINE
- Artificial neural network
- Biological cybernetics
- Biologically-inspired computing
- Cognitive architecture
- Memristor
- Neural network software
- Neuro-fuzzy
- Parallel distributed processing
- Predictive analytics
- Radial basis function network
- Simulated reality
- Support vector machine
- Tensor product network
- 20Q is a neural network implementation of the 20 questions game
- Cultured neuronal networks

- Neuroscience
- Cognitive science
- Recurrent neural networks

References

[1] Arbib, p.666
[2] http://palisade.com/neuraltools/neural_networks.asp
[3] http://www.nasa.gov/centers/dryden/news/NewsReleases/2003/03-49.html
[4] http://www.visa.ca/en/about/visabenefits/innovation.cfm
[5] Roger Bridgman's defence of neural networks (http://members.fortunecity.com/templarseries/popper.html)

Further reading

- Arbib, Michael A. (Ed.) (1995). *The Handbook of Brain Theory and Neural Networks*.
- Alspector, U.S. Patent 4874963 (http://www.google.com/patents?vid=4874963) "*Neuromorphic learning networks*". October 17, 1989.
- Agre, Philip E. (1997). *Computation and Human Experience*. Cambridge University Press. ISBN 0-521-38603-9., p. 80
- Yaneer Bar-Yam (2003). *Dynamics of Complex Systems, Chapter 2 (http://necsi.org/publications/dcs/Bar-YamChap2.pdf)*.
- Yaneer Bar-Yam (2003). *Dynamics of Complex Systems, Chapter 3 (http://necsi.org/publications/dcs/Bar-YamChap3.pdf)*.
- Yaneer Bar-Yam (2005). *Making Things Work (http://necsi.org/publications/mtw/)*. See chapter 3.
- Bertsekas, Dimitri P. (1999). *Nonlinear Programming*.
- Bertsekas, Dimitri P. & Tsitsiklis, John N. (1996). *Neuro-dynamic Programming*.
- Bhadeshia H. K. D. H. (1992). " Neural Networks in Materials Science (http://www.msm.cam.ac.uk/phase-trans/abstracts/neural.review.pdf)". *ISIJ International* **39**: 966–979. doi: 10.2355/isijinternational.39.966 (http://dx.doi.org/10.2355/isijinternational.39.966).
- Boyd, Stephen & Vandenberghe, Lieven (2004). *Convex Optimization (http://www.stanford.edu/~boyd/cvxbook/)*.
- Dewdney, A. K. (1997). *Yes, We Have No Neutrons: An Eye-Opening Tour through the Twists and Turns of Bad Science*. Wiley, 192 pp. See chapter 5.
- Egmont-Petersen, M., de Ridder, D., Handels, H. (2002). "Image processing with neural networks - a review". *Pattern Recognition* **35** (10): 2279–2301. doi: 10.1016/S0031-3203(01)00178-9 (http://dx.doi.org/10.1016/S0031-3203(01)00178-9).
- Fukushima, K. (1975). "Cognitron: A Self-Organizing Multilayered Neural Network". *Biological Cybernetics* **20**: 121–136. doi: 10.1007/BF00342633 (http://dx.doi.org/10.1007/BF00342633).
- Frank, Michael J. (2005). "Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Non-medicated Parkinsonism". *Journal of Cognitive Neuroscience* **17**: 51–72. doi: 10.1162/0898929052880093 (http://dx.doi.org/10.1162/0898929052880093).
- Gardner, E.J., & Derrida, B. (1988). "Optimal storage properties of neural network models". *Journal of Physics a* **21**: 271–284. doi: 10.1088/0305-4470/21/1/031 (http://dx.doi.org/10.1088/0305-4470/21/1/031).

- Krauth, W., & Mezard, M. (1989). "Storage capacity of memory with binary couplings". *Journal de Physique* **50**: 3057–3066. doi: 10.1051/jphys:0198900500200305700 (http://dx.doi.org/10.1051/jphys:0198900500200305700).
- Maass, W., & Markram, H. (2002). " On the computational power of recurrent circuits of spiking neurons (http://www.igi.tugraz.at/maass/publications.html)". *Journal of Computer and System Sciences* **69(4)**: 593–616.
- MacKay, David (2003). *Information Theory, Inference, and Learning Algorithms (http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html)*.
- Mandic, D. & Chambers, J. (2001). *Recurrent Neural Networks for Prediction: Architectures, Learning algorithms and Stability*. Wiley.
- Minsky, M. & Papert, S. (1969). *An Introduction to Computational Geometry*. MIT Press.
- Muller, P. & Insua, D.R. (1995). "Issues in Bayesian Analysis of Neural Network Models". *Neural Computation* **10**: 571–592.
- Reilly, D.L., Cooper, L.N. & Elbaum, C. (1982). "A Neural Model for Category Learning". *Biological Cybernetics* **45**: 35–41. doi: 10.1007/BF00387211 (http://dx.doi.org/10.1007/BF00387211).
- Rosenblatt, F. (1962). *Principles of Neurodynamics*. Spartan Books.
- Sun, R. & Bookman,L. (eds.) (1994.). *Computational Architectures Integrating Neural and Symbolic Processes.*. Kluwer Academic Publishers, Needham, MA..
- Sutton, Richard S. & Barto, Andrew G. (1998). *Reinforcement Learning : An introduction (http://www.cs.ualberta.ca/~sutton/book/the-book.html)*.
- Van den Bergh, F. Engelbrecht, AP. *Cooperative Learning in Neural Networks using Particle Swarm Optimizers*. CIRG 2000.
- Wilkes, A.L. & Wade, N.J. (1997). "Bain on Neural Networks". *Brain and Cognition* **33**: 295–305. doi: 10.1006/brcg.1997.0869 (http://dx.doi.org/10.1006/brcg.1997.0869).
- Wasserman, P.D. (1989). *Neural computing theory and practice*. Van Nostrand Reinhold.
- Jeffrey T. Spooner, Manfredi Maggiore, Raul Ord onez, and Kevin M. Passino, Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques, John Wiley and Sons, NY, 2002.
- http://www.cs.stir.ac.uk/courses/31YF/Notes/Notes_PL.html
- http://www.shef.ac.uk/psychology/gurney/notes/l1/section3_3.html
- Peter Dayan, L.F. Abbott. *Theoretical Neuroscience*. MIT Press.
- Wulfram Gerstner, Werner Kistler. *Spiking Neuron Models:Single Neurons, Populations, Plasticity*. Cambridge University Press.

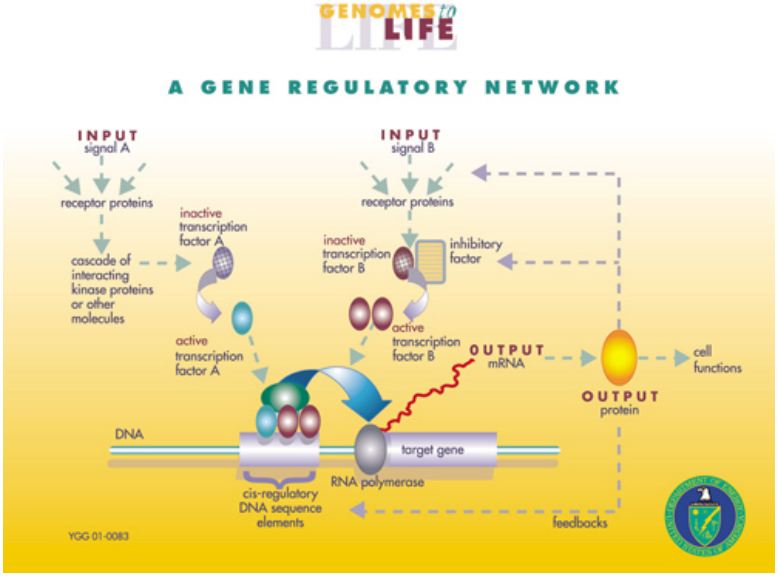
External links

- International Neural Network Society (INNS) (http://www.inns.org)
- European Neural Network Society (ENNS) (http://www.e-nns.org)
- Japanese Neural Network Society (JNNS) (http://www.jnns.org)
- IEEE Computational Intelligence Society (IEEE CIS) (http://www.ieee-cis.org)
- LearnArtificialNeuralNetworks (http://www.learnartificialneuralnetworks.com/robotcontrol.html) - Robot control and neural networks
- Review of Neural Networks in Materials Science (http://www.msm.cam.ac.uk/phase-trans/abstracts/neural.review.html)
- Artificial Neural Networks Tutorial in three languages (Univ. Politécnica de Madrid) (http://www.gc.ssr.upm.es/inves/neural/ann1/anntutorial.html)

- Introduction to Neural Networks and Knowledge Modeling (http://www.makhfi.com/tutorial/introduction.htm)
- Introduction to Artificial Neural Networks (http://www.willamette.edu/~gorr/classes/cs449/intro.html)
- In Situ Adaptive Tabulation: (http://www.hedengren.net/research/isat.htm) - A neural network alternative.
- Another introduction to ANN (http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html)
- Prediction with neural networks (http://www.obitko.com/tutorials/neural-network-prediction/) - includes Java applet for online experimenting with prediction of a function
- Next Generation of Neural Networks (http://pl.youtube.com/watch?v=AyzOUbkUf3M) - Google Tech Talks
- Perceptual Learning (http://pages.sbcglobal.net/louis.savain/AI/perceptual_network.htm) - Artificial Perceptual Neural Network used for machine learning to play Chess
- European Centre for Soft Computing (http://www.softcomputing.es/en/home.php)

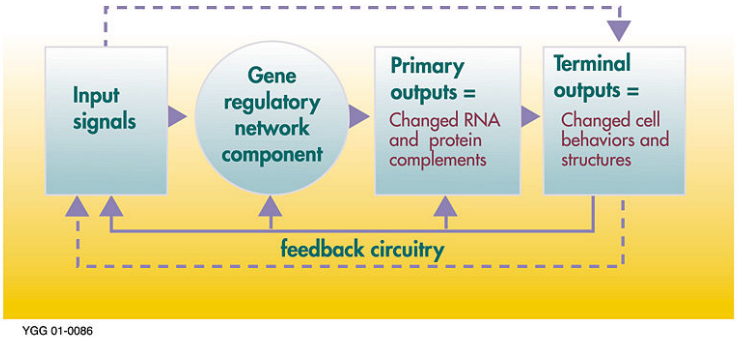
Gene regulatory network

A **gene regulatory network** or **genetic regulatory network (GRN)** is a collection of DNA segments in a cell which interact with each other (indirectly through their RNA and protein expression products) and with other substances in the cell, thereby governing the rates at which genes in the network are transcribed into mRNA. In general, each mRNA molecule goes on to make a specific protein (or set of proteins). In some cases this protein will be structural, and will accumulate at the cell-wall or within the cell to give it particular structural properties. In other cases the protein will be an enzyme; a micro-machine that catalyses a certain reaction, such as the breakdown of a food source or



Structure of a Gene Regulatory Network.

toxin. Some proteins though serve only to activate other genes, and these are the transcription factors that are the main players in regulatory networks or cascades. By binding to the promoter region at the start of other genes they turn them on, initiating the production of another protein, and so on. Some transcription factors are inhibitory.



Control process of a Gene Regulatory Network.

In single-celled organisms regulatory networks respond to the external environment, optimising the cell at a given time for survival in this environment. Thus a yeast cell, finding itself in a sugar solution, will turn on genes to make enzymes that process the sugar to alcohol.[1] This process, which we associate with wine-making, is how the yeast cell makes its living, gaining energy to multiply, which under normal circumstances would enhance its survival prospects.

In multicellular animals the same principle has been put in the service of gene cascades that control body-shape.[2] Each time a cell divides, two cells result which, although they contain the same genome in full, can differ in which genes are turned on and making proteins. Sometimes a 'self-sustaining feedback loop' ensures that a cell maintains its identity and passes it on. Less understood is the mechanism of epigenetics by which chromatin modification may provide cellular memory by blocking or allowing transcription. A major feature of multicellular animals is the use of morphogen gradients, which in effect provide a positioning system that tells a cell where in the body it is, and hence what sort of cell to become. A gene that is turned on in one cell may make a product that leaves the cell and diffuses through adjacent cells, entering them and turning on genes only when it is present above a certain threshold level. These cells are thus induced into a new fate, and may even generate other morphogens that signal back to the original cell. Over longer distances morphogens may use the active process of signal transduction. Such signalling controls embryogenesis, the building of a body plan from scratch through a series of sequential steps. They also control maintain adult bodies through feedback processes, and the loss of such feedback because of a mutation can be responsible for the cell proliferation that is seen in cancer. In parallel with this process of building structure, the gene cascade turns on genes that make structural proteins that give each cell the physical properties it needs.

Overview

At one level, biological cells can be thought of as "partially-mixed bags" of biological chemicals – in the discussion of gene regulatory networks, these chemicals are mostly the mRNAs and proteins that arise from gene expression. These mRNA and proteins interact with each other with various degrees of specificity. Some diffuse around the cell. Others are bound to cell membranes, interacting with molecules in the environment. Still others pass through cell membranes and mediate long range signals to other cells in a multi-cellular

organism. These molecules and their interactions comprise a *gene regulatory network*. A typical gene regulatory network looks something like this:

The nodes of this network are proteins, their corresponding mRNAs, and protein/protein complexes. Nodes that are depicted as lying along vertical lines are associated with the cell/environment interfaces, while the others are free-floating and diffusible. Implied are genes, the DNA sequences which are transcribed into the mRNAs that translate into proteins. Edges between nodes represent individual molecular reactions, the protein/protein and protein/mRNA interactions through which the products of one gene affect those of another, though the lack of experimentally obtained information often implies that some reactions are not modeled at such a fine level of detail. These interactions can be inductive (the arrowheads), with an increase in the concentration of one leading to an increase in the other, or inhibitory (the filled circles), with an increase in one leading to a decrease in the other. A series of edges indicates a chain of such dependences, with cycles corresponding to feedback loops. The network structure is an abstraction of the system's chemical dynamics, describing the manifold ways in which one substance affects all the others to which it is connected. In practice, such GRNs are inferred from the biological literature on a given system and represent a distillation of the collective knowledge about a set of related biochemical reactions.

Genes can be viewed as nodes in the network, with input being proteins such as transcription factors, and outputs being the level of gene expression. The node itself can also be viewed as a function which can be obtained by combining basic functions upon the inputs (in the Boolean network described below these are Boolean functions, typically AND, OR, and NOT). These functions have been interpreted as performing a kind of information processing within the cell, which determines cellular behavior. The basic drivers within cells are concentrations of some proteins, which determine both spatial (location within the cell or tissue) and temporal (cell cycle or developmental stage) coordinates of the cell, as a kind of "cellular memory". The gene networks are only beginning to be understood, and it is a next step for biology to attempt to deduce the functions for each gene "node", to help understand the behavior of the system in increasing levels of complexity, from gene to signaling pathway, cell or tissue level (see systems biology).

Mathematical models of GRNs have been developed to capture the behavior of the system being modeled, and in some cases generate predictions corresponding with experimental observations. In some other cases, models have proven to make accurate novel predictions, which can be tested experimentally, thus suggesting new approaches to explore in an experiment that sometimes wouldn't be considered in the design of the protocol of an experimental laboratory. The most common modeling technique involves the use of coupled ordinary differential equations (ODEs). Several other promising modeling techniques have been used, including Boolean networks, Petri nets, Bayesian networks, graphical Gaussian models, Stochastic, and Process Calculi. Conversely, techniques have been proposed for generating models of GRNs that best explain a set of time series observations.

Modelling

Coupled ODEs

It is common to model such a network with a set of coupled ordinary differential equations (ODEs) or stochastic ODEs, describing the reaction kinetics of the constituent parts. Suppose that our regulatory network has N nodes, and let $S_1(t), S_2(t), \ldots, S_N(t)$ represent the concentrations of the N corresponding substances at time t. Then the temporal evolution of the system can be described approximately by

$$\frac{dS_j}{dt} = f_j\left(S_1, S_2, \ldots, S_N\right)$$

where the functions f_j express the dependence of S_j on the concentrations of other substances present in the cell. The functions f_j are ultimately derived from basic principles of chemical kinetics or simple expressions derived from these e.g. Michaelis-Menten enzymatic kinetics. Hence, the functional forms of the f_j are usually chosen as low-order polynomials or Hill functions that serve as an ansatz for the real molecular dynamics. Such models are then studied using the mathematics of nonlinear dynamics. System-specific information, like reaction rate constants and sensitivities, are encoded as constant parameters.
By solving for the fixed point of the system:

$$\frac{dS_j}{dt} = 0$$

for all j, one obtains (possibly several) concentration profiles of proteins and mRNAs that are theoretically sustainable (though not necessarily stable). Steady states of kinetic equations thus correspond to potential cell types, and oscillatory solutions to the above equation to naturally cyclic cell types. Mathematical stability of these attractors can usually be characterized by the sign of higher derivatives at critical points, and then correspond to biochemical stability of the concentration profile. Critical points and bifurcations in the equations correspond to critical cell states in which small state or parameter perturbations could switch the system between one of several stable differentiation fates. Trajectories correspond to the unfolding of biological pathways and transients of the equations to short-term biological events. For a more mathematical discussion, see the articles on nonlinearity, dynamical systems, bifurcation theory, and chaos theory.

Boolean network

The following example illustrates how a Boolean network can model a GRN together with its gene products (the outputs) and the substances from the environment that affect it (the inputs). Stuart Kauffman was amongst the first biologists to use the metaphor of Boolean networks to model genetic regulatory networks.[3]

1. Each gene, each input, and each output is represented by a node in a directed graph in which there is an arrow from one node to another if and only if there is a causal link between the two nodes.
2. Each node in the graph can be in one of two states: on or off.
3. For a gene, "on" corresponds to the gene being expressed; for inputs and outputs, "on" corresponds to the substance being present.
4. Time is viewed as proceeding in discrete steps. At each step, the new state of a node is a Boolean function of the prior states of the nodes with arrows pointing towards it.

The validity of the model can be tested by comparing simulation results with time series observations.

Continuous networks

Continuous network models of GRNs are an extension of the boolean networks described above. Nodes still represent genes and connections between them regulatory influences on gene expression. Genes in biological systems display a continuous range of activity levels and it has been argued that using a continuous representation captures several properties of gene regulatory networks not present in the Boolean model.[4] Formally most of these approaches are similar to an artificial neural network, as inputs to a node are summed up and the result serves as input to a sigmoid function, e.g.,[5] but proteins do often control gene expression in a synergistic, i.e. non-linear, way.[6] However there is now a continuous network model[7] that allows grouping of inputs to a node thus realizing another level of regulation. This model is formally closer to a higher order recurrent neural network. The same model has also been used to mimic the evolution of cellular differentiation[8] and even multicellular morphogenesis.[9]

Stochastic gene networks

Recent experimental results[10] [11] have demonstrated that gene expression is a stochastic process. Thus, many authors are now using the stochastic formalism, after the first work by.[12] Works on single gene expression[13] and small synthetic genetic networks,[14] [15] such as the genetic toggle switch of Tim Gardner and Jim Collins, provided additional experimental data on the phenotypic variability and the stochastic nature of gene expression. The first versions of stochastic models of gene expression involved only instantaneous reactions and were driven by the Gillespie algorithm.[16]

Since some processes, such as gene transcription, involve many reactions and could not be correctly modeled as an instantaneous reaction in a single step, it was proposed to model these reactions as single step multiple delayed reactions in order to account for the time it takes for the entire process to be complete.[17]

From here, a set of reactions were proposed[18] that allow generating GRNs. These are then simulated using a modified version of the Gillespie algorithm, that can simulate multiple time delayed reactions (chemical reactions where each of the products is provided a time delay that determines when will it be released in the system as a "finished product").

For example, basic transcription of a gene can be represented by the following single-step reaction (RNAP is the RNA polymerase, RBS is the RNA ribosome binding site, and Pro_i is the promoter region of gene *i*):

$$\mathrm{RNAP} + \mathrm{Pro}_i \xrightarrow{k_{i,bas}} \mathrm{Pro}_i(\tau_i^1) + \mathrm{RBS}_i(\tau_i^1) + \mathrm{RNAP}(\tau_i^2)$$

A recent work proposed a simulator (SGNSim, *Stochastic Gene Networks Simulator*),[19] that can model GRNs where transcription and translation are modeled as multiple time delayed events and its dynamics is driven by a stochastic simulation algorithm (SSA) able to deal with multiple time delayed events. The time delays can be drawn from several distributions and the reaction rates from complex functions or from physical parameters. SGNSim can generate ensembles of GRNs within a set of user-defined parameters, such as topology. It can also be used to model specific GRNs and systems of chemical reactions. Genetic perturbations such as gene deletions, gene over-expression, insertions, frame shift

mutations can also be modeled as well.

The GRN is created from a graph with the desired topology, imposing in-degree and out-degree distributions. Gene promoter activities are affected by other genes expression products that act as inputs, in the form of monomers or combined into multimers and set as direct or indirect. Next, each direct input is assigned to an operator site and different transcription factors can be allowed, or not, to compete for the same operator site, while indirect inputs are given a target. Finally, a function is assigned to each gene, defining the gene's response to a combination of transcription factors (promoter state). The transfer functions (that is, how genes respond to a combination of inputs) can be assigned to each combination of promoter states as desired.

In other recent work, multiscale models of gene regulatory networks have been developed that focus on synthetic biology applications. Simulations have been used that model all biomolecular interactions in transcription, translation, regulation, and induction of gene regulatory networks, guiding the design of synthetic systems.[20]

Network connectivity

Empirical data indicate that biological gene networks are sparsely connected, and that the average number of upstream-regulators per gene is less than two.[21] Theoretical results show that selection for robust gene networks will favor minimally complex, more sparsely connected, networks.[21] These results suggest that a sparse, minimally connected, genetic architecture may be a fundamental design constraint shaping the evolution of gene network complexity.

See also

- Operon
- Systems biology
- Synexpression
- Cis-regulatory module
- Body plan
- Morphogen

References

[1] http://web.wi.mit.edu/young/regulator_network/
[2] http://www.pnas.org/cgi/content/full/102/14/4935
[3] Kauffman, Stuart (1993). *The origins of Order*.
[4] Vohradsky, J. (2001). Neural model of the genetic network. The Journal of Biological Chemistry, 276, 36168–36173.
[5] Geard, N. and Wiles, J. A Gene Network Model for Developing Cell Lineages. In Artificial Life, 11 (3): 249–268, 2005.
[6] Schilstra, M. J. and Bolouri, H. The Logic of Gene Regulation., http://strc.herts.ac.uk/bio/maria/NetBuilder/Theory/NetBuilderModelling.htm
[7] Knabe, J. F., Nehaniv, C. L., Schilstra, M. J. and Quick, T. Evolving Biological Clocks using Genetic Regulatory Networks. In Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems, pages 15–21, MIT Press, 2006.
[8] Knabe, J. F., Nehaniv, C. L. and Schilstra, M. J. Evolutionary Robustness of Differentiation in Genetic Regulatory Networks. In Proceedings of the 7th German Workshop on Artificial Life 2006 (GWAL-7), pages 75–84, Akademische Verlagsgesellschaft Aka, Berlin, 2006.
[9] Knabe, J. F., Schilstra, M. J. and Nehaniv, C. L. Evolution and Morphogenesis of Differentiated Multicellular Organisms: Autonomously Generated Diffusion Gradients for Positional Information. In Artificial Life XI:

Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems, MIT Press, 2008.
[10] Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. 2002. Stochastic gene expression in a single cell. Science 297: 1183–1186
[11] Blake, W.J., Kaern, M., Cantor, C.R., and Collins, J.J. 2003. Noise in eukaryotic gene expression. (http://www.bu.edu/abl/publications.html) Nature 422: 633–637
[12] Arkin, A. and McAdams, H.H. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics 149: 1633–1648.
[13] Raser, J.M., and O'Shea, E.K., (2005) Noise in gene expression: origins, consequences, and control, Science, 309, 2010–2013
[14] Elowitz, M. B., and Leibler, S., (2000) A synthetic oscillatory network of transcriptional regulators., Nature, 403, 335–338
[15] Gardner, T. S., Cantor, C. R., and Collins., J. J., (2000) Construction of a genetic toggle switch in Escherichia coli., Nature, 403, 339–342
[16] Gillespie, D.T., A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, 1976, J. Comput. Phys., 22, 403–434.
[17] Roussel, M.R., and Zhu, R., Validation of an algorithm for delay stochastic simulation of transcription and translation in prokaryotic gene expression, 2006, Phys. Biol. 3, 274–284
[18] Ribeiro, Andre S., Zhu, R., Kauffman, S.A. (2006). "A General Modeling Strategy for Gene Regulatory Networks with Stochastic Dynamics", Journal of Computational Biology, 13(9), 1630–1639.
[19] Andre S. Ribeiro and Jason Lloyd-Price, (2007) "SGN Sim, a Stochastic Genetic Networks Simulator", Bioinformatics, 23(6):777–779. doi:10.1093/bioinformatics/btm004., doi:10.1093/bioinformatics/btm004.
[20] Y. N. Kaznessis, (2007) "Models for Synthetic Biology", BMC Systems Biology, 2007, 1:47 doi:10.1186/1752-0509-1-47 (http://www.biomedcentral.com/1752-0509/1/47).
[21] Leclerc R. (August 2008). " Survival of the sparsest: robust gene networks are parsimonious (http://www.nature.com/msb/journal/v4/n1/full/msb200852.html)". *Mol Syst Biol.* **4** (213).

- James M. Bower, Hamid Bolouri (editors), (2001) *Computational Modeling of Genetic and Biochemical Networks* Computational Molecular Biology Series, MIT Press, ISBN 0-262-02481-0
- L. Franke, H. van Bakel, L. Fokkens, E. de Jong, M. Egmont-Petersen, C. Wijmenga, (2006) Reconstruction of a probabilistic human gene network, with an application for prioritizing positional candidate genes, Amer. J. of Human Genetics, 78(6), 1011–25. Human gene network (http://www.genenetwork.nl), Prioritizer software application (http://www.prioritizer.nl).
- S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets", J. Theoret. Biol (1969) 22, 434–467

External links

- Gene Regulatory Networks (http://www.doegenomestolife.org/science/generegulatorynetwork.shtml) — Short introduction
- BIB: Yeast Biological Interaction Browser (http://sergi5.com/bio)
- Graphical Gaussian models for genome data (http://strimmerlab.org/notes/ggm.html) — Inference of gene association networks with GGMs
- A bibliography on learning causal networks of gene interactions (http://www.molgen.mpg.de/~markowet/docs/network-bib.pdf) - regularly updated, contains hundreds of links to papers from bioinformatics, statistics, machine learning.
- http://mips.gsf.de/proj/biorel/ BIOREL is a web-based resource for quantitative estimation of the gene network bias in relation to available database information about gene activity/function/properties/associations/interactio.
- Evolving Biological Clocks using Genetic Regulatory Networks (http://panmental.de/GRNclocks) - Information page with model source code and Java applet.

- Engineered Gene Networks (http://www.bu.edu/abl)
- Tutorial: Genetic Algorithms and their Application to the Artificial Evolution of Genetic Regulatory Networks (http://panmental.de/ICSBtut/)

Genomics

Genomics is the study of the genomes of organisms. The field includes intensive efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping efforts. The field also includes studies of intragenomic phenomena such as heterosis, epistasis, pleiotropy and other interactions between loci and alleles within the genome. In contrast, the investigation of the roles and functions of single genes is a primary focus of molecular biology and is a common topic of modern medical and biological research. Research of single genes does not fall into the definition of genomics unless the aim of this genetic, pathway, and functional information analysis is to elucidate its effect on, place in, and response to the entire genome's networks.

For the United States Environmental Protection Agency, "the term "genomics" encompasses a broader scope of scientific inquiry associated technologies than when genomics was initially considered. A genome is the sum total of all an individual organism's genes. Thus, genomics is the study of all the genes of a cell, or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) levels."[1]

History

Genomics was established by Fred Sanger when he first sequenced the complete genomes of a virus and a mitochondrion. His group established techniques of sequencing, genome mapping, data storage, and bioinformatic analyses in the 1970-1980s. A major branch of genomics is still concerned with sequencing the genomes of various organisms, but the knowledge of full genomes has created the possibility for the field of functional genomics, mainly concerned with patterns of gene expression during various conditions. The most important tools here are microarrays and bioinformatics. Study of the full set of proteins in a cell type or tissue, and the changes during various conditions, is called proteomics. A related concept is materiomics, which is defined as the study of the material properties of biological materials (e.g. hierarchical protein structures and materials, mineralized biological tissues, etc.) and their effect on the macroscopic function and failure in their biological context, linking processes, structure and properties at multiple scales through a materials science approach. The actual term 'genomics' is thought to have been coined by Dr. Tom Roderick, a geneticist at the Jackson Laboratory (Bar Harbor, ME) over beer at a meeting held in Maryland on the mapping of the human genome in 1986.

In 1972, Walter Fiers and his team at the Laboratory of Molecular Biology of the University of Ghent (Ghent, Belgium) were the first to determine the sequence of a gene: the gene for Bacteriophage MS2 coat protein.[2] In 1976, the team determined the complete nucleotide-sequence of bacteriophage MS2-RNA.[3] The first DNA-based genome to be sequenced in its entirety was that of bacteriophage Φ-X174; (5,368 bp), sequenced by Frederick Sanger in 1977.[4]

The first free-living organism to be sequenced was that of *Haemophilus influenzae* (1.8 Mb) in 1995, and since then genomes are being sequenced at a rapid pace. A rough draft of the

human genome was completed by the Human Genome Project in early 2001, creating much fanfare.

As of September 2007, the complete sequence was known of about 1879 viruses [5] , 577 bacterial species and roughly 23 eukaryote organisms, of which about half are fungi. [6] Most of the bacteria whose genomes have been completely sequenced are problematic disease-causing agents, such as *Haemophilus influenzae*. Of the other sequenced species, most were chosen because they were well-studied model organisms or promised to become good models. Yeast (*Saccharomyces cerevisiae*) has long been an important model organism for the eukaryotic cell, while the fruit fly *Drosophila melanogaster* has been a very important tool (notably in early pre-molecular genetics). The worm *Caenorhabditis elegans* is an often used simple model for multicellular organisms. The zebrafish *Brachydanio rerio* is used for many developmental studies on the molecular level and the flower *Arabidopsis thaliana* is a model organism for flowering plants. The Japanese pufferfish (*Takifugu rubripes*) and the spotted green pufferfish (*Tetraodon nigroviridis*) are interesting because of their small and compact genomes, containing very little non-coding DNA compared to most species. [7] [8] The mammals dog (*Canis familiaris*), [9] brown rat (*Rattus norvegicus*), mouse (*Mus musculus*), and chimpanzee (*Pan troglodytes*) are all important model animals in medical research.

Bacteriophage genomics

Bacteriophages have played and continue to play a key role in bacterial genetics and molecular biology. Historically, they were used to define gene structure and gene regulation. Also the first genome to be sequenced was a bacteriophage. However, bacteriophage research did not lead the genomics revolution, which is clearly dominated by bacterial genomics. Only very recently has the study of bacteriophage genomes become prominent, thereby enabling researchers to understand the mechanisms underlying phage evolution. Bacteriophage genome sequences can be obtained through direct sequencing of isolated bacteriophages, but can also be derived as part of microbial genomes. Analysis of bacterial genomes has shown that a substantial amount of microbial DNA consists of prophage sequences and prophage-like elements. A detailed database mining of these sequences offers insights into the role of prophages in shaping the bacterial genome.[10]

Cyanobacteria genomics

At present there are 24 cyanobacteria for which a total genome sequence is available. 15 of these cyanobacteria come from the marine environment. These are six *Prochlorococcus* strains, seven marine *Synechococcus* strains, *Trichodesmium erythraeum* IMS101 and *Crocosphaera watsonii* WH8501. Several studies have demonstrated how these sequences could be used very successfully to infer important ecological and physiological characteristics of marine cyanobacteria. However, there are many more genome projects currently in progress, amongst those there are further *Prochlorococcus* and marine *Synechococcus* isolates, *Acaryochloris* and *Prochloron*, the N_2-fixing filamentous cyanobacteria *Nodularia spumigena*, *Lyngbya aestuarii* and *Lyngbya majuscula*, as well as bacteriophages infecting marine cyanobaceria. Thus, the growing body of genome information can also be tapped in a more general way to address global problems by applying a comparative approach. Some new and exciting examples of progress in this field are the identification of genes for regulatory RNAs, insights into the evolutionary origin of

photosynthesis, or estimation of the contribution of horizontal gene transfer to the genomes that have been analyzed.[11]

See also

- Full Genome Sequencing
- Computational genomics
- Nitrogenomics
- Metagenomics
- Predictive Medicine
- Personal genomics

References

[1] EPA Interim Genomics Policy (http://epa.gov/osa/spc/pdfs/genomics.pdf)
[2] Min Jou W, Haegeman G, Ysebaert M, Fiers W (1972). "Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein". *Nature* **237** (5350): 82–88. PMID 4555447.
[3] Fiers W, Contreras R, Duerinck F, Haegeman G, Iserentant D, Merregaert J, Min Jou W, Molemans F, Raeymaekers A, Van den Berghe A, Volckaert G, Ysebaert M (1976). "Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene". *Nature* **260** (5551): 500–507. PMID 1264203.
[4] Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, Hutchison CA, Slocombe PM, Smith M (1977). "Nucleotide sequence of bacteriophage phi X174 DNA". *Nature* **265** (5596): 687–695. PMID 870828.
[5] *The Viral Genomes Resource*, NCBI Friday, 14 September 2007 (http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/virostat.html)
[6] *Genome Project Statistic*, NCBI Friday, 14 September 2007 (http://www.ncbi.nlm.nih.gov/genomes/static/gpstat.html)
[7] BBC article *Human gene number slashed* from Wednesday, 20 October 2004 (http://news.bbc.co.uk/1/hi/sci/tech/3760766.stm)
[8] CBSE News, Thursday, 16 October 2003 (http://www.cbse.ucsc.edu/news/2003/10/16/pufferfish_fruitfly/index.shtml)
[9] NHGRI, pressrelease of the publishing of the dog genome (http://www.genome.gov/12511476)
[10] McGrath S and van Sinderen D, ed (2007). *Bacteriophage: Genetics and Molecular Biology (http://www.horizonpress.com/phage)* (1st ed.). Caister Academic Press. ISBN 978-1-904455-14-1. .
[11] Herrero A and Flores E, ed (2008). *The Cyanobacteria: Molecular Biology, Genomics and Evolution (http://www.horizonpress.com/cyan)* (1st ed.). Caister Academic Press. ISBN 978-1-904455-15-8. .

External links

- Genomics Directory (http://www.genomicsdirectory.com): A one-stop biotechnology resource center for bioentrepreneurs, scientists, and students
- Annual Review of Genomics and Human Genetics (http://arjournals.annualreviews.org/loi/genom/)
- BMC Genomics (http://www.biomedcentral.com/bmcgenomics/): A BMC journal on Genomics
- Genomics (http://www.genomics.co.uk/companylist.php): UK companies and laboratories* Genomics journal (http://www.elsevier.com/wps/find/journaldescription.cws_home/622838/description#description)
- Genomics.org (http://genomics.org): An openfree wiki based Genomics portal
- NHGRI (http://www.genome.gov/): US government's genome institute
- Pharmacogenomics in Drug Discovery and Development (http://www.springer.com/humana+press/pharmacology+and+toxicology/book/978-1-58829-887-4), a book on pharmacogenomics, diseases, personalized medicine, and therapeutics

- Tishchenko P. D. Genomics: New Science in the New Cultural Situation (http://www.zpu-journal.ru/en/articles/detail.php?ID=342)
- Undergraduate program on Genomic Sciences (spanish) (http://www.lcg.unam.mx/): One of the first undergraduate programs in the world
- JCVI Comprehensive Microbial Resource (http://cmr.jcvi.org/)
- Pathema: A Clade Specific Bioinformatics Resource Center (http://pathema.jcvi.org/)
- KoreaGenome.org (http://koreagenome.org): The first Korean Genome published and the sequence is available freely.
- GenomicsNetwork (http://genomicsnetwork.ac.uk): Looks at the development and use of the science and technologies of genomics.

Genetic algorithm

A **genetic algorithm (GA)** is a search technique used in computing to find exact or approximate solutions to optimization and search problems. Genetic algorithms are categorized as global search heuristics. Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover (also called recombination).

Methodology

Genetic algorithms are implemented in a computer simulation in which a population of abstract representations (called chromosomes or the genotype of the genome) of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

Genetic algorithms find application in bioinformatics, phylogenetics, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics and other fields.

A typical genetic algorithm requires:

1. a genetic representation of the solution domain,
2. a fitness function to evaluate the solution domain.

A standard representation of the solution is as an array of bits. Arrays of other types and structures can be used in essentially the same way. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size, which facilitates simple crossover operations. Variable length representations may also be used, but crossover implementation is more complex in this case. Tree-like representations are explored in genetic programming and graph-form representations are

explored in evolutionary programming.

The fitness function is defined over the genetic representation and measures the *quality* of the represented solution. The fitness function is always problem dependent. For instance, in the knapsack problem one wants to maximize the total value of objects that can be put in a knapsack of some fixed capacity. A representation of a solution might be an array of bits, where each bit represents a different object, and the value of the bit (0 or 1) represents whether or not the object is in the knapsack. Not every such representation is valid, as the size of objects may exceed the capacity of the knapsack. The *fitness* of the solution is the sum of values of all objects in the knapsack if the representation is valid, or 0 otherwise. In some problems, it is hard or even impossible to define the fitness expression; in these cases, interactive genetic algorithms are used.

Once we have the genetic representation and the fitness function defined, GA proceeds to initialize a population of solutions randomly, then improve it through repetitive application of mutation, crossover, inversion and selection operators.

Initialization

Initially many individual solutions are randomly generated to form an initial population. The population size depends on the nature of the problem, but typically contains several hundreds or thousands of possible solutions. Traditionally, the population is generated randomly, covering the entire range of possible solutions (the *search space*). Occasionally, the solutions may be "seeded" in areas where optimal solutions are likely to be found.

Selection

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a *fitness-based* process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions. Other methods rate only a random sample of the population, as this process may be very time-consuming.

Most functions are stochastic and designed so that a small proportion of less fit solutions are selected. This helps keep the diversity of the population large, preventing premature convergence on poor solutions. Popular and well-studied selection methods include roulette wheel selection and tournament selection.

Reproduction

The next step is to generate a second generation population of solutions from those selected through genetic operators: crossover (also called recombination), and/or mutation.

For each new solution to be produced, a pair of "parent" solutions is selected for breeding from the pool selected previously. By producing a "child" solution using the above methods of crossover and mutation, a new solution is created which typically shares many of the characteristics of its "parents". New parents are selected for each child, and the process continues until a new population of solutions of appropriate size is generated. Although reproduction methods that are based on the use of two parents are more "biology inspired", recent researches (Islam Abou El Ata 2006) suggested more than two "parents" are better to be used to reproduce a good quality chromosome.

These processes ultimately result in the next generation population of chromosomes that is different from the initial generation. Generally the average fitness will have increased by this procedure for the population, since only the best organisms from the first generation are selected for breeding, along with a small proportion of less fit solutions, for reasons already mentioned above.

Termination

This generational process is repeated until a termination condition has been reached. Common terminating conditions are:

- A solution is found that satisfies minimum criteria
- Fixed number of generations reached
- Allocated budget (computation time/money) reached
- The highest ranking solution's fitness is reaching or has reached a plateau such that successive iterations no longer produce better results
- Manual inspection
- Combinations of the above

Simple generational genetic algorithm pseudocode

1. Choose initial population
2. Evaluate the fitness of each individual in the population
3. Repeat until termination: (time limit or sufficient fitness achieved)
    1. Select best-ranking individuals to reproduce
    2. Breed new generation through crossover and/or mutation (genetic operations) and give birth to offspring
    3. Evaluate the individual fitnesses of the offspring
    4. Replace worst ranked part of population with offspring

Observations

There are several general observations about the generation of solutions via a genetic algorithm:

- Repeated fitness function evaluation for complex problems is often the most prohibitive and limiting segment of artificial evolutionary algorithms. Finding optimal solution to complex high dimensional, multimodal problems often requires very expensive fitness function evaluations. In real world problems such as structural optimization problems, one single function evaluation may require several hours to several days of complete simulation. Typical optimization method can not deal with such a type of problem. In this case, it may be necessary to forgo an exact evaluation and use an approximated fitness that is computationally efficient. It is apparent that amalgamation of approximate models may be one of the most promising approaches to convincingly use EA to solve complex real life problems.
- The "better" is only in comparison to other solution. As a result, the stop criterion is not clear.
- In many problems, GAs may have a tendency to converge towards local optima or even arbitrary points rather than the global optimum of the problem. This means that it does not "know how" to sacrifice short-term fitness to gain longer-term fitness. The likelihood of this occurring depends on the shape of the fitness landscape: certain problems may

provide an easy ascent towards a global optimum, others may make it easier for the function to find the local optima. This problem may be alleviated by using a different fitness function, increasing the rate of mutation, or by using selection techniques that maintain a diverse population of solutions, although the No Free Lunch theorem proves that there is no general solution to this problem. A common technique to maintain diversity is to impose a "niche penalty", wherein, any group of individuals of sufficient similarity (niche radius) have a penalty added, which will reduce the representation of that group in subsequent generations, permitting other (less similar) individuals to be maintained in the population. This trick, however, may not be effective, depending on the landscape of the problem. Diversity is important in genetic algorithms (and genetic programming) because crossing over a homogeneous population does not yield new solutions. In evolution strategies and evolutionary programming, diversity is not essential because of a greater reliance on mutation.

- Operating on dynamic data sets is difficult, as genomes begin to converge early on towards solutions which may no longer be valid for later data. Several methods have been proposed to remedy this by increasing genetic diversity somehow and preventing early convergence, either by increasing the probability of mutation when the solution quality drops (called *triggered hypermutation*), or by occasionally introducing entirely new, randomly generated elements into the gene pool (called *random immigrants*). Again, evolution strategies and evolutionary programming can be implemented with a so-called "comma strategy" in which parents are not maintained and new parents are selected only from offspring. This can be more effective on dynamic problems.
- GAs cannot effectively solve problems in which the only fitness measure is a single right/wrong measure (like decision problems), as there is no way to converge on the solution (no hill to climb). In these cases, a random search may find a solution as quickly as a GA. *However*, if the situation allows the success/failure trial to be repeated giving (possibly) different results, then the ratio of successes to failures provides a suitable fitness measure.
- Selection is clearly an important genetic operator, but opinion is divided over the importance of crossover versus mutation. Some argue that crossover is the most important, while mutation is only necessary to ensure that potential solutions are not lost. Others argue that crossover in a largely uniform population only serves to propagate innovations originally found by mutation, and in a non-uniform population crossover is nearly always equivalent to a very large mutation (which is likely to be catastrophic). There are many references in Fogel (2006) that support the importance of mutation-based search, but across all problems the No Free Lunch theorem holds, so these opinions are without merit unless the discussion is restricted to a particular problem.
- Often, GAs can rapidly locate *good* solutions, even for difficult search spaces. The same is of course also true for evolution strategies and evolutionary programming.
- For specific optimization problems and problem instances, other optimization algorithms may find better solutions than genetic algorithms (given the same amount of computation time). Alternative and complementary algorithms include evolution strategies, evolutionary programming, simulated annealing, Gaussian adaptation, hill climbing, and swarm intelligence (e.g.: ant colony optimization, particle swarm optimization) and methods based on integer linear programming. The question of which, if any, problems are suited to genetic algorithms (in the sense that such algorithms are better than

others) is open and controversial.
- As with all current machine learning problems it is worth tuning the parameters such as mutation probability, recombination probability and population size to find reasonable settings for the problem class being worked on. A very small mutation rate may lead to genetic drift (which is non-ergodic in nature). A recombination rate that is too high may lead to premature convergence of the genetic algorithm. A mutation rate that is too high may lead to loss of good solutions unless there is elitist selection. There are theoretical but not yet practical upper and lower bounds for these parameters that can help guide selection.
- The implementation and evaluation of the fitness function is an important factor in the speed and efficiency of the algorithm.

Variants

The simplest algorithm represents each chromosome as a bit string. Typically, numeric parameters can be represented by integers, though it is possible to use floating point representations. The floating point representation is natural to evolution strategies and evolutionary programming. The notion of real-valued genetic algorithms has been offered but is really a misnomer because it does not really represent the building block theory that was proposed by Holland in the 1970s. This theory is not without support though, based on theoretical and experimental results (see below). The basic algorithm performs crossover and mutation at the bit level. Other variants treat the chromosome as a list of numbers which are indexes into an instruction table, nodes in a linked list, hashes, objects, or any other imaginable data structure. Crossover and mutation are performed so as to respect data element boundaries. For most data types, specific variation operators can be designed. Different chromosomal data types seem to work better or worse for different specific problem domains.

When bit strings representations of integers are used, Gray coding is often employed. In this way, small changes in the integer can be readily effected through mutations or crossovers. This has been found to help prevent premature convergence at so called *Hamming walls*, in which too many simultaneous mutations (or crossover events) must occur in order to change the chromosome to a better solution.

Other approaches involve using arrays of real-valued numbers instead of bit strings to represent chromosomes. Theoretically, the smaller the alphabet, the better the performance, but paradoxically, good results have been obtained from using real-valued chromosomes.

A very successful (slight) variant of the general process of constructing a new population is to allow some of the better organisms from the current generation to carry over to the next, unaltered. This strategy is known as *elitist selection*.

Parallel implementations of genetic algorithms come in two flavours. Coarse grained parallel genetic algorithms assume a population on each of the computer nodes and migration of individuals among the nodes. Fine grained parallel genetic algorithms assume an individual on each processor node which acts with neighboring individuals for selection and reproduction. Other variants, like genetic algorithms for online optimization problems, introduce time-dependence or noise in the fitness function.

It can be quite effective to combine GA with other optimization methods. GA tends to be quite good at finding generally good global solutions, but quite inefficient at finding the last

few mutations to find the absolute optimum. Other techniques (such as simple hill climbing) are quite efficient at finding absolute optimum in a limited region. Alternating GA and hill climbing can improve the efficiency of GA while overcoming the lack of robustness of hill climbing.

This means that the rules of genetic variation may have a different meaning in the natural case. For instance - provided that steps are stored in consecutive order - crossing over may sum a number of steps from maternal DNA adding a number of steps from paternal DNA and so on. This is like adding vectors that more probably may follow a ridge in the phenotypic landscape. Thus, the efficiency of the process may be increased by many orders of magnitude. Moreover, the inversion operator has the opportunity to place steps in consecutive order or any other suitable order in favour of survival or efficiency. (See for instance [1] or example in travelling salesman problem.)

Population-based incremental learning is a variation where the population as a whole is evolved rather than its individual members.

Problem domains

Problems which appear to be particularly appropriate for solution by genetic algorithms include timetabling and scheduling problems, and many scheduling software packages are based on GAs. GAs have also been applied to engineering. Genetic algorithms are often applied as an approach to solve global optimization problems.

As a general rule of thumb genetic algorithms might be useful in problem domains that have a complex fitness landscape as recombination is designed to move the population away from local optima that a traditional hill climbing algorithm might get stuck in.

History

Computer simulations of evolution started as early as in 1954 with the work of Nils Aall Barricelli, who was using the computer at the Institute for Advanced Study in Princeton, New Jersey.[2] [3] His 1954 publication was not widely noticed. Starting in 1957 [4] , the Australian quantitative geneticist Alex Fraser published a series of papers on simulation of artificial selection of organisms with multiple loci controlling a measurable trait. From these beginnings, computer simulation of evolution by biologists became more common in the early 1960s, and the methods were described in books by Fraser and Burnell (1970)[5] and Crosby (1973)[6] . Fraser's simulations included all of the essential elements of modern genetic algorithms. In addition, Hans Bremermann published a series of papers in the 1960s that also adopted a population of solution to optimization problems, undergoing recombination, mutation, and selection. Bremermann's research also included the elements of modern genetic algorithms. Other noteworthy early pioneers include Richard Friedberg, George Friedman, and Michael Conrad. Many early papers are reprinted by Fogel (1998).[7]

Although Barricelli, in work he reported in 1963, had simulated the evolution of ability to play a simple game,[8] artificial evolution became a widely recognized optimization method as a result of the work of Ingo Rechenberg and Hans-Paul Schwefel in the 1960s and early 1970s - Rechenberg's group was able to solve complex engineering problems through evolution strategies [9] [10] [11] [12] . Another approach was the evolutionary programming technique of Lawrence J. Fogel, which was proposed for generating artificial intelligence. Evolutionary programming originally used finite state machines for predicting

environments, and used variation and selection to optimize the predictive logics. Genetic algorithms in particular became popular through the work of John Holland in the early 1970s, and particularly his book *Adaptation in Natural and Artificial Systems* (1975). His work originated with studies of cellular automata, conducted by Holland and his students at the University of Michigan. Holland introduced a formalized framework for predicting the quality of the next generation, known as Holland's Schema Theorem. Research in GAs remained largely theoretical until the mid-1980s, when The First International Conference on Genetic Algorithms was held in Pittsburgh, Pennsylvania.

As academic interest grew, the dramatic increase in desktop computational power allowed for practical application of the new technique. In the late 1980s, General Electric started selling the world's first genetic algorithm product, a mainframe-based toolkit designed for industrial processes. In 1989, Axcelis, Inc. released Evolver, the world's second GA product and the first for desktop computers. The New York Times technology writer John Markoff wrote[13] about Evolver in 1990.

Related techniques

- Ant colony optimization (ACO) uses many ants (or agents) to traverse the solution space and find locally productive areas. While usually inferior to genetic algorithms and other forms of local search, it is able to produce results in problems where no global or up-to-date perspective can be obtained, and thus the other methods cannot be applied.
- Bacteriologic algorithms (BA) inspired by evolutionary ecology and, more particularly, bacteriologic adaptation. Evolutionary ecology is the study of living organisms in the context of their environment, with the aim of discovering how they adapt. Its basic concept is that in a heterogeneous environment, you can’t find one individual that fits the whole environment. So, you need to reason at the population level. BAs have shown better results than GAs on problems such as complex positioning problems (antennas for cell phones, urban planning, and so on) or data mining.[14]
- Cross-entropy method The cross-entropy (CE) method generates candidates solutions via a parameterized probability distribution. The parameters are updated via cross-entropy minimization, so as to generate better samples in the next iteration.
- Cultural algorithm (CA) consists of the population component almost identical to that of the genetic algorithm and, in addition, a knowledge component called the belief space.
- Evolution strategies (ES, see Rechenberg, 1994) evolve individuals by means of mutation and intermediate and discrete recombination. ES algorithms are designed particularly to solve problems in the real-value domain. They use self-adaptation to adjust control parameters of the search.
- Evolutionary programming (EP) involves populations of solutions with primarily mutation and selection and arbitrary representations. They use self-adaptation to adjust parameters, and can include other variation operations such as combining information from multiple parents.
- Extremal optimization (EO) Unlike GAs, which work with a population of candidate solutions, EO evolves a single solution and makes local modifications to the worst components. This requires that a suitable representation be selected which permits individual solution components to be assigned a quality measure ("fitness"). The governing principle behind this algorithm is that of *emergent* improvement through

selectively removing low-quality components and replacing them with a randomly selected component. This is decidedly at odds with a GA that selects good solutions in an attempt to make better solutions.

- Gaussian adaptation (normal or natural adaptation, abbreviated NA to avoid confusion with GA) is intended for the maximisation of manufacturing yield of signal processing systems. It may also be used for ordinary parametric optimisation. It relies on a certain theorem valid for all regions of acceptability and all Gaussian distributions. The efficiency of NA relies on information theory and a certain theorem of efficiency. Its efficiency is defined as information divided by the work needed to get the information[15] . Because NA maximises mean fitness rather than the fitness of the individual, the landscape is smoothed such that valleys between peaks may disappear. Therefore it has a certain “ambition” to avoid local peaks in the fitness landscape. NA is also good at climbing sharp crests by adaptation of the moment matrix, because NA may maximise the disorder (average information) of the Gaussian simultaneously keeping the mean fitness constant.
- Genetic programming (GP) is a related technique popularized by John Koza in which computer programs, rather than function parameters, are optimized. Genetic programming often uses tree-based internal data structures to represent the computer programs for adaptation instead of the list structures typical of genetic algorithms.
- Grouping genetic algorithm (GGA) is an evolution of the GA where the focus is shifted from individual items, like in classical GAs, to groups or subset of items.[16] The idea behind this GA evolution proposed by Emanuel Falkenauer is that solving some complex problems, a.k.a. *clustering* or *partitioning* problems where a set of items must be split into disjoint group of items in an optimal way, would better be achieved by making characteristics of the groups of items equivalent to genes. These kind of problems include Bin Packing, Line Balancing, Clustering w.r.t. a distance measure, Equal Piles, etc., on which classic GAs proved to perform poorly. Making genes equivalent to groups implies chromosomes that are in general of variable length, and special genetic operators that manipulate whole groups of items. For Bin Packing in particular, a GGA hybridized with the Dominance Criterion of Martello and Toth, is arguably the best technique to date.
- Harmony search (HS) is an algorithm mimicking musicians behaviors in improvisation process.
- Interactive evolutionary algorithms are evolutionary algorithms that use human evaluation. They are usually applied to domains where it is hard to design a computational fitness function, for example, evolving images, music, artistic designs and forms to fit users' aesthetic preference.
- Memetic algorithm (MA), also called *hybrid genetic algorithm* among others, is a relatively new evolutionary method where local search is applied during the evolutionary cycle. The idea of memetic algorithms comes from memes, which unlike genes, can adapt themselves. In some problem areas they are shown to be more efficient than traditional evolutionary algorithms.
- Simulated annealing (SA) is a related global optimization technique that traverses the search space by testing random mutations on an individual solution. A mutation that increases fitness is always accepted. A mutation that lowers fitness is accepted probabilistically based on the difference in fitness and a decreasing temperature parameter. In SA parlance, one speaks of seeking the lowest energy instead of the

maximum fitness. SA can also be used within a standard GA algorithm by starting with a relatively high rate of mutation and decreasing it over time along a given schedule.

- Stochastic optimization is an umbrella set of methods that includes GAs and numerous other approaches.
- Tabu search (TS) is similar to simulated annealing in that both traverse the solution space by testing mutations of an individual solution. While simulated annealing generates only one mutated solution, tabu search generates many mutated solutions and moves to the solution with the lowest energy of those generated. In order to prevent cycling and encourage greater movement through the solution space, a tabu list is maintained of partial or complete solutions. It is forbidden to move to a solution that contains elements of the tabu list, which is updated as the solution traverses the solution space.

Building block hypothesis

Genetic algorithms are relatively simple to implement, but their behavior is difficult to understand. In particular it is difficult to understand why they are often successful in generating solutions of high fitness. The building block hypothesis (BBH) consists of:

1. A description of an abstract adaptive mechanism that performs adaptation by recombining "building blocks", i.e. low order, low defining-length schemata with above average fitness.
2. A hypothesis that a genetic algorithm performs adaptation by implicitly and efficiently implementing this abstract adaptive mechanism.

(Goldberg 1989:41) describes the abstract adaptive mechanism as follows:

> Short, low order, and highly fit schemata are sampled, recombined [crossed over], and resampled to form strings of potentially higher fitness. In a way, by working with these particular schemata [the building blocks], we have reduced the complexity of our problem; instead of building high-performance strings by trying every conceivable combination, we construct better and better strings from the best partial solutions of past samplings.
>
> Just as a child creates magnificent fortresses through the arrangement of simple blocks of wood [building blocks], so does a genetic algorithm seek near optimal performance through the juxtaposition of short, low-order, high-performance schemata, or building blocks.

(Goldberg 1989) claims that the building block hypothesis is supported by Holland's schema theorem.

The building block hypothesis has been sharply criticized on the grounds that it lacks theoretical justification and experimental results have been published that draw its veracity into question. On the theoretical side, for example, Wright et al. state that

> "The various claims about GAs that are traditionally made under the name of the *building block hypothesis* have, to date, no basis in theory and, in some cases, are simply incoherent"[17]

On the experimental side uniform crossover was seen to outperform one-point and two-point crossover on many of the fitness functions studied by Syswerda.[18] Summarizing these results, Fogel remarks that

> "Generally, uniform crossover yielded better performance than two-point crossover, which in turn yielded better performance than one-point crossover"[19]

Syswerda's results contradict the building block hypothesis because uniform crossover is extremely disruptive of short schemata whereas one and two-point crossover are more likely to conserve short schemata and combine their defining bits in children produced during recombination.

The debate over the building block hypothesis demonstrates that the issue of how GAs "work", (i.e. perform adaptation) is currently far from settled.

See also

- Algorithmic efficiency
- Holland's schema theorem
- Genetic programming
- Fitness approximation

Applications

- Artificial creativity
- Automated design, including research on composite material design and multi-objective design of automotive components for crashworthiness, weight savings, and other characteristics.
- Automated design of mechatronic systems using bond graphs and genetic programming (NSF).
- Automated design of industrial equipment using catalogs of exemplar lever patterns.
- Automated design of sophisticated trading systems in the financial sector.
- Building phylogenetic trees.[20]
- Calculation of bound states and local-density approximations.
- Chemical kinetics (gas [21] and solid [22] phases)
- Configuration applications, particularly physics applications of optimal molecule configurations for particular systems like C60 (buckyballs).
- Container loading optimization.
- Code-breaking, using the GA to search large solution spaces of ciphers for the one correct decryption.[23]
- Design of water distribution systems.
- Distributed computer network topologies.
- Electronic circuit design, known as Evolvable hardware.
- File allocation for a distributed system.
- Game Theory Equilibrium Resolution.
- Gene expression profiling analysis.[24]
- Genetic Algorithm for Rule Set Production
- Learning Robot behavior using Genetic Algorithms.
- Learning fuzzy rule base using genetic algorithms.
- Linguistic analysis, including Grammar induction and other aspects of Natural language processing (NLP) such as word sense disambiguation.
- Marketing Mix Analysis
- Mobile communications infrastructure optimization.
- Molecular Structure Optimization (Chemistry).
- Multiple criteria production scheduling.[25]
- Multiple population topologies and interchange methodologies.

- Mutation testing
- Neural Networks; particularly recurrent neural networks[26]
- Operon prediction.[27]
- Optimisation of data compression systems, for example using wavelets.
- Parallelization of GAs/GPs including use of hierarchical decomposition of problem domains and design spaces nesting of irregular shapes using feature matching and GAs.
- Pop music record producer[28] .
- Protein folding and protein/ligand docking.[29]
- Plant floor layout.
- Representing rational agents in economic models such as the cobweb model.
- Bioinformatics: RNA structure prediction.[30]
- Bioinformatics: [Multiple Sequence Alignment].[31] . SAGA is available on: [32].
- Bioinformatics Multiple sequence alignment.[33]
- Scheduling applications, including job-shop scheduling. The objective being to schedule jobs in a sequence dependent or non-sequence dependent setup environment in order to maximize the volume of production while minimizing penalties such as tardiness.
- Selection of optimal mathematical model to describe biological systems.
- Software engineering
- Solving the machine-component grouping problem required for cellular manufacturing systems.
- Tactical asset allocation and international equity strategies.
- Timetabling problems, such as designing a non-conflicting class timetable for a large university.
- Training artificial neural networks when pre-classified training examples are not readily obtainable (neuroevolution).
- Traveling Salesman Problem.
- Finding hardware bugs. [34] [35]
- Wireless Sensor/Ad-hoc Networks. [36]
- Data Center/Server Farm. [37]

Notes

[1] http://www.evolution-in-a-nutshell.se/traveller.htm
[2] Barricelli, Nils Aall (1954). "Esempi numerici di processi di evoluzione". *Methodos*: 45–68.
[3] Barricelli, Nils Aall (1957). "Symbiogenetic evolution processes realized by artificial methods". *Methodos*: 143–182.
[4] Fraser, Alex (1957). "Simulation of genetic systems by automatic digital computers. I. Introduction". *Aust. J. Biol. Sci.* **10**: 484–491.
[5] Fraser, Alex; Donald Burnell (1970). *Computer Models in Genetics*. New York: McGraw-Hill.
[6] Crosby, Jack L. (1973). *Computer Simulation in Genetics*. London: John Wiley & Sons.
[7] Fogel, David B. (editor) (1998). *Evolutionary Computation: The Fossil Record*. New York: IEEE Press.
[8] Barricelli, Nils Aall (1963). "Numerical testing of evolution theories. Part II. Preliminary tests of performance, symbiogenesis and terrestrial life". *Acta Biotheoretica* (16): 99–126.
[9] Rechenberg, Ingo (1973). *Evolutionsstrategie*. Stuttgart: Holzmann-Froboog.
[10] Schwefel, Hans-Paul (1974). *Numerische Optimierung von Computer-Modellen (PhD thesis)*.
[11] Schwefel, Hans-Paul (1977). *Numerische Optimierung von Computor-Modellen mittels der Evolutionsstrategie : mit einer vergleichenden Einführung in die Hill-Climbing- und Zufallsstrategie*. Basel; Stuttgart: Birkhäuser. ISBN 3764308761.
[12] Schwefel, Hans-Paul (1981). *Numerical optimization of computer models (Translation of 1977 Numerische Optimierung von Computor-Modellen mittels der Evolutionsstrategie*. Chichester ; New York: Wiley. ISBN 0471099880.

[13] Markoff, John (1989). "What's the Best Answer? It's Survival of the Fittest" (http://query.nytimes.com/gst/fullpage.html?res=9C0CE1D6153BF93AA1575BC0A966958260). New York Times. .
[14] Baudry, Benoit; Franck Fleurey, Jean-Marc Jézéquel, and Yves Le Traon (March/April 2005). " Automatic Test Case Optimization: A Bacteriologic Algorithm (http://www.irisa.fr/triskell/publis/2005/Baudry05d.pdf)" (PDF). *IEEE Software* (IEEE Computer Society) **22**: 76–82. doi: 10.1109/MS.2005.30 (http://dx.doi.org/10.1109/MS.2005.30). .
[15] Kjellström, G. (December 1991). "On the Efficiency of Gaussian Adaptation". *Journal of Optimization Theory and Applications* **71** (3): 589–597. doi: 10.1007/BF00941405 (http://dx.doi.org/10.1007/BF00941405).
[16] Falkenauer, Emanuel (1997). *Genetic Algorithms and Grouping Problems*. Chichester, England: John Wiley & Sons Ltd. ISBN 978-0-471-97150-4.
[17] Wright, A.H.; et al. (2003). "Implicit Parallelism". *Proceedings of the Genetic and Evolutionary Computation Conference*.
[18] Syswerda, G. (1989). "Uniform crossover in genetic algorithms". J. D. Schaffer *Proceedings of the Third International Conference on Genetic Algorithms*, Morgan Kaufmann.
[19] Fogel, David B. (2000). *Evolutionary Computation: Towards a New Philosophy of Machine Intelligence*. New York: IEEE Press. pp. 140.
[20] Hill T, Lundgren A, Fredriksson R, Schiöth HB (2005). "Genetic algorithm for large-scale maximum parsimony phylogenetic analysis of proteins". *Biochimica et Biophysica Acta* **1725**: 19–29. PMID 15990235.
[21] http://www.personal.leeds.ac.uk/~fuensm/project.html
[22] http://repositories.cdlib.org/postprints/1154
[23] Joachim De Zutter "Codebreaking" (http://byterage.hackaholic.org/kb/codebreaking.pdf)
[24] To CC, Vohradsky J (2007). "A parallel genetic algorithm for single class pattern classification and its application for gene expression profiling in Streptomyces coelicolor". *BMC Genomics* **8**: 49. doi: 10.1186/1471-2164-8-49 (http://dx.doi.org/10.1186/1471-2164-8-49). PMID 17298664.
[25] Bagchi Tapan P (1999). *Multiobjective Scheduling by Genetic Algorithms*.
[26] Applying Genetic Algorithms to Recurrent Neural Networks for Learning Network Parameters and Architecture (http://arimaa.com/arimaa/about/Thesis/)
[27] Wang S, Wang Y, Du W, Sun F, Wang X, Zhou C, Liang Y (2007). "A multi-approaches-guided genetic algorithm with application to operon prediction". *Artificial Intelligence in Medicine* **41**: 151–159. doi: 10.1016/j.artmed.2007.07.010 (http://dx.doi.org/10.1016/j.artmed.2007.07.010). PMID 17869072.
[28] BBC News | Entertainment | To the beat of the byte (http://news.bbc.co.uk/1/hi/entertainment/123983.stm)
[29] Willett P (1995). "Genetic algorithms in molecular recognition and design". *Trends in Biotechnology* **13**: 516–521. doi: 10.1016/S0167-7799(00)89015-0 (http://dx.doi.org/10.1016/S0167-7799(00)89015-0). PMID 8595137.
[30] van Batenburg FH, Gultyaev AP, Pleij CW (1995). "An APL-programmed genetic algorithm for the prediction of RNA secondary structure". *Journal of Theoretical Biology* **174**: 269–280. doi: 10.1006/jtbi.1995.0098 (http://dx.doi.org/10.1006/jtbi.1995.0098). PMID 7545258.
[31] Notredame C, Higgins DG (1995). "SAGA a Genetic Algorithm for Multiple Sequence Alignment". *Nulceic Acids Research* **174**: 1515. PMID 8628686.
[32] http://www.tcoffee.org/homepage.html
[33] Gondro C, Kinghorn BP (2007). "A simple genetic algorithm for multiple sequence alignment". *Genetics and Molecular Research* **6**: 964–982. PMID 18058716.
[34] Hitoshi Iba, Sumitaka Akiba, Tetsuya Higuchi, Taisuke Sato: BUGS: A Bug-Based Search Strategy using Genetic Algorithms. PPSN 1992:
[35] Ibrahim, W. and Amer, H.: An Adaptive Genetic Algorithm for VLSI Test Vector Selection
[36] BiSNET/e - Distributed Software Systems Group, University of Massachusetts, Boston (http://dssg.cs.umb.edu/wiki/index.php/BiSNET/e)
[37] SymbioticSphere - Distributed Software Systems Group, University of Massachusetts, Boston (http://dssg.cs.umb.edu/wiki/index.php/SymbioticSphere)

References

- Banzhaf, Wolfgang; Nordin, Peter; Keller, Robert; Francone, Frank (1998) *Genetic Programming - An Introduction*, Morgan Kaufmann, San Francisco, CA.
- Bies, Robert R; Muldoon, Matthew F; Pollock, Bruce G; Manuck, Steven; Smith, Gwenn and Sale, Mark E (2006). "A Genetic Algorithm-Based, Hybrid Machine Learning Approach to Model Selection". *Journal of Pharmacokinetics and Pharmacodynamics* (Netherlands: Springer): 196–221.
- Fraser, Alex S. (1957). "Simulation of Genetic Systems by Automatic Digital Computers. I. Introduction". *Australian Journal of Biological Sciences* **10**: 484–491.
- Goldberg, David E (1989), *Genetic Algorithms in Search, Optimization and Machine Learning,* Kluwer Academic Publishers, Boston, MA.
- Goldberg, David E (2002), *The Design of Innovation: Lessons from and for Competent Genetic Algorithms,* Addison-Wesley, Reading, MA.
- Fogel, David B (2006), *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence,* IEEE Press, Piscataway, NJ. Third Edition
- Holland, John H (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor
- Koza, John (1992), *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press. ISBN 0-262-11170-5
- Michalewicz, Zbigniew (1999), *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag.
- Mitchell, Melanie, (1996), *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA.
- Poli, R., Langdon, W. B., McPhee, N. F. (2008). *A Field Guide to Genetic Programming*. Lulu.com, freely available from the internet. ISBN 978-1-4092-0073-4.
- Rechenberg, Ingo (1994): Evolutionsstrategie '94, Stuttgart: Fromman-Holzboog.
- Schmitt, Lothar M; Nehaniv, Chrystopher L; Fujii, Robert H (1998), *Linear analysis of genetic algorithms*, Theoretical Computer Science 208: 111-148
- Schmitt, Lothar M (2001), *Theory of Genetic Algorithms*, Theoretical Computer Science 259: 1-61
- Schmitt, Lothar M (2004), *Theory of Genetic Algorithms II: models for genetic operators over the string-tensor representation of populations and convergence to global optima for arbitrary fitness function under scaling*, Theoretical Computer Science 310: 181-231
- Schwefel, Hans-Paul (1974): Numerische Optimierung von Computer-Modellen (PhD thesis). Reprinted by Birkhäuser (1977).
- Vose, Michael D (1999), *The Simple Genetic Algorithm: Foundations and Theory*, MIT Press, Cambridge, MA.
- Whitley, D. (1994). *A genetic algorithm tutorial*. Statistics and Computing 4, 65–85.

External links

- (http://twtmas.mpei.ac.ru/mas/Worksheets/Minimum.mcd) Search of Global Minimum by genetic algorithme

Controversy

- The Fundamental Problem with the Building Block Hypothesis (http://blog.hackingevolution.org/2008/10/18/new-manuscript-the-fundamental-problem-with-the-building-block-hypothesis/) A description and critique of the assumptions that undergird the building block hypothesis

Applications

- Demo applet of a evolutionary algorithm for solving TSP's and VRPTW problems (http://www.dna-evolutions.com/dnaappletsample.html)
- Genetic Arm (http://www.e-nuts.net/en/genetic-algorithms) Simulation of a mechanical arm trained using genetic algorithms. Custom goals can be defined using a scripting language. A sample video is available on page.
- Antenna optimization for NASA (http://ti.arc.nasa.gov/projects/esg/research/antenna.htm) A successful application of genetic algorithms.
- Genesis-SGA Seo genetic Algorithm (http://seo.witinside.net/genetic-algorithms/) Genetic algorithms applied to the theme SEO (Search Engine Optimization)

Resources

- DigitalBiology.NET (http://www.digitalbiology.net) Vertical search engine for GA/GP resources
- Genetic Algorithms Index (http://www.geneticprogramming.com/ga/index.htm) The site Genetic Programming Notebook provides a structured resource pointer to web pages in genetic algorithms field

Tutorials

- A Field Guide to Genetic Programming (http://www.gp-field-guide.org.uk/) A book, freely downloadable under a Creative Commons license.
- Introduction to Genetic Algorithms with interactive Java applets (http://www.obitko.com/tutorials/genetic-algorithms/) For experimenting with GAs online
- A Practical Tutorial on Genetic Algorithm (http://fog.neopages.org/helloworldgeneticalgorithms.php) Programming a Genetic Algorithm step by step.
- A Genetic Algorithm Tutorial by Darrell Whitley Computer Science Department Colorado State University (http://samizdat.mines.edu/ga_tutorial/ga_tutorial.ps) An excellent tutorial with lots of theory
- Cross discipline example applications for GAs with references. (http://www.toarchive.org/faqs/genalg/genalg.html)
- Global Optimization Algorithms - Theory and Application (http://www.it-weise.de/projects/book.pdf)

Libraries

- Demo applet of JOpt.SDK (http://www.dna-evolutions.com/dnaappletsample.html) an evolutionary algorithm software library for Java or .NET for solving TSP's and VRPTW problems
- Evoptool (http://airwiki.elet.polimi.it/mediawiki/index.php/Evoptool:_Evolutive_Optimization_Tool) A framework and a set of libraries written in C++ for the Evolutive Computation, including several Genetic Algorithms and EDAs.
- Jenes (http://sites.google.com/a/ciselab.org/jenes) An optimized Java library for Genetic Algorithms.
- Pyevolve (http://pyevolve.sourceforge.net/) A python framework for Genetic Algorithms.
- ParadisEO (http://paradiseo.gforge.inria.fr) A powerful C++ framework dedicated to the reusable design of metaheuristics, included genetic algorithms.
- Genetic Algorithms in Ruby (http://ai4r.rubyforge.org/geneticAlgorithms.html)
- GAlib (http://lancet.mit.edu/ga/) A C++ Library of Genetic Algorithm Components
- GAEDALib (http://laurel.datsi.fi.upm.es/projects/gaedalib) A C++ Library of Evolutive Algotithms (GAs, EDAs, DEs and others) based in GAlib, and supporting to MOS and parallel computing
- Jenetics (http://jenetics.sourceforge.net/) Genetic Algorithm Library written in Java.
- A Fortran code (PIKAIA) with a tutorial by Paul Charbonneau and Barry Knapp, National Center for Atmospheric Research. (http://www.hao.ucar.edu/Public/models/pikaia/pikaia.html) An excellent tutorial and a versatile public domain code. PIKAIA is also available in a version for Microsoft Excel (http://www.ecy.wa.gov/programs/eap/models.html), as well as a parallel processing version (http://whitedwarf.org/index.html?parallel/&0).
- ga (http://www.mathworks.com/access/helpdesk/help/toolbox/gads/ga.html) Genetic Algorithm in MATLAB (How GA in MATLAB works (http://www.mathworks.com/access/helpdesk/help/toolbox/gads/index.html?/access/helpdesk/help/toolbox/gads/f6187.html))
- gamultiobj (http://www.mathworks.com/access/helpdesk/help/toolbox/gads/gamultiobj.html) Multitobjective Genetic Algorithm in MATLAB
- GARAGe (http://garage.cse.msu.edu/) Michigan State University's Genetic Algorithm library in C, GALLOPS
- GAOT (http://www.ise.ncsu.edu/mirage/GAToolBox/gaot/) The Genetic Algorithm Optimization Toolbox (GAOT) for Matlab, by NCSU
- JGAP (http://jgap.sourceforge.net/) Java Genetic Algorithms Package features comprehensive unit tests
- speedyGA (http://blog.hackingevolution.net/2009/02/04/speedyga-v13/) A fast lightweight genetic algorithm in Matlab
- turboGA (http://blog.hackingevolution.net/2009/05/08/testing-the-efficacy-of-clamping/) An experimental genetic algorithm based on speedyGA

Metabolic network

A **metabolic network** is the complete set of metabolic and physical processes that determine the physiological and biochemical properties of a cell. As such, these networks comprise the chemical reactions of metabolism as well as the regulatory interactions that guide these reactions.

With the sequencing of complete genomes, it is now possible to reconstruct the network of biochemical reactions in many organisms, from bacteria to human. Several of these networks are available online: Kyoto Encyclopedia of Genes and Genomes (KEGG)[1], EcoCyc [2] and BioCyc [3]. Metabolic networks are powerful tools, for studying and modelling metabolism. From the study of metabolic networks' topology with graph theory to predictive toxicology and ADME.

See also

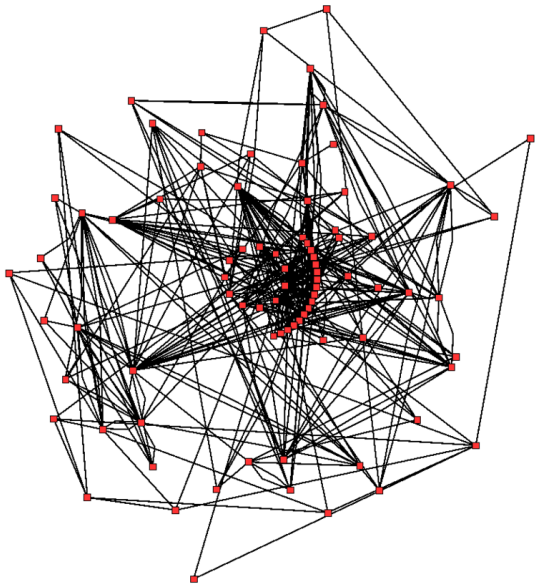
- Metabolic network modelling
- Metabolic pathway

References

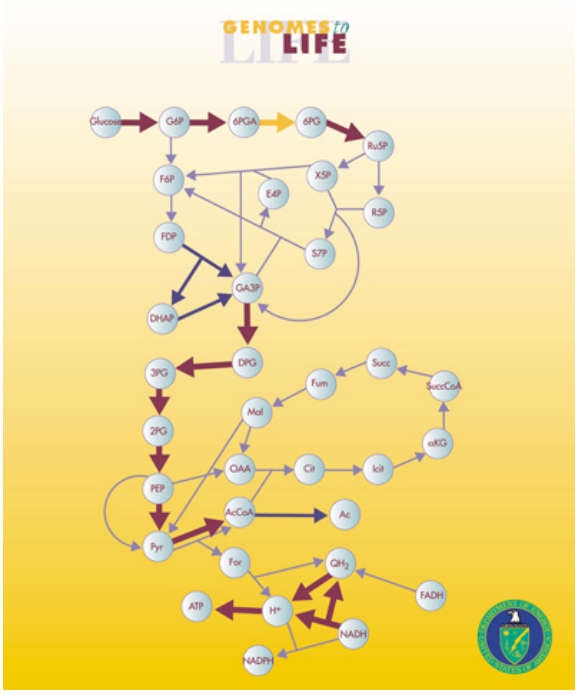
[1] http://www.genome.ad.jp
[2] http://www.ecocyc.org
[3] http://biocyc.org

Metabolic network modelling

Metabolic network reconstruction and simulation allows for an in depth insight into comprehending the molecular mechanisms of a particular organism, especially correlating the genome with molecular physiology (Francke, Siezen, and Teusink 2005). A reconstruction breaks down metabolism pathways into their respective reactions and enzymes, and analyzes them within the perspective of the entire network. Examples of various metabolic pathways include glycolysis, Krebs cycle, pentose phosphate pathway, etc. In simplified terms, a reconstruction involves collecting all of the relevant metabolic information of an organism and then compiling it in a way that makes sense for various types of analyses to be performed. The correlation between the genome and metabolism is made by searching gene databases, such as KEGG [1], GeneDB [2], etc., for particular genes by inputting enzyme or protein names. For example, a search can be conducted based on the protein name or the EC number (a number that represents the catalytic function of the enzyme of interest) in order to find the associated gene (Francke *et al.* 2005).



Metabolic network showing interactions between enzymes and metabolites in the *Arabidopsis thaliana* citric acid cycle. Enzymes and metabolites are the red dots and interactions between them are the lines.



Metabolic Network Model for Escherichia coli.

Beginning steps of a reconstruction

Resources

Below is more detailed description of a few gene/enzyme/reaction/pathway databases that are crucial to a metabolic reconstruction:

- **Kyoto Encyclopedia of Genes and Genomes (KEGG)**: This is a bioinformatics database containing information on genes, proteins, reactions, and pathways. The ‘KEGG Organisms’ section, which is divided into eukaryotes and prokaryotes, encompasses many organisms for which gene and DNA information can be searched by typing in the enzyme of choice. This resource can be extremely useful when building the association between metabolism enzymes, reactions and genes.

- **Gene DataBase (GeneDB)**: Similar to the KEGG resource, the Gene DataBase provides access to genomes of various organisms. If a search for hexokinase is carried out, genes for the organism of interest can be easily found. Moreover, the metabolic process associated with the enzyme is also listed along with the information on the genes (in the case of hexokinase, the pathway is glycolysis). Therefore, with one click, it is very easy to access all the different genes that are associated with glycolysis. Furthermore, GeneDB has a hierarchical organizational structure for metabolism, and it is possible to see at what level of the chain one is currently working on. This helps broaden an understanding of the biological and chemical processes that are involved in the organism.
- **BioCyc, EcoCyc and MetaCyc**: BioCyc is a collection of over 200 pathway/genome databases, containing whole databases dedicated to certain organisms. For example, EcoCyc which falls under the giant umbrella of BioCyc, is a highly detailed bioinformatics database on the genome and metabolic reconstruction of *Escherichia Coli*, including thorough descriptions of the various signaling pathways. The EcoCyc database can serve as a paradigm and model for any reconstruction. Additionally, MetaCyc, an encyclopedia of metabolic pathways, contains a wealth of information on metabolic reactions derived from over 600 different organisms.
- **Pathway Tools [3]**: This is a bioinformatics package that assists in the construction of pathway/genome databases such as EcoCyc (Francke *et al.* 2005). Developed by Peter Karp and associates at the SRI International Bioinformatics Group, Pathway Tools comprises several separate units that work together to generate new pathway/genome databases. First, PathoLogic takes an annotated genome for an organism and infers probable metabolic pathways to produce a new pathway/genome database. This can be followed by application of the Pathway Hole Filler, which predicts likely genes to fill "holes" (missing steps) in predicted pathways. Afterward, the Pathway Tools Navigator and Editor functions let users visualize, analyze, access and update the database. Thus, using PathoLogic and encyclopedias like MetaCyc, an initial fast reconstruction can be developed automatically, and then using the other units of Pathway Tools, a very detailed manual update, curation and verification step can be carried out (SRI 2005).
- **ENZYME**: This is an enzyme nomenclature database (part of the ExPASY [4] proteonomics server of the Swiss Institute of Bioinformatics). After searching for a particular enzyme on the database, this resource gives you the reaction that is catalyzed. Additionally, ENZYME has direct links to various other gene/enzyme/medical literature databases such as KEGG, BRENDA, PUBMED, and PUMA2 to name a few.
- **BRENDA**: A comprehensive enzyme database, BRENDA, allows you to search for an enzyme by name or EC number. You can also search for an organism and find all the relevant enzyme information. Moreover, when an enzyme search is carried out, BRENDA provides a list of all organisms containing the particular enzyme of interest.
- **PUBMED**: This is an online library developed by the National Center for Biotechnology Information, which contains a massive collection of medical journals. Using the link provided by ENZYME, the search can be directed towards the organism of interest, thus recovering literature on the enzyme and its use inside of the organism.

Next steps of the reconstruction

After the initial stages of the reconstruction, a systematic verification is made in order to make sure no inconsistencies are present and that all the entries listed are correct and accurate (Francke *et al.* 2005). Furthermore, previous literature can be researched in order to support any information obtained from one of the many metabolic reaction and genome databases. This provides an added level of assurance for the reconstruction that the enzyme and the reaction it catalyzes do actually occur in the organism.

Any new reactions not present in the databases need to be added to the reconstruction. The presence or absence of certain reactions of the metabolism will affect the amount of reactants/products that are present for other reactions within the particular pathway. This is because products in one reaction go on to become the reactants for another reaction, i.e. products of one reaction can combine with other proteins or compounds to form new proteins/compounds in the presence of different enzymes or catalysts (Francke *et al.* 2005).

Francke *et al.* (2005) provide an excellent example as to why the verification step of the project needs to be performed in significant detail. During a metabolic network reconstruction of *Lactobacillus plantarum*, the model showed that succinyl-CoA was one of the reactants for a reaction that was a part of the biosynthesis of methionine. However, an understanding of the physiology of the organism would have revealed that due to an incomplete tricarboxylic acid pathway, *Lactobacillus plantarum* does not actually produce succinyl-CoA, and the correct reactant for that part of the reaction was acetyl-CoA.

Therefore, systematic verification of the initial reconstruction will bring to light several inconsistencies that can adversely affect the final interpretation of the reconstruction, which is to accurately comprehend the molecular mechanisms of the organism. Furthermore, the simulation step also ensures that all the reactions present in the reconstruction are properly balanced. To sum up, a reconstruction that is fully accurate can lead to greater insight about understanding the functioning of the organism of interest (Francke *et al.* 2005).

Advantages of a reconstruction

- Several inconsistencies exist between gene, enzyme, and reaction databases and published literature sources regarding the metabolic information of an organism. A reconstruction is a systematic verification and compilation of data from various sources that takes into account all of the discrepancies.
- A reconstruction combines the relevant metabolic and genomic information of an organism.
- A reconstruction also allows for metabolic comparisons to be performed between various species of the same organism as well as between different organisms.

Metabolic network simulation

A metabolic network can be broken down into a stoichiometric matrix where the rows represent the compounds of the reactions, while the columns of the matrix correspond to the reactions themselves. Stoichiometry is a quantitative relationship between substrates of a chemical reaction (Merriam 2002). In order to deduce what the metabolic network suggests, recent research has centered on two approaches; namely extreme pathways and elementary mode analysis (Papin, Stelling, Price, Klamt, Schuster, and Palsson 2004).

Extreme Pathways

Price, Reed, Papin, Wiback and Palsson (2003) use a method of singular value decomposition (SVD) of extreme pathways in order to understand regulation of a human red blood cell metabolism. Extreme pathways are convex basis vectors that consist of steady state functions of a metabolic network (Papin, Price, and Palsson 2002). For any particular metabolic network, there is always a unique set of extreme pathways available (Papin *et al.* 2004). Furthermore, Price *et al.* (2003) define a constraint-based approach, where through the help of constraints like mass balance and maximum reaction rates, it is possible to develop a 'solution space' where all the feasible options fall within. Then, using a kinetic model approach, a single solution that falls within the extreme pathway solution space can be determined (Price *et al.* 2003). Therefore, in their study, Price *et al.* (2003) use both constraint and kinetic approaches to understand the human red blood cell metabolism. In conclusion, using extreme pathways, the regulatory mechanisms of a metabolic network can be studied in further detail.

Elementary mode analysis

Elementary mode analysis closely matches the approach used by extreme pathways. Similar to extreme pathways, there is always a unique set of elementary modes available for a particular metabolic network (Papin *et al.* 2004). These are the smallest sub-networks that allow a metabolic reconstruction network to function in steady state (Schuster, Fell, and Dandekar 2000; Stelling, Klamt, Bettenbrock, Schuster, and Gilles 2002). According to Shelling *et al.* (2002), elementary modes can be used to understand cellular objectives for the overall metabolic network. Furthermore, elementary mode analysis takes into account stoichiometrics and thermodynamics when evaluating whether a particular metabolic route or network is feasible and likely for a set of proteins/enzymes (Schuster *et al.* 2000).

Minimal metabolic behaviors (MMBs)

Recently, Larhlimi and Bockmayr (2008) presented a new approach called "minimal metabolic behaviors" for the analysis of metabolic networks. Like elementary modes or extreme pathways, these are uniquely determined by the network, and yield a complete description of the flux cone. However, the new description is much more compact. In contrast with elementary modes and extreme pathways, which use an inner description based on generating vectors of the flux cone, MMBs are using an outer description of the flux cone. This approach is based on sets of non-negativity constraints. These can be identified with irreversible reactions, and thus have a direct biochemical interpretation. One can characterize a metabolic network by MMBs and the reversible metabolic space.

Flux balance analysis

A different technique to simulate the metabolic network is to perform flux balance analysis. This method uses linear programming, but in contrast to elementary mode analysis and extreme pathways, only a single solution results in the end. Linear programming is usually used to obtain the maximum potential of the objective function that you are looking at, and therefore, when using flux balance analysis, a single solution is found to the optimization problem (Stelling *et al.* 2002). In a flux balance analysis approach, exchange fluxes are assigned to those metabolites that enter or leave the particular network only. Those metabolites that are consumed within the network are not assigned any exchange flux

value. Also, the exchange fluxes along with the enzymes can have constraints ranging from a negative to positive value (ex: -10 to 10).

Furthermore, this particular approach can accurately define if the reaction stoichiometry is in line with predictions by providing fluxes for the balanced reactions. Also, flux balance analysis can highlight the most effective and efficient pathway through the network in order to achieve a particular objective function. In addition, gene knockout studies can be performed using flux balance analysis. The enzyme that correlates to the gene that needs to be removed is giving a constraint value of 0. Then, the reaction that the particular enzyme catalyzes is completely removed from the analysis.

Conclusion

In conclusion, metabolic network reconstruction and simulation can be effectively used to understand how an organism or parasite functions inside of the host cell. For example, if the parasite serves to compromise the immune system by lysing macrophages, then the goal of metabolic reconstruction/simulation would be to determine the metabolites that are essential to the organism's proliferation inside of macrophages. If the proliferation cycle is inhibited, then the parasite would not continue to evade the host's immune system. A reconstruction model serves as a first step to deciphering the complicated mechanisms surrounding disease. The next step would be to use the predictions and postulates generated from a reconstruction model and apply it to drug delivery and drug-engineering techniques.

Currently, many tropical diseases affecting third world nations are very inadequately characterized, and thus poorly understood. Therefore, a metabolic reconstruction and simulation of the parasites that cause the tropical diseases would aid in developing new and innovative cures and treatments.

See also

- Metabolic network
- Computer simulation
- Computational systems biology
- Metabolic pathway
- Metagenomics
- Metabolic control analysis

References

1. Francke, C., Siezen, R. J. and Teusink, B. (2005). Reconstructing the metabolic network of a bacterium from its genome. Trends in Microbiology. 13(11): 550-558.
2. Merriam Webster's Medical Dictionary. (2002). http://dictionary.reference.com/medical/
3. Papin, J.A., Price, N.D., and Palsson, B.O. (2002). Extreme Pathway Lengths and Reaction Participation in Genome-Scale Metabolic Networks. Genome Research. 12: 1889-1900.
4. Papin, J.A., Stelling, J., Price, N.D., Klamt, S., Schuster, S., and Palsson, B.O. (2004). Comparison of network-based pathway analysis methods. Trends in Biotechnology. 22(8): 400-405.

5. Price, N.D., Reed, J.L., Papin, J.A., Wiback, S.J., and Palsson, B.O. (2003). Network-based analysis of metabolic regulation in the human red blood cell. Journal of Theoretical Biology. 225: 185-194.
6. Schuster, S., Fell, D.A. and Dandekar, T. (2000). A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nature Biotechnology. 18: 326-332.
7. SRI International. (2005). Pathway Tools Information Site. http://bioinformatics.ai.sri.com/ptools/
8. Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. and Gilles, E.D. (2002). Metabolic network structure determines key aspects of functionality and regulation. Nature. 420: 190-193.
9. Larhlimi, A., Bockmayr, A. (2008) A new constraint-based description of the steady-state flux cone of metabolic networks. Discrete Applied Mathematics. doi:10.1016/j.dam.2008.06.039 [5]

External links

- GeneDB [6]
- KEGG [7]
- PathCase [8] Case Western Reserve University
- BRENDA [9]
- BioCyc [10] and Cyclone [11] - provides an open source Java API to the pathway tool BioCyc to extract Metabolic graphs.
- EcoCyc [12]
- MetaCyc [13]
- ENZYME [14]
- SBRI Bioinformatics Tools and Software [15]
- TIGR [16]
- Pathway Tools [17]
- Stanford Genomic Resources [18]
- Pathway Hunter Tool [19]
- IMG [20] The Integrated Microbial Genomes system, for genome analysis by the DOE-JGI.
- Systems Analysis, Modelling and Prediction Group [21] at the University of Oxford, Biochemical reaction pathway inference techniques.

References

[1] http://www.genome.ad.jp
[2] http://www.genedb.org
[3] http://bioinformatics.ai.sri.com/ptools/
[4] http://ca.expasy.org/
[5] http://dx.doi.org/10.1016%2Fj.dam.2008.06.039
[6] http://www.genedb.org/
[7] http://www.genome.ad.jp/
[8] http://nashua.case.edu/pathwaysweb
[9] http://www.brenda.uni-koeln.de/
[10] http://www.biocyc.org/
[11] http://nemo-cyclone.sourceforge.net
[12] http://ecocyc.org/
[13] http://metacyc.org/
[14] http://www.expasy.org/enzyme/

[15] http://apps.sbri.org/Genome/Link/Bioinformatics_Tools_Software.aspx/
[16] http://www.jcvi.org
[17] http://bioinformatics.ai.sri.com/ptools/
[18] http://genome-www.stanford.edu/
[19] http://pht.tu-bs.de/
[20] http://img.jgi.doe.gov/
[21] http://www.eng.ox.ac.uk/samp

Protein-protein interaction

Protein-protein interactions involve not only the direct-contact association of protein molecules but also longer range interactions through the electrolyte, aqueous solution medium surrounding neighbor hydrated proteins over distances from less than one nanometer to distances of several tens of nanometers. Furthermore, such protein-protein interactions are thermodynamically linked functions[1] of dynamically bound ions and water that exchange rapidly with the surrounding solution by comparison with the molecular tumbling rate (or correlation times) of the interacting proteins. Protein associations are also studied from the perspectives of biochemistry, quantum chemistry, molecular dynamics, signal transduction and other metabolic or genetic/epigenetic networks. Indeed, protein-protein interactions are at the core of the entire Interactomics system of any living cell.

The interactions between proteins are important for very numerous—if not all—biological functions. For example, signals from the exterior of a cell are mediated to the inside of that cell by protein-protein interactions of the signaling molecules. This process, called signal transduction, plays a fundamental role in many biological processes and in many diseases (e.g. cancers). Proteins might interact for a long time to form part of a protein complex, a protein may be carrying another protein (for example, from cytoplasm to nucleus or vice versa in the case of the nuclear pore importins), or a protein may interact briefly with another protein just to modify it (for example, a protein kinase will add a phosphate to a target protein). This modification of proteins can itself change protein-protein interactions. For example, some proteins with SH2 domains only bind to other proteins when they are phosphorylated on the amino acid tyrosine while bromodomains specifically recognise acetylated lysines. In conclusion, protein-protein interactions are of central importance for virtually every process in a living cell. Information about these interactions improves our understanding of diseases and can provide the basis for new therapeutic approaches.

Methods to investigate protein-protein interactions

Biochemical methods

As protein-protein interactions are so important there are a multitude of methods to detect them. Each of the approaches has its own strengths and weaknesses, especially with regard to the sensitivity and specificity of the method. A high sensitivity means that many of the interactions that occur in reality are detected by the screen. A high specificity indicates that most of the interactions detected by the screen are also occurring in reality.

- Co-immunoprecipitation is considered to be the gold standard assay for protein-protein interactions, especially when it is performed with endogenous (not overexpressed and

not tagged) proteins. The protein of interest is isolated with a specific antibody. Interaction partners which stick to this protein are subsequently identified by western blotting. Interactions detected by this approach are considered to be real. However, this method can only verify interactions between suspected interaction partners. Thus, it is not a screening approach. A note of caution also is that immunoprecipitation experiments reveal direct and indirect interactions. Thus, positive results may indicate that two proteins interact directly or may interact via a bridging protein.

- Bimolecular Fluorescence Complementation (BiFC) is a new technique in observing the interactions of proteins. Combining with other new techniques, this method can be used to screen protein-protein interactions and their modulators [2] .
- Affinity electrophoresis as used for estimation of binding constants, as for instance in lectin affinity electrophoresis or characterization of molecules with specific features like glycan content or ligand binding.
- Pull-down assays are a common variation of immunoprecipitation and immunoelectrophoresis and are used identically, although this approach is more amenable to an initial screen for interacting proteins.
- Label transfer can be used for screening or confirmation of protein interactions and can provide information about the interface where the interaction takes place. Label transfer can also detect weak or transient interactions that are difficult to capture using other *in vitro* detection strategies. In a label transfer reaction, a known protein is tagged with a detectable label. The label is then passed to an interacting protein, which can then be identified by the presence of the label.
- The yeast two-hybrid screen investigates the interaction between artificial fusion proteins inside the nucleus of yeast. This approach can identify binding partners of a protein in an unbiased manner. However, the method has a notorious high false-positive rate which makes it necessary to verify the identified interactions by co-immunoprecipitation.
- *In-vivo* crosslinking of protein complexes using photo-reactive amino acid analogs was introduced in 2005 by researchers from the Max Planck Institute [3] In this method, cells are grown with photoreactive diazirine analogs to leucine and methionine, which are incorporated into proteins. Upon exposure to ultraviolet light, the diazirines are activated and bind to interacting proteins that are within a few angstroms of the photo-reactive amino acid analog.
- Tandem affinity purification (TAP) method allows high throughput identification of protein interactions. In contrast to Y2H approach accuracy of the method can be compared to those of small-scale experiments (Collins et al., 2007) and the interactions are detected within the correct cellular environment as by co-immunoprecipitation. However, the TAP tag method requires two successive steps of protein purification and consequently it can not readily detect transient protein-protein interactions. Recent genome-wide TAP experiments were performed by Krogan et al., 2006 and Gavin et al., 2006 providing updated protein interaction data for yeast organism.
- Chemical crosslinking is often used to "fix" protein interactions in place before trying to isolate/identify interacting proteins. Common crosslinkers for this application include the non-cleavable NHS-ester crosslinker, *bis*-sulfosuccinimidyl suberate (BS3); a cleavable version of BS3, dithiobis(sulfosuccinimidyl propionate) (DTSSP); and the imidoester crosslinker dimethyl dithiobispropionimidate (DTBP) that is popular for fixing interactions in ChIP assays.

- Chemical crosslinking followed by high mass MALDI mass spectrometry can be used to analyze intact protein interactions in place before trying to isolate/identify interacting proteins. This method detects interactions among non-tagged proteins and is available from CovalX.
- SPINE (Strep-protein interaction experiment) [4] uses a combination of reversible crosslinking with formaldehyde and an incorporation of an affinity tag to detect interaction partners *in vivo*.
- Quantitative immunoprecipitation combined with knock-down (QUICK) relies on co-immunoprecipitation, quantitative mass spectrometry (SILAC) and RNA interference (RNAi). This method detects interactions among endogenous non-tagged proteins[5] . Thus, it has the same high confidence as co-immunoprecipitation. However, this method also depends on the availability of suitable antibodies.

Physical/Biophysical and Theoretical methods

- Dual Polarisation Interferometry (DPI) can be used to measure protein-protein interactions. DPI provides real-time, high-resolution measurements of molecular size, density and mass. While tagging is not necessary, one of the protein species must be immobilized on the surface of a waveguide.
- Static Light scattering (SLS) measures changes in the Rayleigh scattering of protein complexes in solution and can non-destructively characterize both weak and strong interactions without tagging or immobilization of the protein. The measurement consists of mixing a series of aliquots of different concentrations or compositions with the anylate, measuring the effect of the changes in light scattering as a result of the interaction, and fitting the correlated light scattering changes with concentration to a model. Weak, non-specific interactions are typically characterized via the second virial coefficient. This type of analysis can determine the equilibrium association constant for associated complexes.[6] . Additional light scattering methods for protein activity determination were previously developed by Timasheff. More recent Dynamic Light scattering (DLS) methods for proteins were reported by H. Chou that are also applicable at high protein concentrations and in protein gels; DLS may thus also be applicable for *in vivo* cytoplasmic observations of various protein-protein interactions.
- Surface plasmon resonance can be used to measure protein-protein interaction.
- With Fluorescence correlation spectroscopy, one protein is labeled with a fluorescent dye and the other is left unlabeled. The two proteins are then mixed and the data outputs the fraction of the labeled protein that is unbound and bound to the other protein, allowing you to get a measure of K_D and binding affinity. You can also take time-course measurements to characterize binding kinetics. FCS also tells you the size of the formed complexes so you can measure the stoichiometry of binding. A more powerful methods is [[fluorescence cross-correlation spectroscopy (FCCS) that employs double labeling techniques and cross-correlation resulting in vastly improved signal-to-noise ratios over FCS. Furthermore, the two-photon and three-photon excitation practically eliminates photobleaching effects and provide ultra-fast recording of FCCS or FCS data.
- Fluorescence resonance energy transfer (FRET) is a common technique when observing the interactions of only two different proteins[7] .
- Protein activity determination by NMR multi-nuclear relaxation measurements, or 2D-FT NMR spectroscopy in solutions, combined with nonlinear regression analysis of NMR relaxation or 2D-FT spectroscopy data sets. Whereas the concept of water activity is

widely known and utilized in the applied biosciences, its complement--the protein activity which quantitates protein-protein interactions-- is much less familiar to bioscientists as it is more difficult to determine in dilute solutions of proteins; protein activity is also much harder to determine for concentrated protein solutions when protein aggregation, not merely transient protein association, is often the dominant process[8] .

- Theoretical modeling of protein-protein interactions involves a detailed physical chemistry/thermodynamic understanding of several effects involved, such as intermolecular forces, ion-binding, proton fluctuations and proton exchange. The theory of thermodynamically linked functions is one such example in which ion-binding and protein-protein interactions are treated as linked processes; this treatment is especially important for proteins that have enzymatic activity which depends on cofactor ions dynamically bound at the enzyme active site, as for example, in the case of oxygen-evolving enzyme system (OES) in photosythetic biosystems where the oxygen molecule binding is linked to the chloride anion binding as well as the linked state transition of the manganese ions present at the active site in Photosystem II(PSII). Another example of thermodynamically linked functions of ions and protein activity is that of divalent calcium and magnesium cations to myosin in mechanical energy transduction in muscle. Last-but-not least, chloride ion and oxygen binding to hemoglobin (from several mammalian sources, including human) is a very well-known example of such thermodynamically linked functions for which a detailed and precise theory has been already developed.
- Molecular dynamics (MD) computations of protein-protein interactions.
- Protein-protein docking, the prediction of protein-protein interactions based only on the three-dimensional protein structures from X-ray diffraction of protein crystals might not be satisfactory.[9] [10]

Network visualization of protein-protein interactions

Visualization of protein-protein interaction networks is a popular application of scientific visualization techniques. Although protein interaction diagrams are common in textbooks, diagrams of whole cell protein interaction networks were not as common since the level of complexity made them difficult to generate. One example of a manually produced molecular interaction map is Kurt Kohn's 1999 map of cell cycle control.[11] Drawing on Kohn's map, in 2000 Schwikowski, Uetz, and Fields published a paper on protein-protein interactions in yeast, linking together 1,548 interacting proteins determined by two-hybrid testing. They used a force-directed (Sugiyama) graph drawing algorithm to automatically generate an image of their network.[12] [13] [14] .

An experimental view of Kurt Kohn's 1999 map gmap [15]. Image was merged via gimp 2.2.17 and then uploaded to maplib.net

See also

- Interactomics
- Signal transduction
- Biophysical techniques
- Biochemistry methods
- Genomics
- Complex systems biology
- Complex systems
- Immunoprecipitation
- Protein-protein interaction prediction
- Protein-protein interaction screening
- BioGRID, a public repository for protein and genetic interactions
- Database of Interacting Proteins (DIP)
- NCIBI National Center for Integrative Biomedical Informatics
- Biotechnology
- Protein nuclear magnetic resonance spectroscopy
- 2D-FT NMRI and Spectroscopy
- Fluorescence correlation spectroscopy
- Fluorescence cross-correlation spectroscopy
- Light scattering
- ConsensusPathDB

References

[1] Kinetic Linked-Function Analysis of the Multiligand Interactions on Mg2+-Activated Yeast Pyruvate Kinase. Thomas J. Bollenbach and Thomas Nowak., *Biochemistry*, 2001, 40 (43), pp. 13097–13106

[2] Lu JP, Beatty LK, Pinthus JH. (2008). "Dual expression recombinase based (DERB) single vector system for high throughput screening and verification of protein interactions in living cells.". *Nature Precedings* <*http://hdl.handle.net/10101/npre.2008.1550.2*>.

[3] Suchanek, M., Radzikowska, A., and Thiele, C. (2005). "Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells". *Nature Methods* **2**: 261–268. doi: 10.1038/nmeth752 (http://dx.doi.org/10.1038/nmeth752). PMID 15782218.

[4] Herzberg C., Weidinger LA., Dörrbecker B., Hübner S., Stülke J. and Commichau FM. (2007). "SPINE: A method for the rapid detection and analysis of protein-protein interactions *in vivo*". *Proteomics* **7(22)**: 4032–4035. doi: 10.1002/pmic.200700491 (http://dx.doi.org/10.1002/pmic.200700491). PMID 17994626.

[5] Selbach, M., Mann, M. (2006). "Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK)". *Nature Methods* **3**: 981–983. doi: 10.1038/nmeth972 (http://dx.doi.org/10.1038/nmeth972). PMID 17072306.

[6] Arun K. Attri and Allen P. Minton (2005). "Composition gradient static light scattering: A new technique for rapid detection and quantitative characterization of reversible macromolecular hetero-associations in solution". *Analytical Biochemistry* **346**: 132–138. doi: 10.1016/j.ab.2005.08.013 (http://dx.doi.org/10.1016/j.ab.2005.08.013). PMID 16188220.

[7] Gadella TW Jr., *FRET and FLIM techniques*, 33. Imprint: Elsevier, ISBN 978-0-08-054958-3. (2008) 560 pages.

[8] #Baianu, I.C.; Kumosinski, Thomas (August 1993). "NMR Principles and Applications to Protein Structure, Activity and Hydration.,". *Ch.9 in* Physical Chemistry of Food Processes: Advanced Techniques and Applications. (New York: Van Nostrand-Reinhold) **2**: 338–420. ISBN 0-442-00582-2.

[9] Bonvin AM (2006). "Flexible protein-protein docking". *Current Opinion in Structural Biology* **16**: 194–200. doi: 10.1016/j.sbi.2006.02.002 (http://dx.doi.org/10.1016/j.sbi.2006.02.002). PMID 16488145.

[10] Gray JJ (2006). "High-resolution protein-protein docking". *Current Opinion in Structural Biology* **16**: 183–193. doi: 10.1016/j.sbi.2006.03.003 (http://dx.doi.org/10.1016/j.sbi.2006.03.003). PMID 16546374.

[11] Kurt W. Kohn (1999). " Molecular Interaction Map of the Mammalian Cell Cycle Control and DNA Repair Systems (http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=10436023)". *Molecular Biology of the Cell* **10** (8): 2703–2734. PMID 10436023. .

[12] Benno Schwikowski1, Peter Uetz, and Stanley Fields (2000). " A network of protein−protein interactions in yeast (http://igtmv1.fzk.de/www/itg/uetz/publications/Schwikowski2000.pdf)". *Nature Biotechnology* **18**: 1257–1261. doi: 10.1038/82360 (http://dx.doi.org/10.1038/82360). PMID 11101803. .
[13] Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol. 17:1030-2.
[14] Prieto C, De Las Rivas J (2006). APID: Agile Protein Interaction DataAnalyzer. Nucleic Acids Res. 34:W298-302.
[15] http://www.maplib.net/map.php?id=1700&lat=-52.67138590320257&lng=34.3817138671875&z=9

Further reading

1. Gadella TW Jr., *FRET and FLIM techniques*, 33. Imprint: Elsevier, ISBN 978-0-08-054958-3. (2008) 560 pages
2. Langel FD, et al., Multiple protein domains mediate interaction between Bcl10 and Malt1, *J. Biol. Chem.*, (2008) 283(47):32419-31
3. Clayton AH. , The polarized AB plot for the frequency-domain analysis and representation of fluorophore rotation and resonance energy homotransfer. J *Microscopy*. (2008) 232(2):306-12
4. Clayton AH, et al., Predominance of activated EGFR higher-order oligomers on the cell surface. *Growth Factors* (2008) 20:1
5. Plowman et al., Electrostatic Interactions Positively Regulate K-Ras Nanocluster Formation and Function. *Molecular and Cellular Biology* (2008) 4377–4385
6. Belanis L, et al., Galectin-1 Is a Novel Structural Component and a Major Regulator of H-Ras Nanoclusters. *Molecular Biology of the Cell* (2008) 19:1404–1414
7. Van Manen HJ, Refractive index sensing of green fluorescent proteins in living cells using fluorescence lifetime imaging microscopy. Biophys J. (2008) 94(8):L67-9
8. Van der Krogt GNM, et al., A Comparison of Donor-Acceptor Pairs for Genetically Encoded FRET Sensors: Application to the Epac cAMP Sensor as an Example, PLoS ONE, (2008) 3(4):e1916
9. Dai X, et al., Fluorescence intensity and lifetime imaging of free and micellar-encapsulated doxorubicin in living cells. *Nanomedicine*. (2008) **4**(1):49-56.
10. Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, BioScience (Ed. Klinge & Owman) p.180.
11. Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In Oil Extraction and Analysis., D. Luthria, Editor pp.241–273, AOCS Press., Champaign, IL
12. Richard R. Ernst. 1992. Nuclear Magnetic Resonance Fourier Transform (2D-FT) Spectroscopy.Nobel Lecture, on December 9, 1992.
13. Baianu, I.C.; Kumosinski, Thomas (August 1993). "NMR Principles and Applications to Protein Structure, Activity and Hydration.,". *Ch.9 in* Physical Chemistry of Food Processes: Advanced Techniques and Applications. (New York: Van Nostrand-Reinhold) **2**: 338–420. ISBN 0-442-00582-2.
14. Kurt Wüthrich in 1982-1986 : 2D-FT NMR of solutions (http://en.wikipedia.org/wiki/Nuclear_magnetic_resonance#Nuclear_spin_and_magnets)
15. Charles P. Slichter.1996. *Principles of Magnetic Resonance.*, Springer: Berlin and New York, Third Edition., 651pp. ISBN 0-387-50157-6.
16. Kurt Wüthrich. *Protein structure determination in solution by NMR spectroscopy* . J Biol Chem. 1990, December 25;265(36):22059-62.

External links

- National Center for Integrative Biomedical Informatics (NCIBI) (http://portal.ncibi.org/gateway/)
- Proteins and Enzymes (http://www.dmoz.org/Science/Biology/Biochemistry_and_Molecular_Biology/Biomolecules/Proteins_and_Enzymes/) at the Open Directory Project
- FLIM Applications (http://www.nikoninstruments.com/infocenter.php?n=FLIM) FLIM is also often used in microspectroscopic/ chemical imaging, or microscopic, studies to monitor spatial and temporal protein-protein interactions, properties of membranes and interactions with nucleic acids in living cells.
- Arabidopsis thaliana protein interaction network (http://bioinfo.esalq.usp.br/atpin)

Proteomics

Proteomics is the large-scale study of proteins, particularly their structures and functions.[1] [2] Proteins are vital parts of living organisms, as they are the main components of the physiological metabolic pathways of cells. The term "proteomics" was first coined in 1997[3] to make an analogy with genomics, the study of the genes. The word "proteome" is a blend of "**prote**in" and "gen**ome**", and was coined by Prof Marc Wilkins in 1994 while working on the concept as a PhD student.[4] [5] The proteome is the entire complement of proteins,[4] including the modifications made to a particular set of proteins, produced by an organism or system. This will vary with time and distinct requirements, or stresses, that a cell or organism undergoes.

Robotic preparation of MALDI mass spectrometry samples on a sample carrier.

Complexity of the Problem

After genomics, proteomics is often considered the next step in the study of biological systems. It is much more complicated than genomics mostly because while an organism's genome is more or less constant, the proteome differs from cell to cell and from time to time. This is because distinct genes are expressed in distinct cell types. This means that even the basic set of proteins which are produced in a cell needs to be determined.

In the past this was done by mRNA analysis, but this was found not to correlate with protein content.[6] [7] It is now known that mRNA is not always translated into protein,[8] and the amount of protein produced for a given amount of mRNA depends on the gene it is

transcribed from and on the current physiological state of the cell. Proteomics confirms the presence of the protein and provides a direct measure of the quantity present.

Examples of post-translational modifications

Phosphorylation

More importantly though, any particular protein may go through a wide variety of alterations which will have critical effects to its function. For example during cell signaling many enzymes and structural proteins can undergo phosphorylation. The addition of a phosphate to particular amino acids—most commonly serine and threonine[9] mediated by serine/threonine kinases, or more rarely tyrosine mediated by tyrosine kinases—causes a protein to become a target for binding or interacting with a distinct set of other proteins that recognize the phosphorylated domain.

Because protein phosphorylation is one of the most-studied protein modifications many "proteomic" efforts are geared to determining the set of phosphorylated proteins in a particular cell or tissue-type under particular circumstances. This alerts the scientist to the signaling pathways that may be active in that instance.

Ubiquitination

Ubiquitin is a small protein that can be affixed to certain protein substrates by enzymes called E3 ubiquitin ligases. Determining which proteins are poly-ubiquitinated can be helpful in understanding how protein pathways are regulated. This is therefore an additional legitimate "proteomic" study. Similarly, once it is determined what substrates are ubiquitinated by each ligase, determining the set of ligases expressed in a particular cell type will be helpful.

Additional modifications

Listing all the protein modifications that might be studied in a "Proteomics" project would require a discussion of most of biochemistry; therefore, a short list will serve here to illustrate the complexity of the problem. In addition to phosphorylation and ubiquitination, proteins can be subjected to methylation, acetylation, glycosylation, oxidation, nitrosylation, etc. Some proteins undergo ALL of these modifications, which nicely illustrates the potential complexity one has to deal with when studying protein structure and function.

Distinct proteins are made under distinct settings

Even if one is studying a particular cell type, that cell may make different sets of proteins at different times, or under different conditions. Furthermore, as mentioned, any one protein can undergo a wide range of post-translational modifications.

Therefore a "proteomics" study can become quite complex very quickly, even if the object of the study is very restricted. In more ambitious settings, such as when a biomarker for a tumor is sought - when the proteomics scientist is obliged to study sera samples from multiple cancer patients - the amount of complexity that must be dealt with is as great as in any modern biological project.

Rationale for proteomics

The key requirement in understanding protein function is to learn to correlate the vast array of potential protein modifications to particular phenotypic settings, and then determine if a particular post-translational modification is required for a function to occur.

Limitations to genomic study

Scientists are very interested in proteomics because it gives a much better understanding of an organism than genomics. First, the level of transcription of a gene gives only a rough estimate of its level of expression into a protein. An mRNA produced in abundance may be degraded rapidly or translated inefficiently, resulting in a small amount of protein. Second, as mentioned above many proteins experience post-translational modifications that profoundly affect their activities; for example some proteins are not active until they become phosphorylated. Methods such as phosphoproteomics and glycoproteomics are used to study post-translational modifications. Third, many transcripts give rise to more than one protein, through alternative splicing or alternative post-translational modifications. Fourth, many proteins form complexes with other proteins or RNA molecules, and only function in the presence of these other molecules. Finally, protein degradation rate plays an important role in protein content.[10]

Methods of studying proteins

Determining proteins which are post-translationally modified

One way in which a particular protein can be studied is to develop an antibody which is specific to that modification. For example, there are antibodies which only recognize certain proteins when they are tyrosine-phosphorylated; also, there are antibodies specific to other modifications. These can be used to determine the set of proteins that have undergone the modification of interest.

For sugar modifications, such as glycosylation of proteins, certain lectins have been discovered which bind sugars. These too can be used.

A more common way to determine post-translational modification of interest is to subject a complex mixture of proteins to electrophoresis in "two-dimensions", which simply means that the proteins are electrophoresed first in one direction, and then in another... this allows small differences in a protein to be visualized by separating a modified protein from its unmodified form. This methodology is known as "two-dimensional gel electrophoresis".

Recently, another approach has been developed called PROTOMAP which combines SDS-PAGE with shotgun proteomics to enable detection of changes in gel-migration such as those caused by proteolysis or post translational modification.

Determining the existence of proteins in complex mixtures

Classically, antibodies to particular proteins or to their modified forms have been used in biochemistry and cell biology studies. These are among the most common tools used by practicing biologists today.

For more quantitative determinations of protein amounts, techniques such as ELISAs can be used.

For proteomic study, more recent techniques such as Matrix-assisted laser desorption/ionization have been employed for rapid determination of proteins in particular mixtures.

Establishing protein-protein interactions

Most proteins function in collaboration with other proteins, and one goal of proteomics is to identify which proteins interact. This is especially useful in determining potential partners in cell signaling cascades.

Several methods are available to probe protein-protein interactions. The traditional method is yeast two-hybrid analysis. New methods include protein microarrays, immunoaffinity chromatography followed by mass spectrometry, and experimental methods such as phage display and computational methods.

Practical applications of proteomics

One of the most promising developments to come from the study of human genes and proteins has been the identification of potential new drugs for the treatment of disease. This relies on genome and proteome information to identify proteins associated with a disease, which computer software can then use as targets for new drugs. For example, if a certain protein is implicated in a disease, its 3D structure provides the information to design drugs to interfere with the action of the protein. A molecule that fits the active site of an enzyme, but cannot be released by the enzyme, will inactivate the enzyme. This is the basis of new drug-discovery tools, which aim to find new drugs to inactivate proteins involved in disease. As genetic differences among individuals are found, researchers expect to use these techniques to develop personalized drugs that are more effective for the individual.

A computer technique which attempts to fit millions of small molecules to the three-dimensional structure of a protein is called "virtual ligand screening". The computer rates the quality of the fit to various sites in the protein, with the goal of either enhancing or disabling the function of the protein, depending on its function in the cell. A good example of this is the identification of new drugs to target and inactivate the HIV-1 protease. The HIV-1 protease is an enzyme that cleaves a very large HIV protein into smaller, functional proteins. The virus cannot survive without this enzyme; therefore, it is one of the most effective protein targets for killing HIV.

Biomarkers

Understanding the proteome, the structure and function of each protein and the complexities of protein-protein interactions will be critical for developing the most effective diagnostic techniques and disease treatments in the future.

An interesting use of proteomics is using specific protein biomarkers to diagnose disease. A number of techniques allow to test for proteins produced during a particular disease, which helps to diagnose the disease quickly. Techniques include western blot, immunohistochemical staining, enzyme linked immunosorbent assay (ELISA) or mass spectrometry. The following are some of the diseases that have characteristic biomarkers that physicians can use for diagnosis.

Alzheimer's disease

In Alzheimer’s disease, elevations in beta secretase create amyloid/beta-protein, which causes plaque to build up in the patient's brain, which is thought to play a role in dementia. Targeting this enzyme decreases the amyloid/beta-protein and so slows the progression of the disease. A procedure to test for the increase in amyloid/beta-protein is immunohistochemical staining, in which antibodies bind to specific antigens or biological tissue of amyloid/beta-protein.

Heart disease

Heart disease is commonly assessed using several key protein based biomarkers. Standard protein biomarkers for CVD include interleukin-6, interleukin-8, serum amyloid A protein, fibrinogen, and troponins. cTnI cardiac troponin I increases in concentration within 3 to 12 hours of initial cardiac injury and can be found elevated days after an acute myocardial infarction. A number of commercial antibody based assays as well as other methods are used in hospitals as primary tests for acute MI.

See also

- proteomic chemistry
- bioinformatics
- cytomics
- genomics
- List of omics topics in biology
- metabolomics
- lipidomics
- Shotgun proteomics
- Top-down proteomics
- Bottom-up proteomics
- systems biology
- transcriptomics
- phosphoproteomics
- PEGylation

Protein databases

- UniProt
- Protein Information Resource (PIR)
- Swiss-Prot
- Protein Data Bank (PDB)
- National Center for Biotechnology Information (NCBI)
- Human Protein Reference Database
- Proteopedia The collaborative, 3D encyclopedia of proteins and other molecules.

References

[1] Anderson NL, Anderson NG (1998). "Proteome and proteomics: new technologies, new concepts, and new words". *Electrophoresis* **19** (11): 1853–61. doi: 10.1002/elps.1150191103 (http://dx.doi.org/10.1002/elps.1150191103). PMID 9740045.

[2] Blackstock WP, Weir MP (1999). "Proteomics: quantitative and physical mapping of cellular proteins". *Trends Biotechnol.* **17** (3): 121–7. doi: 10.1016/S0167-7799(98)01245-1 (http://dx.doi.org/10.1016/S0167-7799(98)01245-1). PMID 10189717.

[3] P. James (1997). "Protein identification in the post-genome era: the rapid rise of proteomics.". *Quarterly reviews of biophysics* **30** (4): 279–331. doi: oi:10.1017/S0033583597003399 (http://dx.doi.org/oi:10.1017/S0033583597003399). PMID 9634650.

[4] Marc R. Wilkins, Christian Pasquali, Ron D. Appel, Keli Ou, Olivier Golaz, Jean-Charles Sanchez, Jun X. Yan, Andrew. A. Gooley, Graham Hughes, Ian Humphery-Smith, Keith L. Williams & Denis F. Hochstrasser (1996). "From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Arnino Acid Analysis". *Nature Biotechnology* **14** (1): 61–65. doi: doi:10.1038/nbt0196-61 (http://dx.doi.org/doi:10.1038/nbt0196-61). PMID 9636313.

[5] UNSW Staff Bio: Professor Marc Wilkins (http://www.babs.unsw.edu.au/directory.php?personnelID=12)

[6] Simon Rogers, Mark Girolami, Walter Kolch, Katrina M. Waters, Tao Liu, Brian Thrall and H. Steven Wiley (2008). "Investigating the correspondence between transcriptomic and proteomic expression profiles using coupled cluster models". *Bioinformatics* **24** (24): 2894–2900. doi: 10.1093/bioinformatics/btn553 (http://dx.doi.org/10.1093/bioinformatics/btn553). PMID 18974169.

[7] Vikas Dhingraa, Mukta Gupta, Tracy Andacht and Zhen F. Fu (2005). "New frontiers in proteomics research: A perspective". *International Journal of Pharmaceutics* **299** (1–2): 1–18. doi: 10.1016/j.ijpharm.2005.04.010 (http://dx.doi.org/10.1016/j.ijpharm.2005.04.010). PMID 15979831.

[8] Buckingham, Steven (5 2003). "The major world of microRNAs" (http://www.nature.com/horizon/rna/background/micrornas.html). . Retrieved on 2009-01-14.

[9] Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. (2006). "Global, in vivo, and site-specific phosphorylation dynamics in signaling networks". *Cell* **127**: 635-648. doi: 10.1016/j.cell.2006.09.026 (http://dx.doi.org/10.1016/j.cell.2006.09.026). PMID 17081983.

[10] Archana Belle, Amos Tanay, Ledion Bitincka, Ron Shamir and Erin K. O'Shea (2006). "Quantification of protein half-lives in the budding yeast proteome". *PNAS* **103** (35): 13004–13009. doi: 10.1073/pnas.0605420103 (http://dx.doi.org/10.1073/pnas.0605420103). PMID 16916930.

Bibliography

- Belhajjame, K. et al. Proteome Data Integration: Characteristics and Challenges (http://www.allhands.org.uk/2005/proceedings/papers/525.pdf). Proceedings of the UK e-Science All Hands Meeting, ISBN 1-904425-53-4, September 2005, Nottingham, UK.
- Twyman RM (2004). *Principles Of Proteomics (Advanced Text Series)*. Oxford, UK: BIOS Scientific Publishers. ISBN 1-85996-273-4. (covers almost all branches of proteomics)
- Naven T, Westermeier R (2002). *Proteomics in Practice: A Laboratory Manual of Proteome Analysis*. Weinheim: Wiley-VCH. ISBN 3-527-30354-5. (focused on 2D-gels, good on detail)

- Liebler DC (2002). *Introduction to proteomics: tools for the new biology*. Totowa, NJ: Humana Press. ISBN 0-89603-992-7. ISBN 0-585-41879-9 (electronic, on Netlibrary?), ISBN 0-89603-991-9 hbk
- Wilkins MR, Williams KL, Appel RD, Hochstrasser DF (1997). *Proteome Research: New Frontiers in Functional Genomics (Principles and Practice)*. Berlin: Springer. ISBN 3-540-62753-7.
- Arora PS, Yamagiwa H, Srivastava A, Bolander ME, Sarkar G (2005). " Comparative evaluation of two two-dimensional gel electrophoresis image analysis software applications using synovial fluids from patients with joint disease (http://www.springerlink.com/openurl.asp?genre=article&doi=10.1007/s00776-004-0878-0)". *J Orthop Sci* **10** (2): 160–6. doi: 10.1007/s00776-004-0878-0 (http://dx.doi.org/10.1007/s00776-004-0878-0). PMID 15815863. http://www.springerlink.com/openurl.asp?genre=article&doi=10.1007/s00776-004-0878-0.
- Rediscovering Biology Online Textbook. Unit 2 Proteins and Proteomics. 1997-2006.
- Weaver RF (2005). *Molecular biology* (3rd ed.). New York: McGraw-Hill. pp. 840–9. ISBN 0-07-284611-9.
- Reece J, Campbell N (2002). *Biology* (6th ed.). San Francisco: Benjamin Cummings. pp. 392–3. ISBN 0-8053-6624-5.
- Hye A, Lynham S, Thambisetty M, *et al.* (Nov 2006). "Proteome-based plasma biomarkers for Alzheimer's disease". *Brain* **129** (Pt 11): 3042–50. doi: 10.1093/brain/awl279 (http://dx.doi.org/10.1093/brain/awl279). PMID 17071923.
- Perroud B, Lee J, Valkova N, *et al.* (2006). " Pathway analysis of kidney cancer using proteomics and metabolic profiling (http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1665458)". *Mol Cancer* **5**: 64. doi: 10.1186/1476-4598-5-64 (http://dx.doi.org/10.1186/1476-4598-5-64). PMID 17123452.
- Yohannes E, Chang J, Christ GJ, Davies KP, Chance MR (Jul 2008). "Proteomics analysis identifies molecular targets related to diabetes mellitus-associated bladder dysfunction". *Mol. Cell Proteomics* **7** (7): 1270–85. doi: 10.1074/mcp.M700563-MCP200 (http://dx.doi.org/10.1074/mcp.M700563-MCP200). PMID 18337374..
- Macaulay IC, Carr P, Gusnanto A, Ouwehand WH, Fitzgerald D, Watkins NA (Dec 2005). " Platelet genomics and proteomics in human health and disease (http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1297260)". *J Clin Invest.* **115** (12): 3370–7. doi: 10.1172/JCI26885 (http://dx.doi.org/10.1172/JCI26885). PMID 16322782.
- Rogers MA, Clarke P, Noble J, *et al.* (15 Oct 2003). " Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility (http://cancerres.aacrjournals.org/cgi/pmidlookup?view=long&pmid=14583499)". *Cancer Res.* **63** (20): 6971–83. PMID 14583499. http://cancerres.aacrjournals.org/cgi/pmidlookup?view=long&pmid=14583499.
- Vasan RS (May 2006). "Biomarkers of cardiovascular disease: molecular basis and practical considerations". *Circulation* **113** (19): 2335–62. doi: 10.1161/CIRCULATIONAHA.104.482570 (http://dx.doi.org/10.1161/CIRCULATIONAHA.104.482570). PMID 16702488.
- “Myocardial Infarction” (http://medlib.med.utah.edu/WebPath/TUTORIAL/MYOCARD/MYOCARD.html). (Retrieved 29 November 2006)

- World Community Grid (http://www.worldcommunitygrid.org). (Retrieved 29 November 2006)
- Introduction to Antibodies - Enzyme-Linked Immunosorbent Assay (ELISA) (http://www.chemicon.com/resource/ANT101/a2C.asp). (Retrieved 29 November 2006)
- Decramer S, Wittke S, Mischak H, *et al.* (Apr 2006). " Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis (http://www.nature.com/nm/journal/v12/n4/abs/nm1384.html;jsessionid=3F9707D2B671CA69E12EC68E65919D60)". *Nat Med.* **12** (4): 398–400. doi: 10.1038/nm1384 (http://dx.doi.org/10.1038/nm1384). PMID 16550189. http://www.nature.com/nm/journal/v12/n4/abs/nm1384.html;jsessionid=3F9707D2B671CA69E12EC68E65919D60.
- Mayer U (Jan 2008). "Protein Information Crawler (PIC): extensive spidering of multiple protein information resources for large protein sets". *Proteomics* **8** (1): 42–4. doi: 10.1002/pmic.200700865 (http://dx.doi.org/10.1002/pmic.200700865). PMID 18095364.

External links

- Proteomics (http://www.dmoz.org//Science/Biology/Biochemistry_and_Molecular_Biology/Biomolecules/Proteins_and_Enzymes/Proteomics/) at the Open Directory Project

Interactomics

Interactomics is a discipline at the intersection of bioinformatics and biology that deals with studying both the interactions and the consequences of those interactions between and among proteins, and other molecules within a cell[1] . The network of all such interactions is called the Interactome. Interactomics thus aims to compare such networks of interactions (i.e., interactomes) between and within species in order to find how the traits of such networks are either preserved or varied. From a mathematical, or mathematical biology viewpoint an interactome network is a graph or a category representing the most important interactions pertinent to the normal physiological functions of a cell or organism.

Interactomics is an example of "top-down" systems biology, which takes an overhead, as well as overall, view of a biosystem or organism. Large sets of genome-wide and proteomic data are collected, and correlations between different molecules are inferred. From the data new hypotheses are formulated about feedbacks between these molecules. These hypotheses can then be tested by new experiments[2] .

Through the study of the interaction of all of the molecules in a cell the field looks to gain a deeper understanding of genome function and evolution than just examining an individual genome in isolation[1] . Interactomics goes beyond cellular proteomics in that it not only attempts to characterize the interaction between proteins, but between all molecules in the cell.

Methods of interactomics

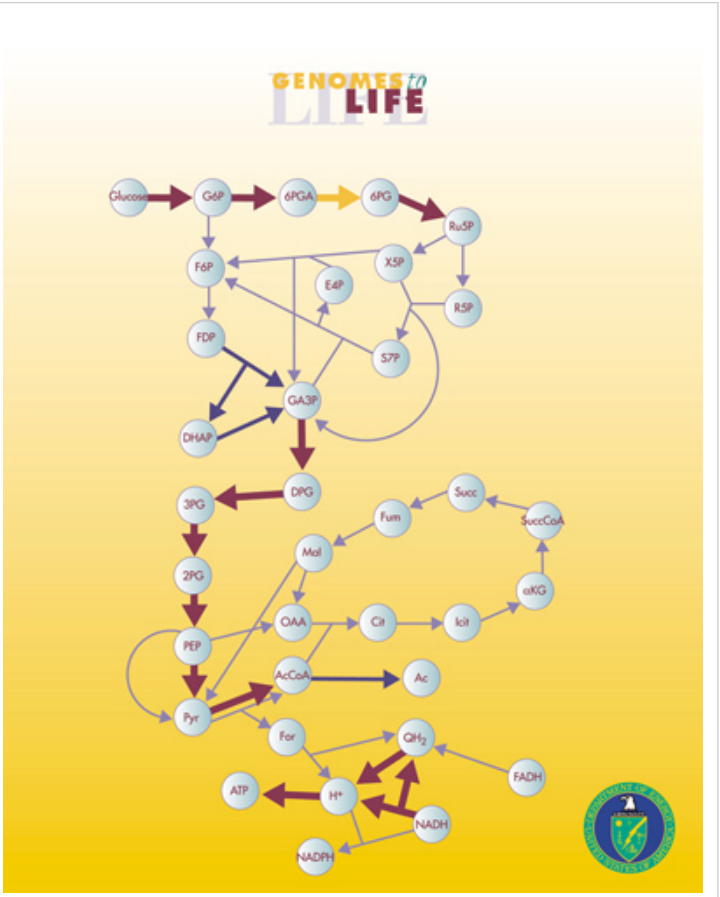
The study of the interactome requires the collection of large amounts of data by way of high throughput experiments. Through these experiments a large number of data points are collected from a single organism under a small number of perturbations[2] These experiments include:

- Two-hybrid screening
- Tandem Affinity Purification
- X-ray tomography
- Optical fluorescence microscopy

Recent developments

The field of interactomics is currently rapidly expanding and developing. While no biological interactomes have been fully characterized. Over 90% of proteins in *Saccharomyces cerevisiae* have been screened and their interactions characterized, making it the first interactome to be nearly fully specified [3] .

Also there have been recent systematic attempts to explore the human interactome[1] and [4] .



Metabolic Network Model for Escherichia coli.

Other species whose interactomes have been studied in some detail include *Caenorhabditis elegans* and *Drosophila melanogaster*.

Criticisms and concerns

Kiemer and Cesareni[1] raise the following concerns with the current state of the field:

- The experimental procedures associated with the field are error prone leading to "noisy results". This leads to 30% of all reported interactions being artifacts. In fact, two groups using the same techniques on the same organism found less than 30% interactions in common.
- Techniques may be biased, i.e. the technique determines which interactions are found.
- Ineractomes are not nearly complete with perhaps the exception of *S. cerivisiae*.
- While genomes are stable, interactomes may vary between tissues and developmental stages.
- Genomics compares amino acids, and nucleotides which are in a sense unchangeable, but interactomics compares proteins and other molecules which are subject to mutation and evolution.
- It is difficult to match evolutionarily related proteins in distantly related species.

See also

- Interaction network
- Proteomics
- Metabolic network
- Metabolic network modelling
- Metabolic pathway
- Genomics
- Mathematical biology
- Systems biology

References

[1] Kiemer, L; G Cesareni (2007). "Comparative interactomics: comparing apples and pears?". *TRENDS in Biochemistry* **25**: 448–454. doi: 10.1016/j.tibtech.2007.08.002 (http://dx.doi.org/10.1016/j.tibtech.2007.08.002).

[2] Bruggeman, F J; H V Westerhoff (2006). "The nature of systems biology". *TRENDS in Microbiology* **15**: 45–50. doi: 10.1016/j.tim.2006.11.003 (http://dx.doi.org/10.1016/j.tim.2006.11.003).

[3] Krogan, NJ; et al. (2006). "Global landscape of protein complexes in the yeast *Saccharomyeses Cerivisiae* ". *Nature* **440**: 637–643. doi: 10.1038/nature04670 (http://dx.doi.org/10.1038/nature04670).

[4] further citation needed

External links

- Interactomics.org (http://interactomics.org). A dedicated interactomics web site operated under BioLicense.
- Interactome.org (http://interactome.org). An interactome wiki site.
- PSIbase (http://psibase.kobic.re.kr) Structural Interactome Map of all Proteins.
- Omics.org (http://omics.org). An omics portal site that is openfree (under BioLicense)
- Genomics.org (http://genomics.org). A Genomics wiki site.
- Comparative Interactomics analysis of protein family interaction networks using PSIMAP (protein structural interactome map) (http://bioinformatics.oxfordjournals.org/cgi/content/full/21/15/3234)
- Interaction interfaces in proteins via the Voronoi diagram of atoms (http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6TYR-4KXVD30-2&_user=10&

_coverDate=11/30/2006&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&
_acct=C000050221&_version=1&_urlVersion=0&_userid=10&
md5=8361bf3fe7834b4642cdda3b979de8bb)
- Using convex hulls to extract interaction interfaces from known structures. Panos Dafas, Dan Bolser, Jacek Gomoluch, Jong Park, and Michael Schroeder. Bioinformatics 2004 20: 1486-1490.
- PSIbase: a database of Protein Structural Interactome map (PSIMAP). Sungsam Gong, Giseok Yoon, Insoo Jang Bioinformatics 2005.
- Mapping Protein Family Interactions : Intramolecular and Intermolecular Protein Family Interaction Repertoires in the PDB and Yeast, Jong Park, Michael Lappe & Sarah A. Teichmann,J.M.B (2001).
- Semantic Systems Biology (http://www.semantic-systems-biology.org)

Mathematical biology

Mathematical biology is also called **theoretical biology**,[1] and sometimes **biomathematics**. It includes at least four major subfields: *biological mathematical modeling*, *relational biology/complex systems biology (CSB)*, *bioinformatics* and *computational biomodeling/biocomputing*. It is an interdisciplinary academic research field with a wide range of applications in biology, medicine[2] and biotechnology.[3]

Mathematical biology aims at the mathematical representation, treatment and modeling of biological processes, using a variety of applied mathematical techniques and tools. It has both theoretical and practical applications in biological, biomedical and biotechnology research. For example, in cell biology, protein interactions are often represented as "cartoon" models, which, although easy to visualize, do not accurately describe the systems studied. In order to do this, precise mathematical models are required. By describing the systems in a quantitative manner, their behavior can be better simulated, and hence properties can be predicted that might not be evident to the experimenter.

Importance

Applying mathematics to biology has a long history, but only recently has there been an explosion of interest in the field. Some reasons for this include:

- the explosion of data-rich information sets, due to the genomics revolution, which are difficult to understand without the use of analytical tools,
- recent development of mathematical tools such as chaos theory to help understand complex, nonlinear mechanisms in biology,
- an increase in computing power which enables calculations and simulations to be performed that were not previously possible, and
- an increasing interest in in silico experimentation due to ethical considerations, risk, unreliability and other complications involved in human and animal research.

For use of basic arithmetics in biology, see relevant topic, such as Serial dilution.

Areas of research

Several areas of specialized research in mathematical and theoretical biology[4] [5] [6] [7] [8] [9] as well as external links to related projects in various universities are concisely presented in the following subsections, including also a large number of appropriate validating references from a list of several thousands of published authors contributing to this field. Many of the included examples are characterised by highly complex, nonlinear, and supercomplex mechanisms, as it is being increasingly recognised that the result of such interactions may only be understood through a combination of mathematical, logical, physical/chemical, molecular and computational models. Due to the wide diversity of specific knowledge involved, biomathematical research is often done in collaboration between mathematicians, biomathematicians, theoretical biologists, physicists, biophysicists, biochemists, bioengineers, engineers, biologists, physiologists, research physicians, biomedical researchers,oncologists, molecular biologists, geneticists, embryologists, zoologists, chemists, etc.

Computer models and automata theory

A monograph on this topic summarizes an extensive amount of published research in this area up to 1987,[10] including subsections in the following areas: computer modeling in biology and medicine, arterial system models, neuron models, biochemical and oscillation networks, quantum automata, [11]quantum computers in molecular biology and genetics, cancer modelling, neural nets, genetic networks, abstract relational biology, metabolic-replication systems, category theory[12] applications in biology and medicine,[13] automata theory,cellular automata, tessallation models[14] [15] and complete self-reproduction [16], chaotic systems in organisms, relational biology and organismic theories.[17] [18] This published report also includes 390 references to peer-reviewed articles by a large number of authors.[19] [20] [21]

Modeling cell and molecular biology

This area has received a boost due to the growing importance of molecular biology.[22]

- Mechanics of biological tissues[23]
- Theoretical enzymology and enzyme kinetics
- Cancer modelling and simulation [24] [25]
- Modelling the movement of interacting cell populations[26]
- Mathematical modelling of scar tissue formation[27]
- Mathematical modelling of intracellular dynamics[28]
- Mathematical modelling of the cell cycle[29]

Modelling physiological systems

- Modelling of arterial disease [30]
- Multi-scale modelling of the heart [31]

Molecular set theory

Molecular set theory was introduced by Anthony Bartholomay, and its applications were developed in mathematical biology and especially in Mathematical Medicine.[32] Molecular set theory (MST) is a mathematical formulation of the wide-sense chemical kinetics of biomolecular reactions in terms of sets of molecules and their chemical transformations represented by set-theoretical mappings between molecular sets. In a more general sense,

MST is the theory of molecular categories defined as categories of molecular sets and their chemical transformations represented as set-theoretical mappings of molecular sets. The theory has also contributed to biostatistics and the formulation of clinical biochemistry problems in mathematical formulations of pathological, biochemical changes of interest to Physiology, Clinical Biochemistry and Medicine.[33] [34]

Population dynamics

Population dynamics has traditionally been the dominant field of mathematical biology. Work in this area dates back to the 19th century. The Lotka–Volterra predator-prey equations are a famous example. In the past 30 years, population dynamics has been complemented by evolutionary game theory, developed first by John Maynard Smith. Under these dynamics, evolutionary biology concepts may take a deterministic mathematical form. Population dynamics overlap with another active area of research in mathematical biology: mathematical epidemiology, the study of infectious disease affecting populations. Various models of viral spread have been proposed and analyzed, and provide important results that may be applied to health policy decisions.

Mathematical methods

A model of a biological system is converted into a system of equations, although the word 'model' is often used synonymously with the system of corresponding equations. The solution of the equations, by either analytical or numerical means, describes how the biological system behaves either over time or at equilibrium. There are many different types of equations and the type of behavior that can occur is dependent on both the model and the equations used. The model often makes assumptions about the system. The equations may also make assumptions about the nature of what may occur.

Mathematical biophysics

The earlier stages of mathematical biology were dominated by mathematical biophysics, described as the application of mathematics in biophysics, often involving specific physical/mathematical models of biosystems and their components or compartments.

The following is a list of mathematical descriptions and their assumptions.

Deterministic processes (dynamical systems)

A fixed mapping between an initial state and a final state. Starting from an initial condition and moving forward in time, a deterministic process will always generate the same trajectory and no two trajectories cross in state space.

- Difference equations – discrete time, continuous state space.
- Ordinary differential equations – continuous time, continuous state space, no spatial derivatives. *See also:* Numerical ordinary differential equations.
- Partial differential equations – continuous time, continuous state space, spatial derivatives. *See also:* Numerical partial differential equations.
- Maps – discrete time, continuous state space.

Stochastic processes (random dynamical systems)

A random mapping between an initial state and a final state, making the state of the system a random variable with a corresponding probability distribution.

- Non-Markovian processes – generalized master equation – continuous time with memory of past events, discrete state space, waiting times of events (or transitions between states) discretely occur and have a generalized probability distribution.
- Jump Markov process – master equation – continuous time with no memory of past events, discrete state space, waiting times between events discretely occur and are exponentially distributed. *See also:* Monte Carlo method for numerical simulation methods, specifically continuous-time Monte Carlo which is also called kinetic Monte Carlo or the stochastic simulation algorithm.
- Continuous Markov process – stochastic differential equations or a Fokker-Planck equation – continuous time, continuous state space, events occur continuously according to a random Wiener process.

Spatial modelling

One classic work in this area is Alan Turing's paper on morphogenesis entitled *The Chemical Basis of Morphogenesis*, published in 1952 in the Philosophical Transactions of the Royal Society.

- Travelling waves in a wound-healing assay[35]
- Swarming behaviour[36]
- A mechanochemical theory of morphogenesis[37]
- Biological pattern formation[38]
- Spatial distribution modeling using plot samples[39]

Phylogenetics

Phylogenetics is an area of mathematical biology that deals with the reconstruction and analysis of phylogenetic (evolutionary) trees and networks based on inherited characteristics. The main mathematical concepts are trees, X-trees and maximum parsimony trees.

Model example: the cell cycle

The eukaryotic cell cycle is very complex and is one of the most studied topics, since its misregulation leads to cancers. It is possibly a good example of a mathematical model as it deals with simple calculus but gives valid results. Two research groups [40] [41] have produced several models of the cell cycle simulating several organisms. They have recently produced a generic eukaryotic cell cycle model which can represent a particular eukaryote depending on the values of the parameters, demonstrating that the idiosyncrasies of the individual cell cycles are due to different protein concentrations and affinities, while the underlying mechanisms are conserved (Csikasz-Nagy et al., 2006).

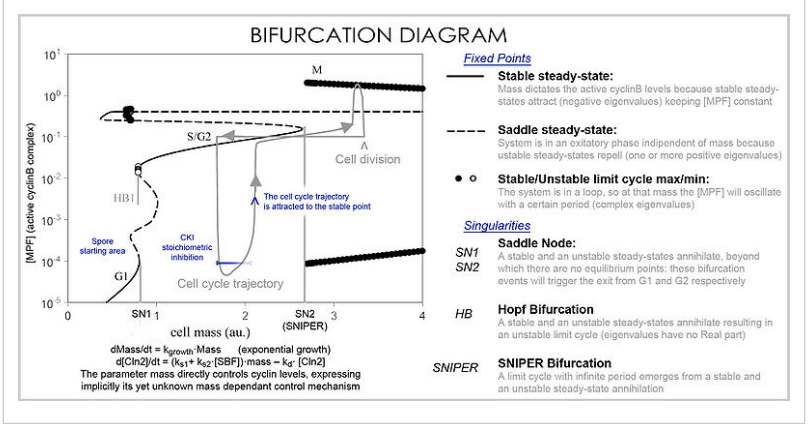
By means of a system of ordinary differential equations these models show the change in time (dynamical system) of the protein inside a single typical cell; this type of model is called a deterministic process (whereas a model describing a statistical distribution of protein concentrations in a population of cells is called a stochastic process).

To obtain these equations an iterative series of steps must be done: first the several models and observations are combined to form a consensus diagram and the appropriate kinetic laws are chosen to write the differential equations, such as rate kinetics for stoichiometric reactions, Michaelis-Menten kinetics for enzyme substrate reactions and Goldbeter–Koshland kinetics for ultrasensitive transcription factors, afterwards the parameters of the equations (rate constants, enzyme efficiency coefficients and Michealis

constants) must be fitted to match observations; when they cannot be fitted the kinetic equation is revised and when that is not possible the wiring diagram is modified. The parameters are fitted and validated using observations of both wild type and mutants, such as protein half-life and cell size.

In order to fit the parameters the differential equations need to be studied. This can be done either by simulation or by analysis.

In a simulation, given a starting vector (list of the values of the variables), the progression of the system is calculated by solving the equations at each time-frame in small increments.



In analysis, the proprieties of the equations are used to investigate the behavior of the system depending of the values of the parameters and variables. A system of differential equations can be represented as a vector field, where each vector described the change (in concentration of two or more protein) determining where and how fast the trajectory (simulation) is heading. Vector fields can have several special points: a stable point, called a sink, that attracts in all directions (forcing the concentrations to be at a certain value), an unstable point, either a source or a saddle point which repels (forcing the concentrations to change away from a certain value), and a limit cycle, a closed trajectory towards which several trajectories spiral towards (making the concentrations oscillate).

A better representation which can handle the large number of variables and parameters is called a bifurcation diagram(Bifurcation theory): the presence of these special steady-state points at certain values of a parameter (e.g. mass) is represented by a point and once the parameter passes a certain value, a qualitative change occurs, called a bifurcation, in which the nature of the space changes, with profound consequences for the protein concentrations: the cell cycle has phases (partially corresponding to G1 and G2) in which mass, via a stable point, controls cyclin levels, and phases (S and M phases) in which the concentrations change independently, but once the phase has changed at a bifurcation event (Cell cycle checkpoint), the system cannot go back to the previous levels since at the current mass the vector field is profoundly different and the mass cannot be reversed back through the bifurcation event, making a checkpoint irreversible. In particular the S and M checkpoints are regulated by means of special bifurcations called a Hopf bifurcation and an infinite period bifurcation.

Mathematical/theoretical biologists

- Pere Alberch
- Anthony F. Bartholomay
- J. T. Bonner
- Jack Cowan
- Gerd B. Müller
- Walter M. Elsasser
- Claus Emmeche
- Andree Ehresmann
- Marc Feldman
- Ronald A. Fisher
- Brian Goodwin
- Bryan Grenfell
- J. B. S. Haldane
- William D. Hamilton
- Lionel G. Harrison
- Michael Hassell
- Sven Erik Jørgensen
- George Karreman
- Stuart Kauffman
- Kalevi Kull
- Herbert D. Landahl
- Richard Lewontin
- Humberto Maturana
- Robert May
- John Maynard Smith
- Howard Pattee
- George R. Price
- Erik Rauch
- Nicolas Rashevsky
- Ronald Brown (mathematician)
- Johannes Reinke
- Robert Rosen
- Rene Thom
- Jakob von Uexküll
- Robert Ulanowicz
- Francisco Varela
- C. H. Waddington
- Arthur Winfree
- Lewis Wolpert
- Sewall Wright
- Christopher Zeeman

Mathematical, theoretical and computational biophysicists

- Nicolas Rashevsky
- Ludwig von Bertalanffy
- Francis Crick
- Manfred Eigen
- Walter Elsasser
- Herbert Frohlich, FRS
- Francois Jacob
- Martin Karplus
- George Karreman
- Herbert D. Landahl
- Ilya, Viscount Prigogine
- SirJohn Randall
- James D. Murray
- Bernard Pullman
- Alberte Pullman
- Erwin Schrodinger
- Klaus Schulten
- Peter Schuster
- Zeno Simon
- D'Arcy Thompson
- Murray Gell-Mann

See also

- Abstract relational biology [42][43] [44]
- Biocybernetics
- Bioinformatics
- Biologically-inspired computing
- Biostatistics
- Cellular automata[45]
- Coalescent theory
- Complex systems biology[46] [47] [48]
- Computational biology
- Dynamical systems in biology[49] [50] [51] [52] [53] [54]
- Epidemiology
- Evolution theories and Population Genetics
  - Population genetics models
  - Molecular evolution theories
- Ewens's sampling formula
- Excitable medium
- Mathematical models
  - Molecular modelling
  - Software for molecular modeling
  - Metabolic-replication systems [55][56]
  - Models of Growth and Form
  - Neighbour-sensing model

- Morphometrics
- Organismic systems (OS) [57][58]
- Organismic supercategories [59][60] [61]
- Population dynamics of fisheries
- Protein folding, also blue Gene and folding@home
- Quantum computers
- Quantum genetics
- Relational biology [62]
- Self-reproduction[63] (also called self-replication in a more general context).
- Computational gene models
- Systems biology[64]
- Theoretical biology[65]
- Topological models of morphogenesis
  - DNA topology
  - DNA sequencing theory

    For use of basic arithmetics in biology, see relevant topic, such as Serial dilution.
- Biographies
  - Charles Darwin
  - D'Arcy Thompson
  - Joseph Fourier
  - Charles S. Peskin
  - Nicolas Rashevsky [66]
  - Robert Rosen
  - Rosalind Franklin
  - Francis Crick
  - René Thom
  - Vito Volterra

References

- Nicolas Rashevsky. (1938)., *Mathematical Biophysics*. Chicago: University of Chicago Press.
- Robert Rosen, Dynamical system theory in biology. New York, Wiley-Interscience (1970) ISBN 0471735507 [67]
- Israel, G., 2005, "Book on mathematical biology" in Grattan-Guinness, I., ed., *Landmark Writings in Western Mathematics*. Elsevier: 936-44.
- Israel, G (1988), "On the contribution of Volterra and Lotka to the development of modern biomathematics. [68]", *History and philosophy of the life sciences* **10** (1): 37-49, PMID:3045853, http://www.ncbi.nlm.nih.gov/pubmed/3045853
- Scudo, F M (1971), "Vito Volterra and theoretical ecology. [69]", *Theoretical population biology* **2** (1): 1-23, 1971 Mar, PMID:4950157, http://www.ncbi.nlm.nih.gov/pubmed/4950157
- S.H. Strogatz, *Nonlinear dynamics and Chaos: Applications to Physics, Biology, Chemistry, and Engineering.* Perseus, 2001, ISBN 0-7382-0453-6
- N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland., 3rd ed. 2001, ISBN 0-444-89349-0

- I. C. Baianu., Computer Models and Automata Theory in Biology and Medicine., *Monograph*, Ch.11 in M. Witten (Editor), *Mathematical Models in Medicine*, vol. **7**., Vol. **7**: 1513-1577 (1987),Pergamon Press:New York, (updated by Hsiao Chen Lin in 2004[70] ,[71] ,[72] ISBN 0080363776 [73] .
- P.G. Drazin, *Nonlinear systems*. C.U.P., 1992. ISBN 0-521-40668-4
- L. Edelstein-Keshet, *Mathematical Models in Biology*. SIAM, 2004. ISBN 0-07-554950-6
- G. Forgacs and S. A. Newman, *Biological Physics of the Developing Embryo*. C.U.P., 2005. ISBN 0-521-78337-2
- A. Goldbeter, *Biochemical oscillations and cellular rhythms*. C.U.P., 1996. ISBN 0-521-59946-6
- L.G. Harrison, *Kinetic theory of living pattern*. C.U.P., 1993. ISBN 0-521-30691-4
- F. Hoppensteadt, *Mathematical theories of populations: demographics, genetics and epidemics*. SIAM, Philadelphia, 1975 (reprinted 1993). ISBN 0-89871-017-0
- D.W. Jordan and P. Smith, *Nonlinear ordinary differential equations*, 2nd ed. O.U.P., 1987. ISBN 0-19-856562-3
- J.D. Murray, *Mathematical Biology*. Springer-Verlag, 3rd ed. in 2 vols.: *Mathematical Biology: I. An Introduction*, 2002 ISBN 0-387-95223-3; *Mathematical Biology: II. Spatial Models and Biomedical Applications*, 2003 ISBN 0-387-95228-4.
- E. Renshaw, *Modelling biological populations in space and time*. C.U.P., 1991. ISBN 0-521-44855-7
- S.I. Rubinow, *Introduction to mathematical biology*. John Wiley, 1975. ISBN 0-471-74446-8
- L.A. Segel, *Modeling dynamic phenomena in molecular and cellular biology*. C.U.P., 1984. ISBN 0-521-27477-X
- L. Preziosi, *Cancer Modelling and Simulation*. Chapman Hall/CRC Press, 2003. ISBN 1-58488-361-8.

Lists of references

- A general list of Theoretical biology/Mathematical biology references, including an updated list of actively contributing authors[74] .
- A list of references for applications of category theory in relational biology[75] .
- An updated list of publications of theoretical biologist Robert Rosen[76]

External

- F. Hoppensteadt, *Getting Started in Mathematical Biology* [77]. Notices of American Mathematical Society, Sept. 1995.
- M. C. Reed, *Why Is Mathematical Biology So Hard?* [78] Notices of American Mathematical Society, March, 2004.
- R. M. May, *Uses and Abuses of Mathematics in Biology* [79]. Science, February 6, 2004.
- J. D. Murray, *How the leopard gets its spots?* [80] Scientific American, 258(3): 80-87, 1988.
- S. Schnell, R. Grima, P. K. Maini, *Multiscale Modeling in Biology* [81], American Scientist, Vol 95, pages 134-142, March-April 2007.
- Chen KC et al. Integrative analysis of cell cycle control in budding yeast. Mol Biol Cell. 2004 Aug;15(8):3841-62.

- Csikász-Nagy A et al. Analysis of a generic model of eukaryotic cell-cycle regulation. Biophys J. 2006 Jun 15;90(12):4361-79.
- Fuss H, et al. Mathematical models of cell cycle regulation. Brief Bioinform. 2005 Jun;6(2):163-77.
- Lovrics A et al. *Time scale and dimension analysis of a budding yeast cell cycle model.* [82] BMC Bioinform. 2006 Nov 9;7:494.

Notes: Inline and online

[1] Mathematical Biology and Theoretical Biophysics-An Outline: What is Life? http://planetmath.org/?op=getobj&from=objects&id=10921
[2] http://www.kli.ac.at/theorylab/EditedVol/W/WittenM1987a.html
[3] http://en.scientificcommons.org/1857372
[4] http://www.kli.ac.at/theorylab/index.html
[5] http://www.springerlink.com/content/w2733h7280521632/
[6] http://en.scientificcommons.org/1857371
[7] http://cogprints.org/3687/
[8] "Research in Mathematical Biology" (http://www.maths.gla.ac.uk/research/groups/biology/kal.htm). Maths.gla.ac.uk. . Retrieved on 2008-09-10.
[9] http://acube.org/volume_23/v23-1p11-36.pdf J. R. Junck. Ten Equations that Changed Biology: Mathematics in Problem-Solving Biology Curricula, *Bioscene*, (1997), 1-36
[10] http://en.scientificcommons.org/1857371
[11] http://planetphysics.org/encyclopedia/QuantumAutomaton.html
[12] http://planetphysics.org/encyclopedia/BibliographyForCategoryTheoryAndAlgebraicTopologyApplicationsInTheoreticalPhysics.html
[13] http://planetphysics.org/encyclopedia/BibliographyForMathematicalBiophysicsAndMathematicalMedicine.html
[14] *Modern Cellular Automata* by Kendall Preston and M. J. B. Duff http://books.google.co.uk/books?id=l0_0q_e-u_UC&dq=cellular+automata+and+tessalation&pg=PP1&ots=ciXYCF3AYm&source=citation&sig=CtaUDhisM7MalS7rZfXvp689y-8&hl=en&sa=X&oi=book_result&resnum=12&ct=result
[15] http://mathworld.wolfram.com/DualTessellation.html
[16] http://planetphysics.org/encyclopedia/ETACAxioms.html
[17] Baianu, I. C. 1987, Computer Models and Automata Theory in Biology and Medicine., in M. Witten (ed.),*Mathematical Models in Medicine*, vol. **7**., Ch.11 Pergamon Press, New York, 1513-1577. http://cogprints.org/3687/
[18] http://www.kli.ac.at/theorylab/EditedVol/W/WittenM1987a.html
[19] http://www.springerlink.com/content/w2733h7280521632/
[20] Currently available for download as an updated PDF: http://cogprints.ecs.soton.ac.uk/archive/00003718/01/COMPUTER_SIMULATIONCOMPUTABILITYBIOSYSTEMSrefnew.pdf
[21] http://planetphysics.org/encyclopedia/BibliographyForMathematicalBiophysics.html
[22] "Research in Mathematical Biology" (http://www.maths.gla.ac.uk/research/groups/biology/kal.htm). Maths.gla.ac.uk. . Retrieved on 2008-09-10.
[23] http://www.maths.gla.ac.uk/~rwo/research_areas.htm
[24] http://www.springerlink.com/content/71958358k273622q/
[25] http://calvino.polito.it/~mcrtn/
[26] http://www.ma.hw.ac.uk/~jas/researchinterests/index.html
[27] http://www.ma.hw.ac.uk/~jas/researchinterests/scartissueformation.html
[28] http://www.sbi.uni-rostock.de/dokumente/p_gilles_paper.pdf
[29] http://mpf.biol.vt.edu/Research.html
[30] http://www.maths.gla.ac.uk/~nah/research_interests.html
[31] http://www.integrativebiology.ox.ac.uk/heartmodel.html
[32] http://planetphysics.org/encyclopedia/CategoryOfMolecularSets2.html
[33] Representation of Uni-molecular and Multimolecular Biochemical Reactions in terms of Molecular Set Transformations http://planetmath.org/?op=getobj&from=objects&id=10770
[34] http://planetphysics.org/encyclopedia/CategoryOfMolecularSets2.html
[35] http://www.maths.ox.ac.uk/~maini/public/gallery/twwha.htm
[36] http://www.math.ubc.ca/people/faculty/keshet/research.html
[37] http://www.maths.ox.ac.uk/~maini/public/gallery/mctom.htm

[38] http://www.maths.ox.ac.uk/~maini/public/gallery/bpf.htm
[39] http://links.jstor.org/sici?sici=0030-1299%28199008%2958%3A3%3C257%3ASDOTMU%3E2.0.CO%3B2-S&size=LARGE&origin=JSTOR-enlargePage
[40] "The JJ Tyson Lab" (http://mpf.biol.vt.edu/Tyson Lab.html). Virginia Tech. . Retrieved on 2008-09-10.
[41] "The Molecular Network Dynamics Research Group" (http://cellcycle.mkt.bme.hu/). Budapest University of Technology and Economics. .
[42] http://www.kli.ac.at/theorylab/ALists/Authors_R.html
[43] http://planetphysics.org/encyclopedia/AbstractRelationalBiologyARB.html
[44] http://www.kli.ac.at/theorylab/EditedVol/M/MatsunoKDose_84.html
[45] Baianu, I. C. 1987, Computer Models and Automata Theory in Biology and Medicine., in M. Witten (ed.),*Mathematical Models in Medicine*, vol. **7**., Ch.11 Pergamon Press, New York, 1513-1577. http://www.springerlink.com/content/w2733h7280521632/
[46] http://www.springerlink.com/content/v1rt05876h74v607/?p=2bd3993c33644512ba7069ed7fad0046&pi=1
[47] http://www.springerlink.com/content/j7t56r530140r88p/?p=2bd3993c33644512ba7069ed7fad0046&pi=3
[48] http://www.springerlink.com/content/98303486x3l07jx3/
[49] Robert Rosen, *Dynamical system theory in biology.* New York, Wiley-Interscience (1970) ISBN 0471735507 http://www.worldcat.org/oclc/101642
[50] http://www.springerlink.com/content/j7t56r530140r88p/?p=2bd3993c33644512ba7069ed7fad0046&pi=3
[51] http://cogprints.org/3674/
[52] http://cogprints.org/3829/
[53] http://www.ncbi.nlm.nih.gov/pubmed/4327361
[54] http://www.springerlink.com/content/98303486x3l07jx3/
[55] http://planetphysics.org/encyclopedia/RSystemsCategoryOfM.html
[56] http://www.kli.ac.at/theorylab/ALists/Authors_R.html
[57] http://planetphysics.org/encyclopedia/OrganismicSetTheory.html
[58] Organisms as Super-complex Systems http://planetmath.org/?op=getobj&from=objects&id=10890
[59] http://planetphysics.org/encyclopedia/OrganismicSetTheory.html
[60] http://www.springerlink.com/content/98303486x3l07jx3/
[61] http://planetmath.org/encyclopedia/SupercategoriesOfComplexSystems.html
[62] http://planetmath.org/?op=getobj&from=objects&id=10921
[63] http://planetmath.org/?method=l2h&from=objects&name=NaturalTransformationsOfOrganismicStructures&op=getobj
[64] http://www.kli.ac.at/theorylab/ALists/Authors_R.html
[65] http://www.kli.ac.at/theorylab/index.html
[66] http://planetphysics.org/encyclopedia/NicolasRashevsky.html
[67] http://www.worldcat.org/oclc/101642
[68] http://www.ncbi.nlm.nih.gov/pubmed/3045853
[69] http://www.ncbi.nlm.nih.gov/pubmed/4950157
[70] http://cogprints.org/3718/1/COMPUTER_SIMULATIONCOMPUTABILITYBIOSYSTEMSrefnew.pdf
[71] http://www.springerlink.com/content/w2733h7280521632/
[72] http://www.springerlink.com/content/n8gw445012267381/
[73] http://www.bookfinder.com/dir/i/Mathematical_Models_in_Medicine/0080363776/
[74] http://www.kli.ac.at/theorylab/index.html
[75] http://planetmath.org/?method=l2h&from=objects&id=10746&op=getobj
[76] Publications list for Robert Rosen http://www.people.vcu.edu/~mikuleck/rosen.htm
[77] http://www.ams.org/notices/199509/hoppensteadt.pdf
[78] http://www.resnet.wm.edu/~jxshix/math490/reed.pdf
[79] http://www.resnet.wm.edu/~jxshix/math490/may.pdf
[80] http://www.resnet.wm.edu/~jxshix/math490/murray.doc
[81] http://eprints.maths.ox.ac.uk/567/01/224.pdf
[82] http://www.biomedcentral.com/content/pdf/1471-2105-7-494.pdf

External links

- Theoretical and mathematical biology website (http://www.kli.ac.at/theorylab/index.html)
- Complexity Discussion Group (http://www.complex.vcu.edu/)
- Integrative cancer biology modeling and Complex systems biology (http://fs512.fshn.uiuc.edu/ComplexSystemsBiology.htm)
- UCLA Biocybernetics Laboratory (http://biocyb.cs.ucla.edu/research.html)
- TUCS Computational Biomodelling Laboratory (http://www.tucs.fi/research/labs/combio.php)
- Nagoya University Division of Biomodeling (http://www.agr.nagoya-u.ac.jp/english/e3senko-1.html)
- Technische Universiteit Biomodeling and Informatics (http://www.bmi2.bmt.tue.nl/Biomedinf/)
- BioCybernetics Wiki, a vertical wiki on biomedical cybernetics and systems biology (http://wiki.biological-cybernetics.de)
- Society for Mathematical Biology (http://www.smb.org/)
- Bulletin of Mathematical Biology (http://www.springerlink.com/content/119979/)
- European Society for Mathematical and Theoretical Biology (http://www.esmtb.org/)
- Journal of Mathematical Biology (http://www.springerlink.com/content/100436/)
- Biomathematics Research Centre at University of Canterbury (http://www.math.canterbury.ac.nz/bio/)
- Centre for Mathematical Biology at Oxford University (http://www.maths.ox.ac.uk/cmb/)
- Mathematical Biology at the National Institute for Medical Research (http://mathbio.nimr.mrc.ac.uk/)
- Institute for Medical BioMathematics (http://www.imbm.org/)
- *Mathematical Biology Systems of Differential Equations* (http://eqworld.ipmnet.ru/en/solutions/syspde/spde-toc2.pdf) from EqWorld: The World of Mathematical Equations
- Systems Biology Workbench - a set of tools for modelling biochemical networks (http://sbw.kgi.edu)
- The Collection of Biostatistics Research Archive (http://www.biostatsresearch.com/repository/)
- Statistical Applications in Genetics and Molecular Biology (http://www.bepress.com/sagmb/)
- The International Journal of Biostatistics (http://www.bepress.com/ijb/)

Theoretical biology

Theoretical biology is a field of academic study and research that involves the use of models and theories in biology.

Many separate areas of biology fall under the concept of theoretical biology, according to the way they are studied. Some of these areas include: animal behaviour (ethology), biomechanics, biorhythms, cell biology, complexity of biological systems, ecology, enzyme kinetics, evolutionary biology, genetics, immunology, membrane transport, microbiology, molecular structures, morphogenesis, physiological mechanisms, systems biology and the origin of life. Neurobiology is an example of a subdiscipline of biology which already has a theoretical version of its own, theoretical or computational neuroscience.

The ultimate goal of the theoretical biologist is to explain the biological world using mainly mathematical and computational tools. Though it is ultimately based on observations and experimental results, the theoretical biologist's product is a model or theory, and it is this that chiefly distinguishes the theoretical biologist from other biologists.

Theoretical biologists

- Pere Alberch
- Anthony F. Bartholomay
- Ervin Bauer
- Ludwig von Bertalanffy
- J. T. Bonner
- Jack Cowan
- Francis Crick
- Gerd B. Müller
- Walter M. Elsasser
- Claus Emmeche
- Andree Ehresmann
- Marc Feldman
- Ronald A. Fisher
- Brian Goodwin
- Bryan Grenfell
- J. B. S. Haldane
- William D. Hamilton
- Lionel G. Harrison
- Michael Hassell
- Sven Erik Jørgensen
- George Karreman
- Stuart Kauffman
- Kalevi Kull
- Herbert D. Landahl
- Richard Lewontin
- Humberto Maturana
- Robert May
- John Maynard Smith
- James D. Murray

- Howard Pattee
- George R. Price
- Erik Rauch
- Nicolas Rashevsky
- Ronald Brown (mathematician)
- Johannes Reinke
- Robert Rosen
- Peter Schuster
- Rene Thom
- D'Arcy Thompson
- Jakob von Uexküll
- Robert Ulanowicz
- Francisco Varela
- C. H. Waddington
- Arthur Winfree
- Lewis Wolpert
- Sewall Wright
- Christopher Zeeman

See also

- Journal of Theoretical Biology
- Bioinformatics
- Biosemiotics
- Mathematical biology
- Theoretical ecology
- Artificial life

Bibliographical references

- Bonner, J. T. 1988. *The Evolution of Complexity by Means of Natural Selection*. Princeton: Princeton University Press.
- Hertel, H. 1963. *Structure, Form, Movement*. New York: Reinhold Publishing Corp.
- Mangel, M. 1990. *Special Issue, Classics of Theoretical Biology* (part 1). Bull. Math. Biol. 52(1/2): 1-318.
- Mangel, M. 2006. *The Theoretical Biologist's Toolbox. Quantitative Methods for Ecology and Evolutionary Biology*. Cambridge University Press.
- Prusinkiewicz, P. & Lindenmeyer, A. 1990. *The Algorithmic Beauty of Plants*. Berlin: Springer-Verlag.
- Reinke, J. 1901. *Einleitung in die theoretische Biologie*. Berlin: Verlag von Gebrüder Paetel.
- Thompson, D.W. 1942. *On Growth and Form*. 2nd ed. Cambridge: Cambridge University Press: 2. vols.
- Uexküll, J.v. 1920. *Theoretische Biologie*. Berlin: Gebr. Paetel.
- Vogel, S. 1988. *Life's Devices: The Physical World of Animals and Plants*. Princeton: Princeton University Press.
- Waddington, C.H. 1968-1972. *Towards a Theoretical Biology*. 4 vols. Edinburg: Edinburg University Press.

External links

- Theory of Biological Anthropology (Documents No. 9 and 10 in English) [1]
- Drawing the Line Between Theoretical and Basic Biology (a forum article by Isidro T. Savillo) [2]

Related Journals

- Acta Biotheoretica [3]
- Bioinformatics [4]
- Biological Theory [5]
- BioSystems [6]
- Bulletin of Mathematical Biology [7]
- Ecological Modelling [8]
- Journal of Mathematical Biology [9]
- Journal of Theoretical Biology [10]
- Journal of the Royal Society Interface [11]
- Mathematical Biosciences [12]
- Medical Hypotheses [13]
- Rivista di Biologia-Biology Forum [14]
- Theoretical and Applied Genetics [15]
- Theoretical Biology and Medical Modelling [16]
- Theoretical Population Biology [17]
- Theory in Biosciences [18] (formerly: Biologisches Zentralblatt)

Related societies

- American Mathematical Society [19]
- British Society of Developmental Biology [20]
- European Mathematical Society [21]
- ESMTB: European Society for Mathematical and Theoretical Biology [22]
- The International Biometric Society [23]
- International Society for Ecological Modelling [24]
- The Israeli Society for Theoretical and Mathematical Biology [25]
- London Mathematical Society [26]
- Société Francophone de Biologie Théorique [27]
- Society for Industrial and Applied Mathematics [28]
- Society for Mathematical Biology [29]
- International Society for Biosemiotic Studies [30]

References

[1] http://homepage.uibk.ac.at/~c720126/humanethologie/ws/medicus/block1/inhalt.html
[2] http://www.scientistsolutions.com/t5844-Drawing+the+line+between+Theoretical+and+Basic+Biology.html
[3] http://www.springerlink.com/link.asp?id=102835
[4] http://bioinformatics.oupjournals.org/
[5] http://www.mitpressjournals.org/loi/biot/
[6] http://www.elsevier.com/locate/biosystems
[7] http://www.springerlink.com/content/119979/
[8] http://www.elsevier.com/locate/issn/03043800
[9] http://www.springerlink.com/content/100436/
[10] http://www.elsevier.com/locate/issn/0022-5193
[11] http://publishing.royalsociety.org/index.cfm?page=1058#
[12] http://www.elsevier.com/locate/mbs
[13] http://www.harcourt-international.com/journals/mehy/
[14] http://www.tilgher.it/biologiae.html
[15] http://www.springerlink.com/content/100386/
[16] http://www.tbiomed.com/
[17] http://www.elsevier.com/locate/issn/00405809
[18] http://www.elsevier.com/wps/product/cws_home/701802
[19] http://www.ams.org/
[20] http://www.dundee.ac.uk/lifesciences/BSDB/
[21] http://www.maths.soton.ac.uk/EMIS
[22] http://www.esmtb.org/
[23] http://www.tibs.org/
[24] http://www.isemna.org/
[25] http://bioinformatics.weizmann.ac.il/istmb/
[26] http://www.lms.ac.uk/
[27] http://www.necker.fr/sfbt/
[28] http://www.siam.org/
[29] http://www.smb.org/
[30] http://www.biosemiotics.org/

Bifurcation theory

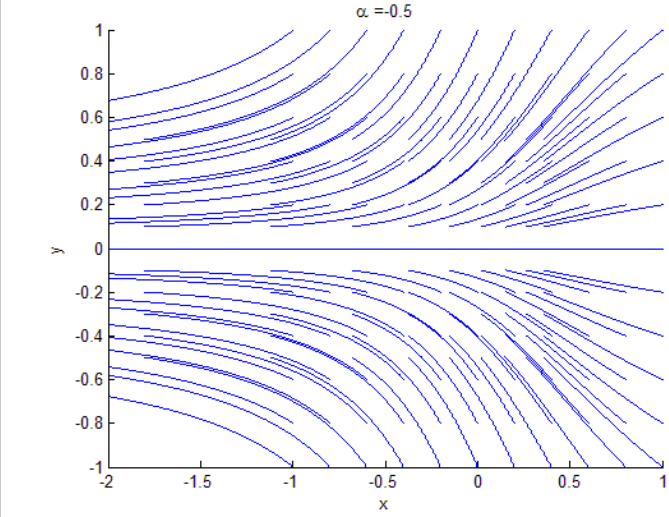
Bifurcation theory is the mathematical study of changes in the qualitative or topological structure of a given family. Examples of such families are the integral curves of a family of vector fields or, the solutions of a family of differential equations. Most commonly applied to the mathematical study of dynamical systems, a **bifurcation** occurs when a small smooth change made to the parameter values (the bifurcation parameters) of a system causes a sudden 'qualitative' or topological change in its behaviour. Bifurcations occur in both continuous systems (described by ODEs, DDEs or PDEs), and discrete systems (described by maps).

Bifurcation Types

It is useful to divide bifurcations into two principal classes:

- Local bifurcations, which can be analysed entirely through changes in the local stability properties of equilibria, periodic orbits or other invariant sets as parameters cross through critical thresholds; and
- Global bifurcations, which often occur when larger invariant sets of the system 'collide' with each other, or with equilibria of the system. They cannot be detected purely by a stability analysis of the equilibria (fixed points).

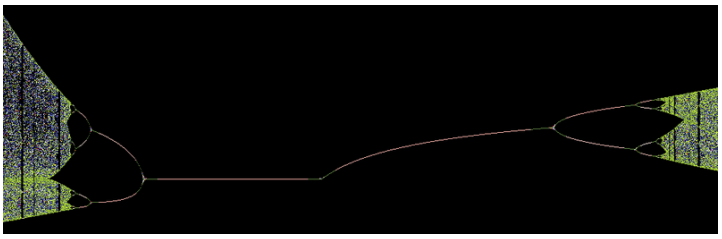
Local bifurcations



Phase portrait showing Saddle-node bifurcation.

A local bifurcation occurs when a parameter change causes the stability of an equilibrium (or fixed point) to change. In continuous systems, this corresponds to the real part of an eigenvalue of an equilibrium passing through zero. In discrete systems (those described by maps rather than ODEs), this corresponds to a fixed point having a Floquet multiplier with modulus equal to one. In both cases, the equilibrium is *non-hyperbolic* at the bifurcation point. The topological changes in the phase portrait of the system can be confined to arbitrarily small neighbourhoods of the bifurcating fixed points by moving the bifurcation parameter close to the bifurcation point (hence 'local').

More technically, consider the continuous dynamical system described by the ODE



Period-halving bifurcations (L) leading to order, followed by period doubling bifurcations (R) leading to chaos.

$$\dot{x} = f(x, \lambda) \quad f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n.$$

A local bifurcation occurs at (x_0, λ_0) if the Jacobian matrix $\mathrm{d}f_{x_0,\lambda_0}$ has an eigenvalue with zero real part. If the eigenvalue is equal to zero, the bifurcation is a steady state bifurcation, but if the eigenvalue is non-zero but purely imaginary, this is a Hopf bifurcation.
For discrete dynamical systems, consider the system

$$x_{n+1} = f(x_n, \lambda)\,.$$

Then a local bifurcation occurs at (x_0, λ_0) if the matrix $\mathrm{d}f_{x_0,\lambda_0}$ has an eigenvalue with modulus equal to one. If the eigenvalue is equal to one, the bifurcation is either a saddle-node (often called fold bifurcation in maps), transcritical or pitchfork bifurcation. If the eigenvalue is equal to -1, it is a period-doubling (or flip) bifurcation, and otherwise, it is a Hopf bifurcation.
Examples of local bifurcations include:

- Saddle-node (fold) bifurcation
- Transcritical bifurcation
- Pitchfork bifurcation
- Period-doubling (flip) bifurcation
- Hopf bifurcation
- Neimark (secondary Hopf) bifurcation

Global bifurcations

Global bifurcations occur when 'larger' invariant sets, such as periodic orbits, collide with equilibria. This causes changes in the topology of the trajectories in the phase space which cannot be confined to a small neighbourhood, as is the case with local bifurcations. In fact, the changes in topology extend out to an arbitrarily large distance (hence 'global').

Examples of global bifurcations include:

- Homoclinic bifurcation in which a limit cycle collides with a saddle point.
- Heteroclinic bifurcation in which a limit cycle collides with two or more saddle points.
- Infinite-period bifurcation in which a stable node and saddle point simultaneously occur on a limit cycle.
- Blue sky catastrophe in which a limit cycle collides with a nonhyperbolic cycle.

Global bifurcations can also involve more complicated sets such as chaotic attractors.

Codimension of a bifurcation

The codimension of a bifurcation is the number of parameters which must be varied for the bifurcation to occur. This corresponds to the codimension of the parameter set for which the bifurcation occurs within the full space of parameters. Saddle-node bifurcations are the only generic local bifurcations which are really codimension-one (the others all having higher codimension). However, often transcritical and pitchfork bifurcations are also often thought of as codimension-one, because the normal forms can be written with only one parameter.

An example of a well-studied codimension-two bifurcation is the Bogdanov-Takens bifurcation.

See also

- Bifurcation diagram
- Catastrophe theory
- Feigenbaum constant
- Phase portrait

References

- Nonlinear dynamics [1]
- Bifurcations and Two Dimensional Flows [2] by Elmer G. Wiens
- Introduction to Bifurcation theory [3] by John David Crawford

References

[1] http://monet.physik.unibas.ch/~elmer/pendulum/nldyn.htm
[2] http://www.egwald.ca/nonlineardynamics/bifurcations.php
[3] http://prola.aps.org/abstract/RMP/v63/i4/p991_1

Catastrophe theory

This article is about the study of dynamical systems. For other meanings, see catastrophe.

In mathematics, **catastrophe theory** is a branch of bifurcation theory in the study of dynamical systems; it is also a particular special case of more general singularity theory in geometry.

Bifurcation theory studies and classifies phenomena characterized by sudden shifts in behavior arising from small changes in circumstances, analysing how the qualitative nature of equation solutions depends on the parameters that appear in the equation. This may lead to sudden and dramatic changes, for example the unpredictable timing and magnitude of a landslide.

Catastrophe theory, which originated with the work of the French mathematician René Thom in the 1960s, and became very popular due to the efforts of Christopher Zeeman in the 1970s, considers the special case where the long-run stable equilibrium can be identified with the minimum of a smooth, well-defined potential function (Lyapunov function).

Small changes in certain parameters of a nonlinear system can cause equilibria to appear or disappear, or to change from attracting to repelling and vice versa, leading to large and sudden changes of the behaviour of the system. However, examined in a larger parameter space, catastrophe theory reveals that such bifurcation points tend to occur as part of well-defined qualitative geometrical structures.

Elementary catastrophes

Catastrophe theory analyses *degenerate critical points* of the potential function — points where not just the first derivative, but one or more higher derivatives of the potential function are also zero. These are called the germs of the catastrophe geometries. The degeneracy of these critical points can be *unfolded* by expanding the potential function as a Taylor series in small perturbations of the parameters.

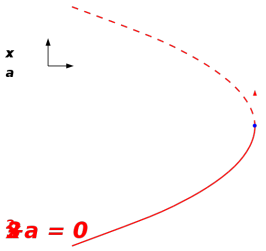
When the degenerate points are not merely accidental, but are structurally stable, the degenerate points exist as organising centres for particular geometric structures of lower degeneracy, with critical features in the parameter space around them. If the potential function depends on two or fewer active variables, and four (resp. five) or fewer active parameters, then there are only seven (resp. eleven) generic structures for these bifurcation geometries, with corresponding standard forms into which the Taylor series around the catastrophe germs can be transformed by diffeomorphism (a smooth transformation whose inverse is also smooth). These seven fundamental types are now presented, with the names that Thom gave them.

Potential functions of one active variable

Fold catastrophe

$$V = x^3 + ax$$

At negative values of *a*, the potential has two extrema - one stable, and one unstable. If the parameter *a* is slowly increased, the system can follow the stable minimum point. But at *a=0* the stable and unstable extrema meet, and annihilate. This is the bifurcation point. At *a>0* there is no longer a stable solution. If a physical system is followed through a fold bifurcation, one therefore finds that as *a* reaches 0, the stability of the *a<0* solution is suddenly lost, and the system will make a sudden transition to a new, very different behaviour. This bifurcation value of the parameter *a* is sometimes called the tipping point.



Stable and unstable pair of extrema disappear at a fold bifurcation

Cusp catastrophe

$$V = x^4 + ax^2 + bx$$

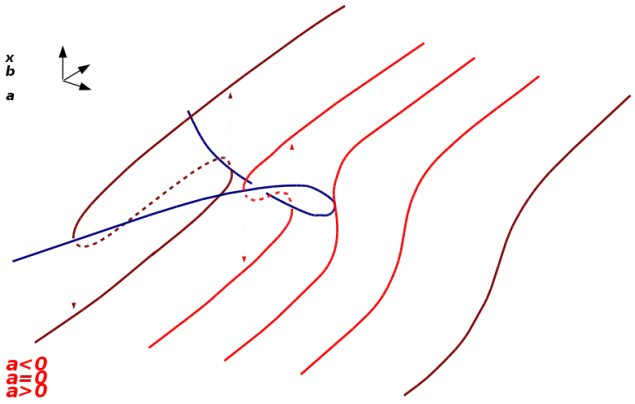
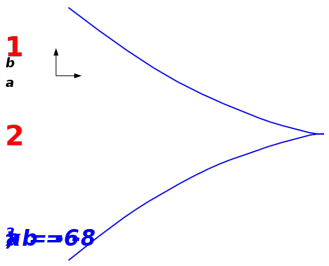
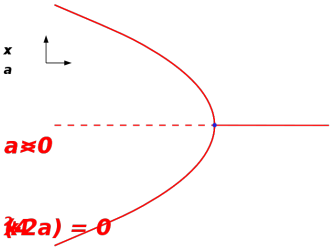


Diagram of cusp catastrophe, showing curves (brown, red) of x satisfying dV / dx = 0 for parameters (a,b), drawn for parameter b continuously varied, for several values of parameter a. Outside the cusp locus of bifurcations (blue), for each point (a,b) in parameter space there is only one extremising value of x. Inside the cusp, there are two different values of x giving local minima of V(x) for each (a,b), separated by a value of x giving a local maximum.



Cusp shape in parameter space *(a,b)* near the catastrophe point, showing the locus of fold bifurcations separating the region with two stable solutions from the region with one.



Pitchfork bifurcation at *a=0* on the surface *b=0*

The cusp geometry is very common, when one explores what happens to a fold bifurcation if a second parameter, *b*, is added to the control space. Varying the parameters, one finds that there is now a *curve* (blue) of points in *(a, b)* space where stability is lost, where the stable solution will suddenly jump to an alternate outcome.

But in a cusp geometry the bifurcation curve loops back on itself, giving a second branch where this alternate solution itself loses stability, and will make a jump back to the original solution set. By repeatedly increasing *b* and then decreasing it, one can therefore observe hysteresis loops, as the system alternately follows one solution, jumps to the other, follows the other back, then jumps back to the first.

However, this is only possible in the region of parameter space $a<0$. As *a* is increased, the hysteresis loops become smaller and smaller, until above $a=0$ they disappear altogether (the cusp catastrophe), and there is only one stable solution.

One can also consider what happens if one holds *b* constant and varies *a*. In the symmetrical case $b=0$, one observes a pitchfork bifurcation as *a* is reduced, with one stable solution suddenly splitting into two stable solutions and one unstable solution as the physical system passes to $a<0$ through the cusp point $a=0$, $b=0$ (an example of spontaneous symmetry breaking). Away from the cusp point, there is no sudden change in a physical solution being followed: when passing through the curve of fold bifurcations, all that happens is an alternate second solution becomes available.

A famous suggestion is that the cusp catastrophe can be used to model the behaviour of a stressed dog, which may respond by becoming cowed or becoming angry. The suggestion is that at moderate stress ($a>0$), the dog will exhibit a smooth transition of response from cowed to angry, depending on how it is provoked. But higher stress levels correspond to moving to the region ($a<0$). Then, if the dog starts cowed, it will remain cowed as it is irritated more and more, until it reaches the 'fold' point, when it will suddenly, discontinuously snap through to angry mode. Once in 'angry' mode, it will remain angry, even if the direct irritation parameter is considerably reduced.

Another application example is for the outer sphere electron transfer frequently encountered in chemical and biological systems (Xu, F. Application of catastrophe theory to the $\Delta G^{\neq}$ to $-\Delta G$ relationship in electron transfer reactions. Zeitschrift für Physikalische Chemie Neue Folge 166, 79-91 (1990)).

Fold bifurcations and the cusp geometry are by far the most important practical consequences of catastrophe theory. They are patterns which reoccur again and again in physics, engineering and mathematical modelling. They are the only way we currently have of detecting black holes and the dark matter of the universe, via the phenomenon of gravitational lensing producing multiple images of distant quasars.

The remaining simple catastrophe geometries are very specialised in comparison, and presented here only for curiosity value.

Swallowtail catastrophe

$$V = x^5 + ax^3 + bx^2 + cx$$

The control parameter space is three dimensional. The bifurcation set in parameter space is made up of three surfaces of fold bifurcations, which meet in two lines of cusp bifurcations, which in turn meet at a single swallowtail bifurcation point.

As the parameters go through the surface of fold bifurcations, one minimum and one maximum of the potential function disappear. At the cusp bifurcations, two minima and one maximum are replaced by one minimum; beyond them the fold bifurcations disappear. At the swallowtail point, two minima and two maxima all meet at a single value of x. For values of $a>0$, beyond the swallowtail, there is either one maximum-minimum pair, or none at all, depending on the values of b and c. Two of the surfaces of fold bifurcations, and the two lines of cusp bifurcations where they meet for $a<0$, therefore disappear at the swallowtail point, to be replaced with only a single surface of fold bifurcations remaining. Salvador Dalí's last painting, *The Swallow's Tail*, was based on this catastrophe.

Butterfly catastrophe

$$V = x^6 + ax^4 + bx^3 + cx^2 + dx$$

Depending on the parameter values, the potential function may have three, two, or one different local minima, separated by the loci of fold bifurcations. At the butterfly point, the different 3-surfaces of fold bifurcations, the 2-surfaces of cusp bifurcations, and the lines of swallowtail bifurcations all meet up and disappear, leaving a single cusp structure remaining when $a>0$

Potential functions of two active variables

Umbilic catastrophes are examples of corank 2 catastrophes. They can be observed in optics in the focal surfaces created by light reflecting off a surface in three dimensions and are intimately connected with the geometry of nearly spherical surfaces. Thom proposed that the Hyperbolic umbilic catastrophe modeled the breaking of a wave and the elliptical umbilic modeled the creation of hair like structures.

Hyperbolic umbilic catastrophe

$$V = x^3 + y^3 + axy + bx + cy$$

Elliptic umbilic catastrophe

$$V = \frac{x^3}{3} - xy^2 + a(x^2 + y^2) + bx + cy$$

Parabolic umbilic catastrophe

$$V = x^2y + y^4 + ax^2 + by^2 + cx + dy$$

Arnold's notation

Vladimir Arnold gave the catastrophes the ADE classification, due to a deep connection with simple Lie groups.

- A_0 - a non-singular point: $V = x$.
- A_1 - a local extrema, either a stable minimum or unstable maximum $V = \pm x^2 + ax$.
- A_2 - the fold
- A_3 - the cusp
- A_4 - the swallowtail
- A_5 - the butterfly
- A_k - an infinite sequence of one variable forms $V = x^{k+1} + \cdots$
- D_4^- - the elliptical umbilic
- D_4^+ - the hyperbolic umbilic
- D_5 - the parabolic umbilic
- D_k - an infinite sequence of further umbilic forms
- E_6 - the symbolic umbilic $V = x^3 + y^4 + axy^2 + bxy + cx + dy + ey^2$
- E_7^6
- E_8

There are objects in singularity theory which correspond to most of the other simple Lie groups.

See also

- broken symmetry
- tipping point
- phase transition
- domino effect
- snowball effect
- butterfly effect
- spontaneous symmetry breaking
- chaos theory

References

- Arnold, Vladimir Igorevich. Catastrophe Theory, 3rd ed. Berlin: Springer-Verlag, 1992.
- Castrigiano, Domenico P. L. and Hayes, Sandra A. Catastrophe Theory, 2nd ed. Boulder: Westview, 2004. ISBN 0-8133-4126-4
- Gilmore, Robert. Catastrophe Theory for Scientists and Engineers. New York: Dover, 1993.
- Petters, Arlie O., Levine, Harold and Wambsganss, Joachim. Singularity Theory and Gravitational Lensing. Boston: Birkhauser, 2001. ISBN 0-8176-3668-4
- Postle, Denis. Catastrophe Theory – Predict and avoid personal disasters. Fontana Paperbacks, 1980. ISBN 0-00-635559-5
- Poston, Tim and Stewart, Ian. Catastrophe: Theory and Its Applications. New York: Dover, 1998. ISBN 0-486-69271-X.
- Sanns, Werner. Catastrophe Theory with Mathematica: A Geometric Approach. Germany: DAV, 2000.
- Saunders, Peter Timothy. An Introduction to Catastrophe Theory. Cambridge, England: Cambridge University Press, 1980.
- Thom, René. Structural Stability and Morphogenesis: An Outline of a General Theory of Models. Reading, MA: Addison-Wesley, 1989. ISBN 0-201-09419-3.
- Thompson, J. Michael T. Instabilities and Catastrophes in Science and Engineering. New York: Wiley, 1982.
- Woodcock, Alexander Edward Richard and Davis, Monte. Catastrophe Theory. New York: E. P. Dutton, 1978. ISBN 0525078126.
- Zeeman, E.C. Catastrophe Theory-Selected Papers 1972–1977. Reading, MA: Addison-Wesley, 1977.

External links

- CompLexicon: Catastrophe Theory [1]
- Catastrophe teacher [2]

References

[1] http://www.exploratorium.edu/complexity/CompLexicon/catastrophe.html
[2] http://perso.wanadoo.fr/l.d.v.dujardin/ct/eng_index.html

Chaos

Chaos (derived from the Ancient Greek Χάος, *Chaos*) typically refers to a state lacking order or predictability. In ancient Greece, it referred to the initial state of the universe, and, by extension, space, darkness, or an abyss.[1] In modern English, it is used in classical studies with this original meaning; in mathematics and science to refer to a very specific kind of unpredictability; and informally to mean a state of confusion.[2]

Chaos in mythology, literature, and religion

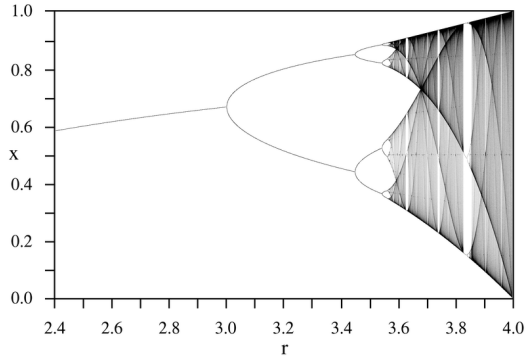
In Greek myth, **Chaos** is the original dark void from which everything else appeared. According to Hesiod's *Theogony* (the origin of the gods), Chaos was the nothingness out of which the first objects of existence appeared. In a similar way, the book of *Genesis* in the Bible refers to the earliest conditions of the Earth as "without form, and void",[3] while Ovid's *Metamorphoses* describes the initial state of the Universe as a disorganised mixture of the four elements:

> Rather a rude and indigested mass:
> A lifeless lump, unfashion'd, and unfram'd,
> Of jarring seeds; and justly **Chaos** nam'd.
> No sun was lighted up, the world to view;
> No moon did yet her blunted horns renew:
> Nor yet was Earth suspended in the sky,
> Nor pois'd, did on her own foundations lye:
> Nor seas about the shores their arms had thrown;
> But earth, and air, and water, were in one.
> Thus air was void of light, and earth unstable,
> And water's dark abyss unnavigable.[4]

Hesiod and the Muse, by Gustave Moreau

Scientific and mathematical chaos

Mathematically, **chaos** means deterministic behaviour which is very sensitive to its initial conditions.[5] In other words, infinitesimal perturbations of initial conditions for a chaotic dynamic system lead to large variations in behaviour.



Bifurcation diagram of a chaotic function

Chaotic systems consequently look random. However, they are actually deterministic systems governed by physical or mathematical laws (predictable in principle, if you have exact information) that are impossible to predict in practice beyond a certain point.[6] A commonly used example is weather forecasting, which is only possible up to about a week ahead.[7]

Edward Lorenz and Henri Poincaré were early pioneers of chaos theory, and James Gleick's 1987 book *Chaos: Making a New Science* helped to popularize the field. A number of philosophers have used the existence of chaos in this sense in arguments about free will.

More recently, computer scientist Christopher Langton in 1990 coined the phrase "edge of chaos" to refer to the behaviour of certain classes of cellular automata.[8] The phrase has since come to refer to a metaphor that some physical, biological, economic, and social systems operate in a region where complexity is maximal, balanced between order, on the one hand, and randomness or chaos, on the other.

Notes

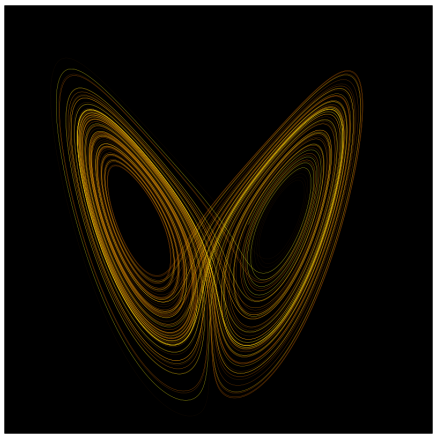
[1] Henry Liddell and Robert Scott, *A Greek-English Lexicon*, online (http://www.perseus.tufts.edu/cgi-bin/ptext?doc=Perseus:text:1999.04.0057:entry=#113326) at the Perseus Project, 2007.
[2] *Merriam-Webster Online Dictionary (http://www.merriam-webster.com/dictionary/chaos)*, accessed 4 April 2009.
[3] *The Holy Bible*, King James Version, online (http://bartleby.com/108/01/1.html).
[4] Ovid, *Metamorphoses*, Book 1, online (http://classics.mit.edu/Ovid/metam.1.first.html)
[5] Saber N. Elaydi, *Discrete Chaos*, Chapman & Hall/CRC, 1999, page 117.
[6] Werndl, Charlotte (2009). *What are the New Implications of Chaos for Unpredictability? (http://bjps.oxfordjournals.org/cgi/content/abstract/60/1/195)*. The British Journal for the Philosophy of Science 60, 195-220.
[7] Robert G. Watts, *Global Warming and the Future of the Earth*, Morgan & Claypool, 2007, page 17.
[8] Christopher G. Langton. "Computation at the edge of chaos". *Physica D*, **42**, 1990.

Chaotic system

1. REDIRECT Chaos theory

Chaos theory

In mathematics, **chaos theory** describes the behavior of certain dynamical systems – that is, systems whose states evolve with time – that may exhibit dynamics that are highly sensitive to initial conditions (popularly referred to as the butterfly effect). As a result of this sensitivity, which manifests itself as an exponential growth of perturbations in the initial conditions, the behavior of chaotic systems appears to be random. This happens even though these systems are deterministic, meaning that their future dynamics are fully defined by their initial conditions with no random elements involved. This behavior is known as deterministic chaos, or simply *chaos*.



A plot of the Lorenz attractor for values $r = 28$, $\sigma = 10$, $b = 8/3$

Chaotic behavior is also observed in natural systems, such as the weather. This may be explained by a chaos-theoretical analysis of a mathematical model of such a system, embodying the laws of physics that are relevant for the natural system.

Overview

Chaotic behavior has been observed in the laboratory in a variety of systems including electrical circuits, lasers, oscillating chemical reactions, fluid dynamics, and mechanical and magneto-mechanical devices. Observations of chaotic behavior in nature include the dynamics of satellites in the solar system, the time evolution of the magnetic field of celestial bodies, population growth in ecology, the dynamics of the action potentials in neurons, and molecular vibrations. Everyday examples of chaotic systems include weather and climate.[1] There is some controversy over the existence of chaotic dynamics in plate tectonics and in economics.[2] [3] [4]

Systems that exhibit mathematical chaos are deterministic and thus orderly in some sense; this technical use of the word *chaos* is at odds with common parlance, which suggests complete disorder. However, even though they are deterministic, chaotic systems show a strong kind of unpredictability not shown by other deterministic systems.[5]

A related field of physics called quantum chaos theory studies systems that follow the laws of quantum mechanics. Recently, another field, called relativistic chaos,[6] has emerged to describe systems that follow the laws of general relativity.

This article tries to describe limits on the degree of disorder that computers can model with simple rules that have complex results. For example, the Lorenz system pictured is chaotic, but has a clearly defined structure. *Bounded chaos* is a useful term for describing models of disorder.

History

The first discoverer of chaos was Henri Poincaré. In the 1880s, while studying the three-body problem, he found that there can be orbits which are nonperiodic, and yet not forever increasing nor approaching a fixed point.[7] [8] In 1898 Jacques Hadamard published an influential study of the chaotic motion of a free particle gliding frictionlessly on a surface of constant negative curvature.[9] In the system studied, "Hadamard's billiards," Hadamard was able to show that all trajectories are unstable in that all particle trajectories diverge exponentially from one another, with a positive Lyapunov exponent.

Fractal fern created using chaos game. Natural forms (ferns, clouds, mountains, etc.) may be recreated through an Iterated function system (IFS).

Much of the earlier theory was developed almost entirely by mathematicians, under the name of ergodic theory. Later studies, also on the topic of nonlinear differential equations, were carried out by G.D. Birkhoff,[10] A. N. Kolmogorov,[11] [12] [13] M.L. Cartwright and J.E. Littlewood,[14] and Stephen Smale.[15] Except for Smale, these studies were all directly inspired by physics: the three-body problem in the case of Birkhoff, turbulence and astronomical problems in the case of Kolmogorov, and radio engineering in the case of Cartwright and Littlewood. Although chaotic planetary motion had not been observed, experimentalists had encountered turbulence in fluid motion and nonperiodic oscillation in radio circuits without the benefit of a theory to explain what they were seeing.

Despite initial insights in the first half of the twentieth century, chaos theory became formalized as such only after mid-century, when it first became evident for some scientists that linear theory, the prevailing system theory at that time, simply could not explain the observed behaviour of certain experiments like that of the logistic map. What had been beforehand excluded as measure imprecision and simple "noise" was considered by chaos theories as a full component of the studied systems.

The main catalyst for the development of chaos theory was the electronic computer. Much of the mathematics of chaos theory involves the repeated iteration of simple mathematical formulas, which would be impractical to do by hand. Electronic computers made these repeated calculations practical, while figures and images made it possible to visualize these systems. One of the earliest electronic digital computers, ENIAC, was used to run simple weather forecasting models.

Turbulence in the tip vortex from an airplane wing. Studies of the critical point beyond which a system creates turbulence was important for Chaos theory, analyzed for example by the Soviet physicist Lev Landau who developed the Landau-Hopf theory of turbulence. David Ruelle and Floris Takens later predicted, against Landau, that fluid turbulence could develop through a strange attractor, a main concept of chaos theory.

An early pioneer of the theory was Edward Lorenz whose interest in chaos came about accidentally through his work on weather prediction in 1961.[16] Lorenz was using a simple digital computer, a Royal McBee LGP-30, to run his weather simulation. He wanted to see a sequence of data again and to save time he started the simulation in the middle of its course. He was able to do this by entering a printout of the data corresponding to conditions in the middle of his simulation which he had calculated last time.

To his surprise the weather that the machine began to predict was completely different from the weather calculated before. Lorenz tracked this down to the computer printout. The computer worked with 6-digit precision, but the printout rounded variables off to a 3-digit number, so a value like 0.506127 was printed as 0.506. This difference is tiny and the consensus at the time would have been that it should have had practically no effect. However Lorenz had discovered that small changes in initial conditions produced large changes in the long-term outcome.[17] Lorenz's discovery, which gave its name to Lorenz attractors, proved that meteorology could not reasonably predict weather beyond a weekly period (at most).

The year before, Benoît Mandelbrot found recurring patterns at every scale in data on cotton prices.[18] Beforehand, he had studied information theory and concluded noise was patterned like a Cantor set: on any scale the proportion of noise-containing periods to error-free periods was a constant – thus errors were inevitable and must be planned for by incorporating redundancy.[19] Mandelbrot described both the "Noah effect" (in which sudden discontinuous changes can occur, e.g., in a stock's prices after bad news, thus challenging normal distribution theory in statistics, aka Bell Curve) and the "Joseph effect" (in which persistence of a value can occur for a while, yet suddenly change afterwards).[20] [21] In 1967, he published "How long is the coast of Britain? Statistical self-similarity and fractional dimension," showing that a coastline's length varies with the scale of the measuring instrument, resembles itself at all scales, and is infinite in length for an infinitesimally small measuring device.[22] Arguing that a ball of twine appears to be a point when viewed from far away (0-dimensional), a ball when viewed from fairly near (3-dimensional), or a curved strand (1-dimensional), he argued that the dimensions of an object are relative to the observer and may be fractional. An object whose irregularity is constant over different scales ("self-similarity") is a fractal (for example, the Koch curve or "snowflake", which is infinitely long yet encloses a finite space and has fractal dimension equal to circa 1.2619, the Menger sponge and the Sierpiński gasket). In 1975 Mandelbrot published *The Fractal Geometry of Nature*, which became a classic of chaos theory. Biological systems such as the branching of the circulatory and bronchial systems proved to fit a fractal model.

Chaos was observed by a number of experimenters before it was recognized; e.g., in 1927 by van der Pol[23] and in 1958 by R.L. Ives.[24] [25] However, Yoshisuke Ueda seems to have been the first experimenter to have recognized chaos as such while using an analog computer on November 27, 1961. Ueda's supervising professor, Hayashi, did not believe in chaos, and thus he prohibited Ueda from publishing his findings until 1970.[26]

In December 1977 the New York Academy of Sciences organized the first symposium on Chaos, attended by David Ruelle, Robert May, James Yorke (coiner of the term "chaos" as used in mathematics), Robert Shaw (a physicist, part of the Eudaemons group with J. Doyne Farmer and Norman Packard who tried to find a mathematical method to beat roulette, and then created with them the Dynamical Systems Collective in Santa Cruz, California), and the meteorologist Edward Lorenz.

The following year, Mitchell Feigenbaum published the noted article "Quantitative Universality for a Class of Nonlinear Transformations", where he described logistic maps.[27] Feigenbaum had applied fractal geometry to the study of natural forms such as coastlines. Feigenbaum notably discovered the universality in chaos, permitting an application of chaos theory to many different phenomena.

In 1979, Albert J. Libchaber, during a symposium organized in Aspen by Pierre Hohenberg, presented his experimental observation of the bifurcation cascade that leads to chaos and turbulence in convective Rayleigh–Benard systems. He was awarded the Wolf Prize in Physics in 1986 along with Mitchell J. Feigenbaum "for his brilliant experimental demonstration of the transition to turbulence and chaos in dynamical systems".[28]

Then in 1986 the New York Academy of Sciences co-organized with the National Institute of Mental Health and the Office of Naval Research the first important conference on Chaos in biology and medicine. Bernardo Huberman thereby presented a mathematical model of the eye tracking disorder among schizophrenics.[29] Chaos theory thereafter renewed physiology in the 1980s, for example in the study of pathological cardiac cycles.

In 1987, Per Bak, Chao Tang and Kurt Wiesenfeld published a paper in *Physical Review Letters*[30] describing for the first time self-organized criticality (SOC), considered to be one of the mechanisms by which complexity arises in nature. Alongside largely lab-based approaches such as the Bak–Tang–Wiesenfeld sandpile, many other investigations have centered around large-scale natural or social systems that are known (or suspected) to display scale-invariant behaviour. 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 behaviour such as the Gutenberg–Richter law describing the statistical distribution of earthquake sizes, and the Omori law[31] 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.

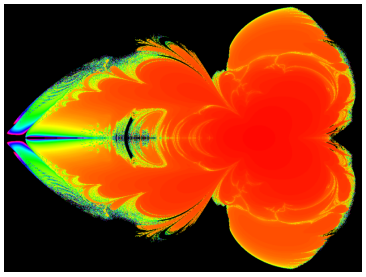
The same year, James Gleick published *Chaos: Making a New Science*, which became a best-seller and introduced general principles of chaos theory as well as its history to the broad public. At first the domains of work of a few, isolated individuals, chaos theory progressively emerged as a transdisciplinary and institutional discipline, mainly under the name of nonlinear systems analysis. Alluding to Thomas Kuhn's concept of a paradigm shift exposed in *The Structure of Scientific Revolutions* (1962), many "chaologists" (as some self-nominated themselves) claimed that this new theory was an example of such as shift, a thesis upheld by J. Gleick.

The availability of cheaper, more powerful computers broadens the applicability of chaos theory. Currently, chaos theory continues to be a very active area of research, involving many different disciplines (mathematics, topology, physics, population biology, biology, meteorology, astrophysics, information theory, etc.).

Chaotic dynamics

For a dynamical system to be classified as chaotic, it must have the following properties:[32]

1. it must be sensitive to initial conditions,
2. it must be topologically mixing, and
3. its periodic orbits must be dense.



Assign z to z^2 minus the conjugate of z, plus the original value of the pixel for each pixel, then count how many cycles it took when the absolute value of z exceeds two; inversion (borders are inner set), so that you can see that it threatens to fail that third condition, even if it meets condition two.

Sensitivity to initial conditions means that each point in such a system is arbitrarily closely approximated by other points with significantly different future trajectories. Thus, an arbitrarily small perturbation of the current trajectory may lead to significantly different future behaviour. However, it has been shown that the first two conditions in fact imply this one.[33]

Sensitivity to initial conditions is popularly known as the "butterfly effect," so called because of the title of a paper given by Edward Lorenz in 1972 to the American Association for the Advancement of Science in Washington, D.C. entitled *Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?* The flapping wing represents a small change in the initial condition of the system, which causes a chain of events leading to large-scale phenomena. Had the butterfly not flapped its wings, the trajectory of the system might have been vastly different.

Sensitivity to initial conditions is often confused with chaos in popular accounts. It can also be a subtle property, since it depends on a choice of metric, or the notion of distance in the phase space of the system. For example, consider the simple dynamical system produced by repeatedly doubling an initial value. This system has sensitive dependence on initial conditions everywhere, since any pair of nearby points will eventually become widely separated. However, it has extremely simple behaviour, as all points except 0 tend to infinity. If instead we use the bounded metric on the line obtained by adding the point at infinity and viewing the result as a circle, the system no longer is sensitive to initial conditions. For this reason, in defining chaos, attention is normally restricted to systems

with bounded metrics, or closed, bounded invariant subsets of unbounded systems.

Even for bounded systems, sensitivity to initial conditions is not identical with chaos. For example, consider the two-dimensional torus described by a pair of angles (x,y), each ranging between zero and 2π. Define a mapping that takes any point (x,y) to $(2x, y + a)$, where a is any number such that $a/2\pi$ is irrational. Because of the doubling in the first coordinate, the mapping exhibits sensitive dependence on initial conditions. However, because of the irrational rotation in the second coordinate, there are no periodic orbits, and hence the mapping is not chaotic according to the definition above.

Topologically mixing means that the system will evolve over time so that any given region or open set of its phase space will eventually overlap with any other given region. Here, "mixing" is really meant to correspond to the standard intuition: the mixing of colored dyes or fluids is an example of a chaotic system.

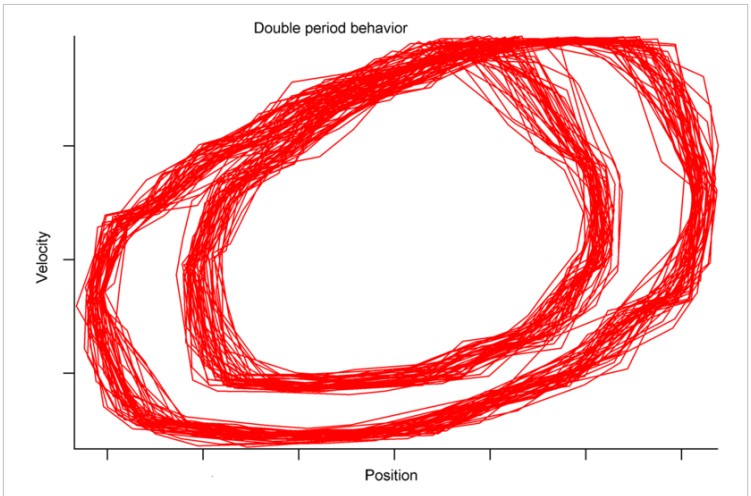
Linear systems are never chaotic; for a dynamical system to display chaotic behaviour it has to be nonlinear. Also, by the Poincaré–Bendixson theorem, a continuous dynamical system on the plane cannot be chaotic; among continuous systems only those whose phase space is non-planar (having dimension at least three, or with a non-Euclidean geometry) can exhibit chaotic behaviour. However, a discrete dynamical system (such as the logistic map) can exhibit chaotic behaviour in a one-dimensional or two-dimensional phase space.

Attractors

Some dynamical systems are chaotic everywhere (see e.g. Anosov diffeomorphisms) but in many cases chaotic behaviour is found only in a subset of phase space. The cases of most interest arise when the chaotic behaviour takes place on an attractor, since then a large set of initial conditions will lead to orbits that converge to this chaotic region.

An easy way to visualize a chaotic attractor is to start with a point in the basin of attraction of the attractor, and then simply plot its subsequent orbit. Because of the topological transitivity condition, this is likely to produce a picture of the entire final attractor.

For instance, in a system describing a pendulum, the phase space might be two-dimensional, consisting of information about position and velocity. One might plot the *position* of a pendulum against its *velocity*. A pendulum at rest will be plotted as a point, and one in periodic motion will be plotted as a simple closed curve. When such a plot forms a closed curve, the curve is called an orbit. Our pendulum has an infinite number of such orbits, forming a pencil of nested ellipses about the origin.



Phase diagram for a damped driven pendulum, with double period motion

Strange attractors

While most of the motion types mentioned above give rise to very simple attractors, such as points and circle-like curves called *limit cycles*, chaotic motion gives rise to what are known as *strange attractors*, attractors that can have great detail and complexity. For instance, a simple three-dimensional model of the Lorenz weather system gives rise to the famous Lorenz attractor. The Lorenz attractor is perhaps one of the best-known chaotic system diagrams, probably because not only was it one of the first, but it is one of the most complex and as such gives rise to a very interesting pattern which looks like the wings of a butterfly. Another such attractor is the Rössler map, which experiences period-two doubling route to chaos, like the logistic map.

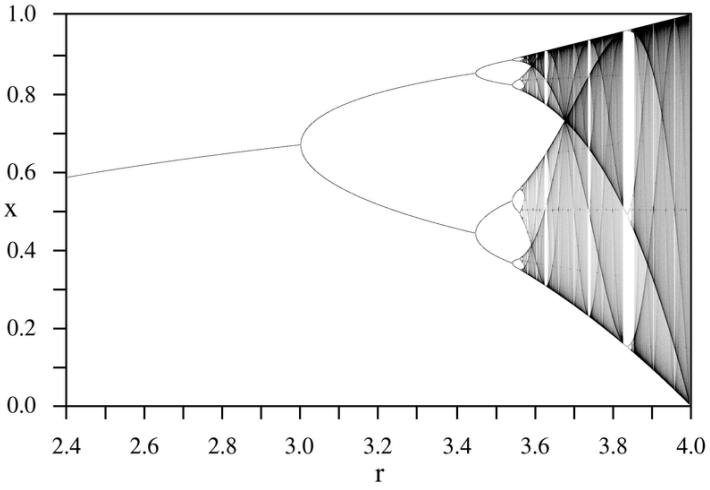
Strange attractors occur in both continuous dynamical systems (such as the Lorenz system) and in some discrete systems (such as the Hénon map). Other discrete dynamical systems have a repelling structure called a Julia set which forms at the boundary between basins of attraction of fixed points - Julia sets can be thought of as strange *repellers*. Both strange attractors and Julia sets typically have a fractal structure.

The Poincaré-Bendixson theorem shows that a strange attractor can only arise in a continuous dynamical system if it has three or more dimensions. However, no such restriction applies to discrete systems, which can exhibit strange attractors in two or even one dimensional systems.

The initial conditions of three or more bodies interacting through gravitational attraction (see the *n*-body problem) can be arranged to produce chaotic motion.

Minimum complexity of a chaotic system

Simple systems can also produce chaos without relying on differential equations. An example is the logistic map, which is a difference equation (recurrence relation) that describes population growth over time. Another example is the Ricker model of population dynamics.



Bifurcation diagram of a logistic map, displaying chaotic behaviour past a threshold

Even the evolution of simple discrete systems, such as cellular automata, can heavily depend on initial conditions. Stephen Wolfram has investigated a cellular automaton with this property, termed by him *rule 30*.

A minimal model for conservative (reversible) chaotic behavior is provided by Arnold's cat map.

Mathematical theory

Sharkovskii's theorem is the basis of the Li and Yorke (1975) proof that any one-dimensional system which exhibits a regular cycle of period three will also display regular cycles of every other length as well as completely chaotic orbits.

Mathematicians have devised many additional ways to make quantitative statements about chaotic systems. These include: fractal dimension of the attractor, Lyapunov exponents, recurrence plots, Poincaré maps, bifurcation diagrams, and transfer operator.

Distinguishing random from chaotic data

It can be difficult to tell from data whether a physical or other observed process is random or chaotic, because in practice no time series consists of pure 'signal.' There will always be some form of corrupting noise, even if it is present as round-off or truncation error. Thus any real time series, even if mostly deterministic, will contain some randomness.[34]

All methods for distinguishing deterministic and stochastic processes rely on the fact that a deterministic system always evolves in the same way from a given starting point.[34] [35] Thus, given a time series to test for determinism, one can:

1. pick a test state;
2. search the time series for a similar or 'nearby' state; and
3. compare their respective time evolutions.

Define the error as the difference between the time evolution of the 'test' state and the time evolution of the nearby state. A deterministic system will have an error that either remains small (stable, regular solution) or increases exponentially with time (chaos). A stochastic system will have a randomly distributed error.[36]

Essentially all measures of determinism taken from time series rely upon finding the closest states to a given 'test' state (i.e., correlation dimension, Lyapunov exponents, etc.). To define the state of a system one typically relies on phase space embedding methods.[37] Typically one chooses an embedding dimension, and investigates the propagation of the error between two nearby states. If the error looks random, one increases the dimension. If you can increase the dimension to obtain a deterministic looking error, then you are done. Though it may sound simple it is not really. One complication is that as the dimension increases the search for a nearby state requires a lot more computation time and a lot of data (the amount of data required increases exponentially with embedding dimension) to find a suitably close candidate. If the embedding dimension (number of measures per state) is chosen too small (less than the 'true' value) deterministic data can appear to be random but in theory there is no problem choosing the dimension too large – the method will work.

When a non-linear deterministic system is attended by external fluctuations, its trajectories present serious and permanent distortions. Furthermore, the noise is amplified due to the inherent non-linearity and reveals totally new dynamical properties. Statistical tests attempting to separate noise from the deterministic skeleton or inversely isolate the deterministic part risk failure. Things become worse when the deterministic component is a non-linear feedback system.[38] In presence of interactions between nonlinear deterministic components and noise the resulting nonlinear series can display dynamics that traditional tests for nonlinearity are sometimes not able to capture.[39]

Applications

Chaos theory is applied in many scientific disciplines: mathematics, biology, computer science, economics,[40] [41] [42] engineering, finance,[43] [44] philosophy, physics, politics, population dynamics, psychology, and robotics.[45]

One of the most successful applications of chaos theory has been in ecology, where dynamical systems such as the Ricker model have been used to show how population growth under density dependence can lead to chaotic dynamics.

Chaos theory is also currently being applied to medical studies of epilepsy, specifically to the prediction of seemingly random seizures by observing initial conditions.[46]

See also

Examples of chaotic systems

- Arnold's cat map
- Bouncing Ball Simulation System
- Chua's circuit
- Double pendulum
- Dynamical billiards
- Economic bubble
- Hénon map
- Horseshoe map
- Logistic map
- Rössler attractor
- Standard map
- Swinging Atwood's machine
- Tilt A Whirl

Other related topics

- Anosov diffeomorphism
- Bifurcation theory
- Butterfly effect
- Chaos theory in organizational development
- Complexity
- Control of chaos
- Edge of chaos
- Fractal
  - Mandelbrot set
  - Julia set
- Predictability
- Santa Fe Institute
- Synchronization of chaos

People

- Mitchell Feigenbaum
- Brosl Hasslacher
- Michel Hénon
- Edward Lorenz
- Aleksandr Lyapunov
- Benoît Mandelbrot
- Henri Poincaré
- Otto Rössler
- David Ruelle
- Oleksandr Mikolaiovich Sharkovsky
- Floris Takens
- James A. Yorke

References

[1] Raymond Sneyers (1997) "Climate Chaotic Instability: Statistical Determination and Theoretical Background", *Environmetrics*, vol. 8, no. 5, pages 517–532.

[2] Apostolos Serletis and Periklis Gogas, *Purchasing Power Parity Nonlinearity and Chaos (http://www.informaworld.com/smpp/content~content=a713761243~db=all~order=page)*, in: Applied Financial Economics, 10, 615–622, 2000.

[3] Apostolos Serletis and Periklis Gogas *The North American Gas Markets are Chaotic (http://mpra.ub.uni-muenchen.de/1576/01/MPRA_paper_1576.pdf)*PDF (918 KB), in: The Energy Journal, 20, 83–103, 1999.

[4] Apostolos Serletis and Periklis Gogas, *Chaos in East European Black Market Exchange Rates (http://ideas.repec.org/a/eee/reecon/v51y1997i4p359-385.html)*, in: Research in Economics, 51, 359–385, 1997.

[5] Werndl, Charlotte (2009). *What are the New Implications of Chaos for Unpredictability? (http://bjps.oxfordjournals.org/cgi/content/abstract/60/1/195)*. The British Journal for the Philosophy of Science 60, 195–220.

[6] A. E. Motter, *Relativistic chaos is coordinate invariant (http://prola.aps.org/abstract/PRL/v91/i23/e231101)*, in: Phys. Rev. Lett. 91, 231101 (2003).

[7] Jules Henri Poincaré (1890) "Sur le problème des trois corps et les équations de la dynamique. Divergence des séries de M. Lindstedt," *Acta Mathematica*, vol. 13, pages 1–270.

[8] Florin Diacu and Philip Holmes (1996) *Celestial Encounters: The Origins of Chaos and Stability*, Princeton University Press.

[9] Hadamard, Jacques (1898). "Les surfaces à courbures opposées et leurs lignes géodesiques". *Journal de Mathématiques Pures et Appliquées* **4**: pp. 27–73.

[10] George D. Birkhoff, *Dynamical Systems*, vol. 9 of the American Mathematical Society Colloquium Publications (Providence, Rhode Island: American Mathematical Society, 1927)

[11] Kolmogorov, A. N. (1941a). “Local structure of turbulence in an incompressible fluid for very large Reynolds numbers,” *Doklady Akademii Nauk SSSR*, vol. 30, no. 4, pages 301–305. Reprinted in: *Proceedings of the Royal Society of London: Mathematical and Physical Sciences* (Series A), vol. 434, pages 9–13 (1991).

[12] Kolmogorov, A. N. (1941b) “On degeneration of isotropic turbulence in an incompressible viscous liquid,” *Doklady Akademii Nauk SSSR*, vol. 31, no. 6, pages 538–540. Reprinted in: *Proceedings of the Royal Society of London: Mathematical and Physical Sciences* (Series A), vol. 434, pages 15–17 (1991).

[13] Andrey N. Kolmogorov (1954) "Preservation of conditionally periodic movements with small change in the Hamiltonian function," *Doklady Akademii Nauk SSSR*, vol. 98, pages 527–530. See also Kolmogorov–Arnold–Moser theorem

[14] Mary L. Cartwright and John E. Littlewood (1945) "On non-linear differential equations of the second order, I: The equation $y'' + k(1-y^2)y' + y = b\lambda k\cos(\lambda t + a)$, k large," *Journal of the London Mathematical Society*, vol. 20, pages 180–189. See also: Van der Pol oscillator

[15] Stephen Smale (January 1960) "Morse inequalities for a dynamical system," *Bulletin of the American Mathematical Society*, vol. 66, pages 43–49.

[16] Edward N. Lorenz, "Deterministic non-periodic flow," *Journal of the Atmospheric Sciences*, vol. 20, pages 130–141 (1963).

[17] Gleick, James (1987). *Chaos: Making a New Science*. London: Cardinal. pp. 17.

[18] Mandelbrot, Benoît (1963). "The variation of certain speculative prices". *Journal of Business* **36**: pp. 394–419.

[19] J.M. Berger and B. Mandelbrot (July 1963) "A new model for error clustering in telephone circuits," *I.B.M. Journal of Research and Development*, vol 7, pages 224–236.

[20] B. Mandelbrot, *The Fractal Geometry of Nature* (N.Y., N.Y.: Freeman, 1977), page 248.

[21] See also: Benoît B. Mandelbrot and Richard L. Hudson, *The (Mis)behavior of Markets: A Fractal View of Risk, Ruin, and Reward* (N.Y., N.Y.: Basic Books, 2004), page 201.

[22] Benoît Mandelbrot (5 May 1967) "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension," *Science*, Vol. 156, No. 3775, pages 636–638.

[23] B. van der Pol and J. van der Mark (1927) "Frequency demultiplication," *Nature*, vol. 120, pages 363–364. See also: Van der Pol oscillator

[24] R.L. Ives (10 October 1958) "Neon oscillator rings," *Electronics*, vol. 31, pages 108–115.

[25] See p. 83 of Lee W. Casperson, "Gas laser instabilities and their interpretation," pages 83–98 in: N. B. Abraham, F. T. Arecchi, and L. A. Lugiato, eds., *Instabilities and Chaos in Quantum Optics II: Proceedings of the NATO Advanced Study Institute, Il Ciocco, Italy, June 28–July 7, 1987* (N.Y., N.Y.: Springer Verlag, 1988).

[26] Ralph H. Abraham and Yoshisuke Ueda, eds., *The Chaos Avant-Garde: Memoirs of the Early Days of Chaos Theory* (Singapore: World Scientific Publishing Co., 2001). See Chapters 3 and 4.

[27] Mitchell Feigenbaum (July 1978) "Quantitative universality for a class of nonlinear transformations," *Journal of Statistical Physics*, vol. 19, no. 1, pages 25–52.

[28] "The Wolf Prize in Physics in 1986." (http://www.wolffund.org.il/cat.asp?id=25&cat_title=PHYSICS). .

[29] Bernardo Huberman, "A Model for Dysfunctions in Smooth Pursuit Eye Movement" Annals of the New York Academy of Sciences, Vol. 504 Page 260 July 1987, Perspectives in Biological Dynamics and Theoretical Medicine

[30] Per Bak, Chao Tang, and Kurt Wiesenfeld, "Self-organized criticality: An explanation of the 1/f noise," *Physical Review Letters*, vol. 59, no. 4, pages 381–384 (27 July 1987). However, the conclusions of this article have been subject to dispute. See: http://www.nslij-genetics.org/wli/1fnoise/1fnoise_square.html . See especially: Lasse Laurson, Mikko J. Alava, and Stefano Zapperi, "Letter: Power spectra of self-organized critical sand piles," *Journal of Statistical Mechanics: Theory and Experiment*, 0511, L001 (15 September 2005).

[31] F. Omori (1894) "On the aftershocks of earthquakes," *Journal of the College of Science, Imperial University of Tokyo*, vol. 7, pages 111–200.

[32] Hasselblatt, Boris; Anatole Katok (2003). *A First Course in Dynamics: With a Panorama of Recent Developments*. Cambridge University Press. ISBN 0521587506.

[33] Saber N. Elaydi, *Discrete Chaos*, Chapman & Hall/CRC, 1999, page 117.

[34] Provenzale A. et al.: "Distinguishing between low-dimensional dynamics and randomness in measured time-series", in: Physica D, 58:31–49, 1992

[35] Sugihara G. and May R.: "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series", in: Nature, 344:734–41, 1990

[36] Casdagli, Martin. "Chaos and Deterministic *versus* Stochastic Non-linear Modelling", in: *Journal Royal Statistics Society*: Series B, 54, nr. 2 (1991), 303-28

[37] Broomhead D. S. and King G. P.: "Extracting Qualitative Dynamics from Experimental Data", in: Physica 20D, 217–36, 1986

[38] Kyrtsou, C., (2008). Re-examining the sources of heteroskedasticity: the paradigm of noisy chaotic models, *Physica A*, 387, pp. 6785–6789.

[39] Kyrtsou, C., (2005). Evidence for neglected linearity in noisy chaotic models, International Journal of Bifurcation and Chaos, 15(10), pp. 3391–3394.
[40] Kyrtsou, C. and W. Labys, (2006). Evidence for chaotic dependence between US inflation and commodity prices, *Journal of Macroeconomics*, 28(1), pp. 256–266.
[41] Kyrtsou, C. and W. Labys, (2007). Detecting positive feedback in multivariate time series: the case of metal prices and US inflation, *Physica A*, 377(1), pp. 227–229.
[42] Kyrtsou, C., and Vorlow, C., (2005). Complex dynamics in macroeconomics: A novel approach, in *New Trends in Macroeconomics*, Diebolt, C., and Kyrtsou, C., (eds.), Springer Verlag.
[43] Hristu-Varsakelis, D., and Kyrtsou, C., (2008): Evidence for nonlinear asymmetric causality in US inflation, metal and stock returns, *Discrete Dynamics in Nature and Society*, Volume 2008, Article ID 138547, 7 pages, doi:10.1155/2008/138547.
[44] Kyrtsou, C. and M. Terraza, (2003). Is it possible to study chaotic and ARCH behaviour jointly? Application of a noisy Mackey-Glass equation with heteroskedastic errors to the Paris Stock Exchange returns series, *Computational Economics*, 21, 257–276.
[45] Metaculture.net, metalinks: Applied Chaos (http://metalinks.metaculture.net/science/fractal/applications/default.asp), 2007.
[46] Comdig.org, Complexity Digest 199.06 (http://www.comdig.org/index.php?id_issue=1999.06#194)

Scientific literature

Articles

- A.N. Sharkovskii, "Co-existence of cycles of a continuous mapping of the line into itself", Ukrainian Math. J., 16:61–71 (1964)
- Li, T. Y. and Yorke, J. A. "Period Three Implies Chaos." American Mathematical Monthly 82, 985–92, 1975.
- Kolyada, S. F. " Li-Yorke sensitivity and other concepts of chaos (http://www.springerlink.com/content/q00627510552020g/?p=93e1f3daf93549d1850365a8800afb30&pi=3)", Ukrainian Math. J. 56 (2004), 1242–1257.
- C.E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July, October, 1948

Textbooks

- Alligood, K. T. (1997). *Chaos: an introduction to dynamical systems*. Springer-Verlag New York, LLC. ISBN 0-387-94677-2.
- Baker, G. L. (1996). *Chaos, Scattering and Statistical Mechanics*. Cambridge University Press. ISBN 0-521-39511-9.
- Badii, R.; Politi A. (1997). "*Complexity: hierarchical structures and scaling in physics" (http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521663857)*. Cambridge University Press. ISBN 0521663857. http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521663857.
- Devaney, Robert L. (2003). *An Introduction to Chaotic Dynamical Systems, 2nd ed,*. Westview Press. ISBN 0-8133-4085-3.
- Gollub, J. P.; Baker, G. L. (1996). *Chaotic dynamics*. Cambridge University Press. ISBN 0-521-47685-2.
- Gutzwiller, Martin (1990). *Chaos in Classical and Quantum Mechanics*. Springer-Verlag New York, LLC. ISBN 0-387-97173-4.
- Hoover, William Graham (1999,2001). *Time Reversibility, Computer Simulation, and Chaos*. World Scientific. ISBN 981-02-4073-2.

- Kiel, L. Douglas; Elliott, Euel W. (1997). *Chaos Theory in the Social Sciences*. Perseus Publishing. ISBN 0-472-08472-0.
- Moon, Francis (1990). *Chaotic and Fractal Dynamics*. Springer-Verlag New York, LLC. ISBN 0-471-54571-6.
- Ott, Edward (2002). *Chaos in Dynamical Systems*. Cambridge University Press New, York. ISBN 0-521-01084-5.
- Strogatz, Steven (2000). *Nonlinear Dynamics and Chaos*. Perseus Publishing. ISBN 0-7382-0453-6.
- Sprott, Julien Clinton (2003). *Chaos and Time-Series Analysis*. Oxford University Press. ISBN 0-19-850840-9.
- Tél, Tamás; Gruiz, Márton (2006). *Chaotic dynamics: An introduction based on classical mechanics*. Cambridge University Press. ISBN 0-521-83912-2.
- Tufillaro, Abbott, Reilly (1992). *An experimental approach to nonlinear dynamics and chaos*. Addison-Wesley New York. ISBN 0-201-55441-0.
- Zaslavsky, George M. (2005). *Hamiltonian Chaos and Fractional Dynamics*. Oxford University Press. ISBN 0-198-52604-0.

Semitechnical and popular works

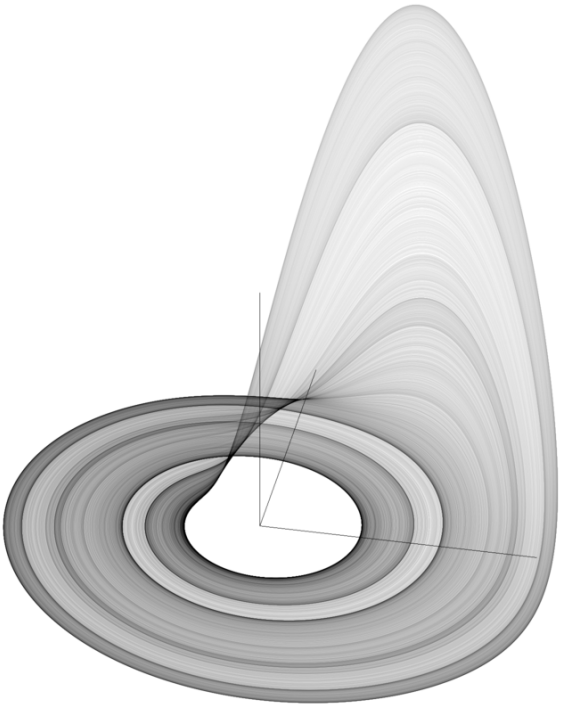
- Ralph H. Abraham and Yoshisuke Ueda (Ed.), *The Chaos Avant-Garde: Memoirs of the Early Days of Chaos Theory*, World Scientific Publishing Company, 2001, 232 pp.
- Michael Barnsley, *Fractals Everywhere*, Academic Press 1988, 394 pp.
- Richard J Bird, *Chaos and Life: Complexity and Order in Evolution and Thought*, Columbia University Press 2003, 352 pp.
- John Briggs and David Peat, *Turbulent Mirror: : An Illustrated Guide to Chaos Theory and the Science of Wholeness*, Harper Perennial 1990, 224 pp.
- John Briggs and David Peat, *Seven Life Lessons of Chaos: Spiritual Wisdom from the Science of Change*, Harper Perennial 2000, 224 pp.
- Lawrence A. Cunningham, *From Random Walks to Chaotic Crashes: The Linear Genealogy of the Efficient Capital Market Hypothesis,* George Washington Law Review, Vol. 62, 1994, 546 pp.
- Leon Glass and Michael C. Mackey, *From Clocks to Chaos: The Rhythms of Life,* Princeton University Press 1988, 272 pp.
- James Gleick, *Chaos: Making a New Science*, New York: Penguin, 1988. 368 pp.
- John Gribbin, *Deep Simplicity*,
- L Douglas Kiel, Euel W Elliott (ed.), *Chaos Theory in the Social Sciences: Foundations and Applications*, University of Michigan Press, 1997, 360 pp.
- Arvind Kumar, *Chaos, Fractals and Self-Organisation; New Perspectives on Complexity in Nature* , National Book Trust, 2003.
- Hans Lauwerier, *Fractals*, Princeton University Press, 1991.
- Edward Lorenz, *The Essence of Chaos*, University of Washington Press, 1996.
- Chapter 5 of Alan Marshall (2002) The Unity of nature, Imperial College Press: London
- Heinz-Otto Peitgen and Dietmar Saupe (Eds.), *The Science of Fractal Images*, Springer 1988, 312 pp.
- Clifford A. Pickover, *Computers, Pattern, Chaos, and Beauty: Graphics from an Unseen World* , St Martins Pr 1991.
- Ilya Prigogine and Isabelle Stengers, *Order Out of Chaos*, Bantam 1984.

- Heinz-Otto Peitgen and P. H. Richter, *The Beauty of Fractals : Images of Complex Dynamical Systems*, Springer 1986, 211 pp.
- David Ruelle, *Chance and Chaos*, Princeton University Press 1993.
- Ivars Peterson, *Newton's Clock: Chaos in the Solar System*, Freeman, 1993.
- David Ruelle, *Chaotic Evolution and Strange Attractors*, Cambridge University Press, 1989.
- Peter Smith, *Explaining Chaos*, Cambridge University Press, 1998.
- Ian Stewart, *Does God Play Dice?: The Mathematics of Chaos* , Blackwell Publishers, 1990.
- Steven Strogatz, *Sync: The emerging science of spontaneous order*, Hyperion, 2003.
- Yoshisuke Ueda, *The Road To Chaos*, Aerial Pr, 1993.
- M. Mitchell Waldrop, *Complexity : The Emerging Science at the Edge of Order and Chaos*, Simon & Schuster, 1992.

External links

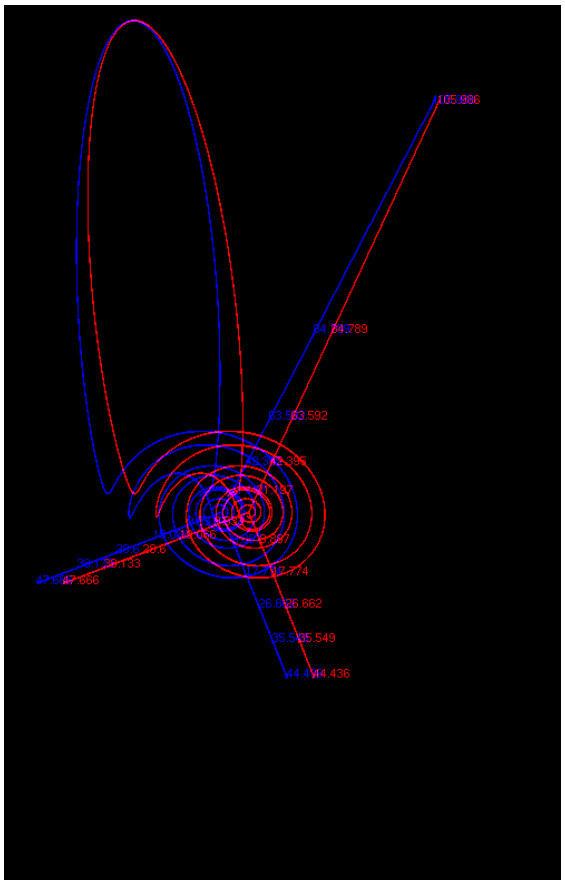
- Nonlinear Dynamics Research Group (http://lagrange.physics.drexel.edu) with Animations in Flash
- The Chaos group at the University of Maryland (http://www.chaos.umd.edu)
- The Chaos Hypertextbook (http://hypertextbook.com/chaos/). An introductory primer on chaos and fractals.
- Society for Chaos Theory in Psychology & Life Sciences (http://www.societyforchaostheory.org/)
- Interactive live chaotic pendulum experiment (http://physics.mercer.edu/pendulum/), allows users to interact and sample data from a real working damped driven chaotic pendulum
- Nonlinear dynamics: how science comprehends chaos (http://www.creatingtechnology.org/papers/chaos.htm), talk presented by Sunny Auyang, 1998.
- Nonlinear Dynamics (http://www.egwald.ca/nonlineardynamics/index.php). Models of bifurcation and chaos by Elmer G. Wiens
- Gleick's *Chaos* (excerpt) (http://www.around.com/chaos.html)
- Systems Analysis, Modelling and Prediction Group (http://www.eng.ox.ac.uk/samp) at the University of Oxford.
- A page about the Mackey-Glass equation (http://www.mgix.com/snippets/?MackeyGlass).

Rössler attractor



The Rössler attractor

The **Rössler attractor** (pronounced /ˈrɒslər/) is the attractor for the **Rössler system**, a system of three non-linear ordinary differential equations. These differential equations define a continuous-time dynamical system that exhibits chaotic dynamics associated with the fractal properties of the attractor. Some properties of the Rössler system can be deduced via linear methods such as eigenvectors, but the main features of the system require non-linear methods such as Poincaré maps and bifurcation diagrams. The original Rössler paper says the Rössler attractor was intended to behave similarly to the Lorenz attractor, but also be easier to analyze qualitatively. An orbit within the attractor follows an outward spiral close to the x, y plane around an unstable fixed point. Once the graph spirals out enough, a second fixed point influences the graph, causing a rise and twist in the z-dimension. In the time domain, it becomes apparent that although each variable is oscillating within a fixed range of values, the oscillations are chaotic. This attractor has some similarities to the Lorenz attractor, but is simpler and has only one manifold. Otto Rössler designed the Rössler attractor in 1976, but the originally theoretical equations were later found to be useful in modeling equilibrium in chemical reactions. The defining equations are:



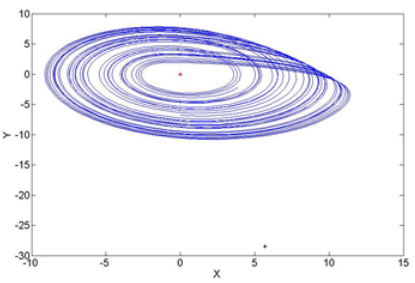
Rössler attractor as a stereogram with $a = 0.2$, $b = 0.2$, $c = 14$

$$\frac{dx}{dt} = -y - z$$
$$\frac{dy}{dt} = x + ay$$
$$\frac{dz}{dt} = b + z(x - c)$$

Rössler studied the chaotic attractor with $a = 0.2$, $b = 0.2$, and $c = 5.7$, though properties of $a = 0.1$, $b = 0.1$, and $c = 14$ have been more commonly used since.

An analysis

Some of the Rössler attractor's elegance is due to two of its equations being linear; setting $z = 0$, allows examination of the behavior on the x, y plane



x, y plane of Rössler attractor with $a = 0.2$, $b = 0.2$, $c = 5.7$

$$\frac{dx}{dt} = -y$$
$$\frac{dy}{dt} = x + ay$$

The stability in the x, y plane can then be found by calculating the eigenvalues of the Jacobian $\begin{pmatrix} 0 & -1 \\ 1 & a \end{pmatrix}$, which are $(a \pm \sqrt{a^2 - 4})/2$. From this, we can see that when $0 < a < 2$, the eigenvalues are complex and at least one has a real component, making the origin unstable with an outwards spiral on the x, y plane. Now consider the z plane behavior within the context of this range for a. So long as x is smaller than c, the c term will keep the orbit close to the x, y plane. As the orbit approaches x greater than c, the z -values begin to climb. As z climbs, though, the $-z$ in the equation for dx/dt stops the growth in x .

Fixed points

In order to find the fixed points, the three Rössler equations are set to zero and the (x , y , z) coordinates of each fixed point were determined by solving the resulting equations. This yields the general equations of each of the fixed point coordinates:

$$x = \frac{c \pm \sqrt{c^2 - 4ab}}{2}$$
$$y = -\left(\frac{c \pm \sqrt{c^2 - 4ab}}{2a}\right)$$
$$z = \frac{c \pm \sqrt{c^2 - 4ab}}{2a}$$

Which in turn can be used to show the actual fixed points for a given set of parameter values:

$$\left(\frac{c + \sqrt{c^2 - 4ab}}{2}, \frac{-c - \sqrt{c^2 - 4ab}}{2a}, \frac{c + \sqrt{c^2 - 4ab}}{2a}\right)$$
$$\left(\frac{c - \sqrt{c^2 - 4ab}}{2}, \frac{-c + \sqrt{c^2 - 4ab}}{2a}, \frac{c - \sqrt{c^2 - 4ab}}{2a}\right)$$

As shown in the general plots of the Rössler Attractor above, one of these fixed points resides in the center of the attractor loop and the other lies comparatively removed from the attractor.

Eigenvalues and eigenvectors

The stability of each of these fixed points can be analyzed by determining their respective eigenvalues and eigenvectors. Beginning with the Jacobian: $\begin{pmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ z & 0 & x - c \end{pmatrix}$, the eigenvalues can be determined by solving the following cubic:

$$-\lambda^3 + \lambda^2(a + x - c) + \lambda(ac - ax - 1 - z) + x - c + az$$

For the centrally located fixed point, Rössler's original parameter values of a=0.2, b=0.2, and c=5.7 yield eigenvalues of:

$$\lambda_1 = 0.0971028 + 0.995786i$$

$$\lambda_2 = 0.0971028 - 0.995786i$$
$$\lambda_3 = -5.68718$$

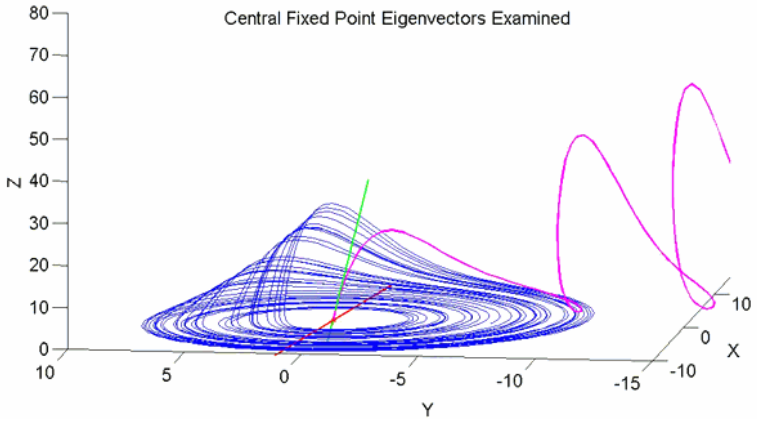
(Using Mathematica 7)

The magnitude of a negative eigenvalue characterizes the level of attraction along the corresponding eigenvector. Similarly the magnitude of a positive eigenvalue characterizes the level of repulsion along the corresponding eigenvector.

The eigenvectors corresponding to these eigenvalues are:

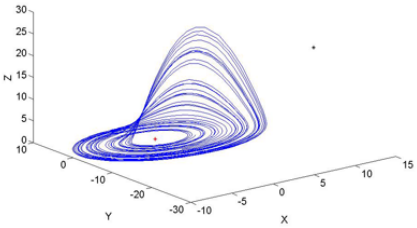
$$v_1 = \begin{pmatrix} 0.7073 \\ -0.07278 - 0.7032i \\ 0.0042 - 0.0007i \end{pmatrix}$$
$$v_2 = \begin{pmatrix} 0.7073 \\ 0.07278 + 0.7032i \\ 0.0042 + 0.0007i \end{pmatrix}$$
$$v_3 = \begin{pmatrix} 0.1682 \\ -0.0286 \\ 0.9853 \end{pmatrix}$$

These eigenvectors have several interesting implications. First, the two eigenvalue/eigenvector pairs (v_1 and v_2) are responsible for the steady outward slide that occurs in the main disk of the attractor. The last eigenvalue/eigenvector pair is attracting along an axis that runs through the center of the manifold and accounts for the z motion that occurs within the attractor. This effect is roughly demonstrated with the figure below.



Examination of central fixed point eigenvectors: The blue line corresponds to the standard Rössler attractor generated with $a = 0.2$, $b = 0.2$, and $c = 5.7$.

The figure examines the central fixed point eigenvectors. The blue line corresponds to the standard Rössler attractor generated with $a = 0.2$, $b = 0.2$, and $c = 5.7$. The red dot in the center of this attractor is FP_1. The red line intersecting that fixed point is an illustration of the repulsing plane generated by v_1 and v_2. The green line is an illustration of the attracting v_3. The magenta line is generated by stepping backwards through time from a point on the attracting eigenvector which is slightly above FP_1– it illustrates the behavior of points that become completely dominated by that vector. Note



Rössler attractor with $a = 0.2$, $b = 0.2$, $c = 5.7$

that the magenta line nearly touches the plane of the attractor before being pulled upwards into the fixed point; this suggests that the general appearance and behavior of the Rössler attractor is largely a product of the interaction between the attracting v_3 and the repelling v_1 and v_2 plane. Specifically it implies that a sequence generated from the Rössler equations will begin to loop around FP_1, start being pulled upwards into the v_3 vector, creating the upward arm of a curve that bends slightly inward toward the vector before being pushed outward again as it is pulled back towards the repelling plane.

For the outlier fixed point, Rössler's original parameter values of $a = 0.2$, $b = 0.2$, and $c = 5.7$ yield eigenvalues of:

$$\lambda_1 = -0.0000046 + 5.4280259i$$
$$\lambda_2 = -0.0000046 - 5.4280259i$$
$$\lambda_3 = 0.1929830$$

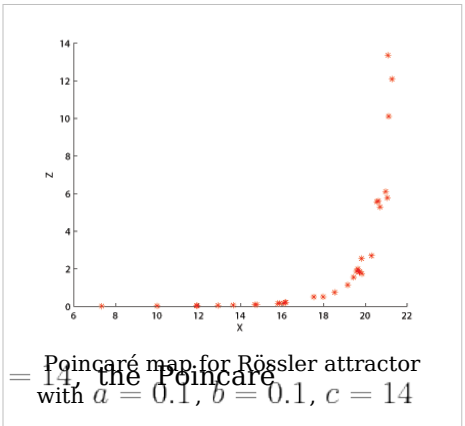
The eigenvectors corresponding to these eigenvalues are:

$$v_1 = \begin{pmatrix} 0.0002422 + 0.1872055i \\ 0.0344403 - 0.0013136i \\ 0.9817159 \end{pmatrix}$$
$$v_2 = \begin{pmatrix} 0.0002422 - 0.1872055i \\ 0.0344403 + 0.0013136i \\ 0.9817159 \end{pmatrix}$$
$$v_3 = \begin{pmatrix} 0.0049651 \\ -0.7075770 \\ 0.7066188 \end{pmatrix}$$

Although these eigenvalues and eigenvectors exist in the Rössler attractor, their influence is confined to iterations of the Rössler system whose initial conditions are in the general vicinity of this outlier fixed point. Except in those cases where the initial conditions lie on the attracting plane generated by λ_1 and λ_2, this influence effectively involves pushing the resulting system towards the general Rössler attractor. As the resulting sequence approaches the central fixed point and the attractor itself, the influence of this distant fixed point (and its eigenvectors) will wane.

Poincaré map

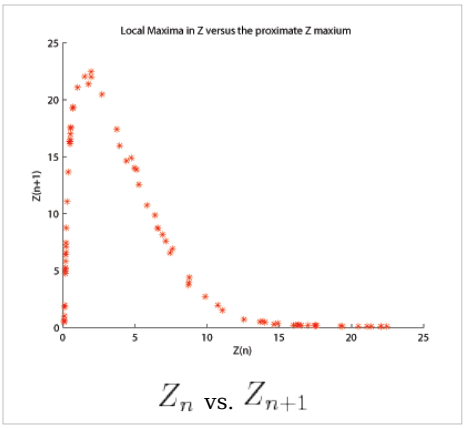
The Poincaré map is constructed by plotting the value of the function every time it passes through a set plane in a specific direction. An example would be plotting the y, z value every time it passes through the $x = 0$ plane where x is changing from negative to positive, commonly done when studying the Lorenz attractor. In the case of the Rössler attractor, the $x = 0$ plane is uninteresting, as the map always crosses the $x = 0$ plane at $z = 0$ due to the nature of the Rössler equations. In the $a = 0.1$ plane for $a = 0.1$, $b = 0.1$, $c = 14$, the Poincaré map shows the upswing in z values as x increases, as is to be expected due to the upswing and twist section of the Rössler plot. The number of points in this specific Poincaré plot is infinite, but when a different c value is used, the number of points can vary. For



Poincaré map for Rössler attractor with $a = 0.1$, $b = 0.1$, $c = 14$

example, with a c value of 4, there is only one point on the Poincaré map, because the function yields a periodic orbit of period one, or if the c value is set to 12.8, there would be six points corresponding to a period six orbit.

Mapping local maxima



Z_n vs. Z_{n+1}

In the original paper on the Lorenz Attractor, Edward Lorenz analyzed the local maxima of z against the immediately preceding local maxima. When visualized, the plot resembled the tent map, implying that similar analysis can be used between the map and attractor. For the Rössler attractor, when the z_n local maximum is plotted against the next local z maximum, z_{n+1}, the resulting plot (shown here for $a = 0.2$, $b = 0.2$, $c = 5.7$) is unimodal, resembling a skewed Henon map. Knowing that the Rössler attractor can be used to create a pseudo 1-d map, it then follows to use similar analysis methods. The bifurcation diagram is specifically a useful analysis method.

Variation of parameters

Rössler attractor's behavior is largely a factor of the values of its constant parameters (a, b, and c). In general varying each parameter has a comparable effect by causing the system to converge toward a periodic orbit, fixed point, or escape towards infinity, however the specific ranges and behaviors induced vary substantially for each parameter. Periodic orbits, or "unit cycles," of the Rössler system are defined by the number of loops around the central point that occur before the loops series begins to repeat itself.

Bifurcation diagrams are a common tool for analyzing the behavior of chaotic systems. Bifurcation diagrams for the Rössler attractor are created by iterating through the Rössler ODEs holding two of the parameters constant while conducting a parameter sweep over a range of possible values for the third. The local x maxima for each varying parameter value is then plotted against that parameter value. These maxima are determined after the attractor has reached steady state and any initial transient behaviors have disappeared. This is useful in determining the relationship between periodicity and the selected parameter. Increasing numbers of points in a vertical line on a bifurcation diagram indicates the Rössler attractor behaves chaotically that value of the parameter being examined.

Varying a

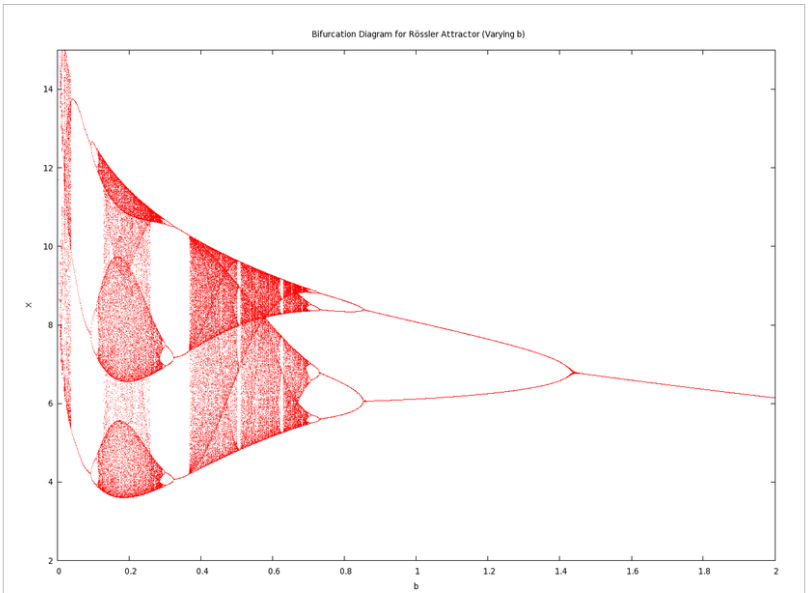
In order to examine the behavior of the Rössler attractor for different values of a, b was fixed at 0.2, c was fixed at 5.7. Numerical examination of attractor's behavior over changing a suggests it has a disproportional influence over the attractor's behavior. Some examples of this relationship include:

- $a \leq 0$: converges to the centrally located fixed point
- $a = 0.1$: unit cycle of period 1
- $a = 0.2$: standard parameter value selected by Rössler, chaotic
- $a = 0.3$: chaotic attractor, significantly more Möbius strip-like (folding over itself).
- $a = 0.35$: similar to .3, but increasingly chaotic
- $a = 0.38$: similar to .35, but increasingly chaotic

If a gets even slightly larger than .38, it causes MATLAB to hang. Note this suggests that the practical range of a is very narrow.

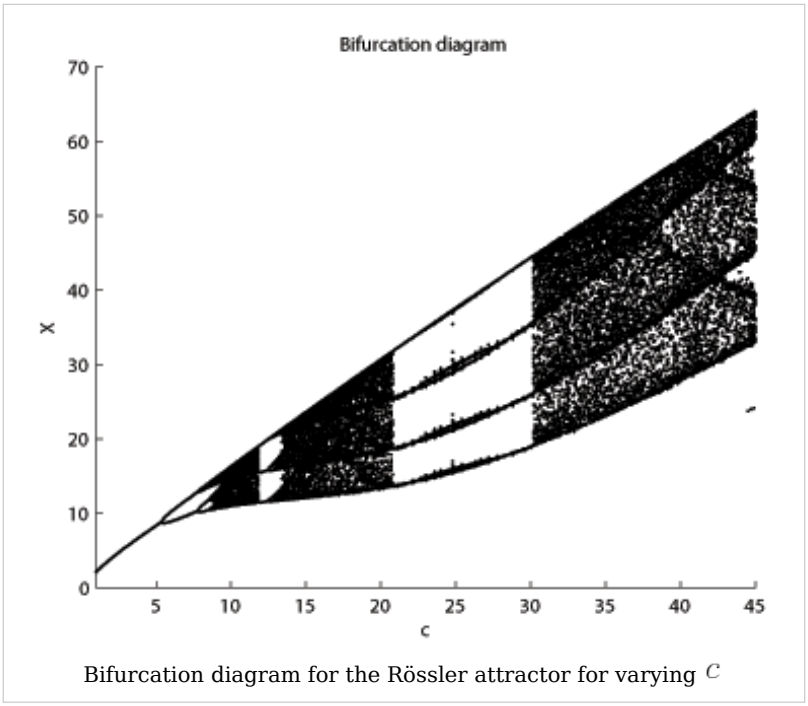
Varying b

The effect of b on the Rössler attractor's behavior is best illustrated through a bifurcation diagram. This bifurcation diagram was created with $a = 0.2$, $c = 5.7$. As shown in the accompanying diagram, as b approaches 0 the attractor approaches infinity (note the upswing for very small values of b. Comparative to the other parameters, varying b seems to generate a greater range when period-3 and period-6 orbits will occur. In contrast to a and c, higher values of b systems that converge on a period-1 orbit instead of higher level orbits or chaotic attractors.



Bifurcation diagram for the Rössler attractor for varying b

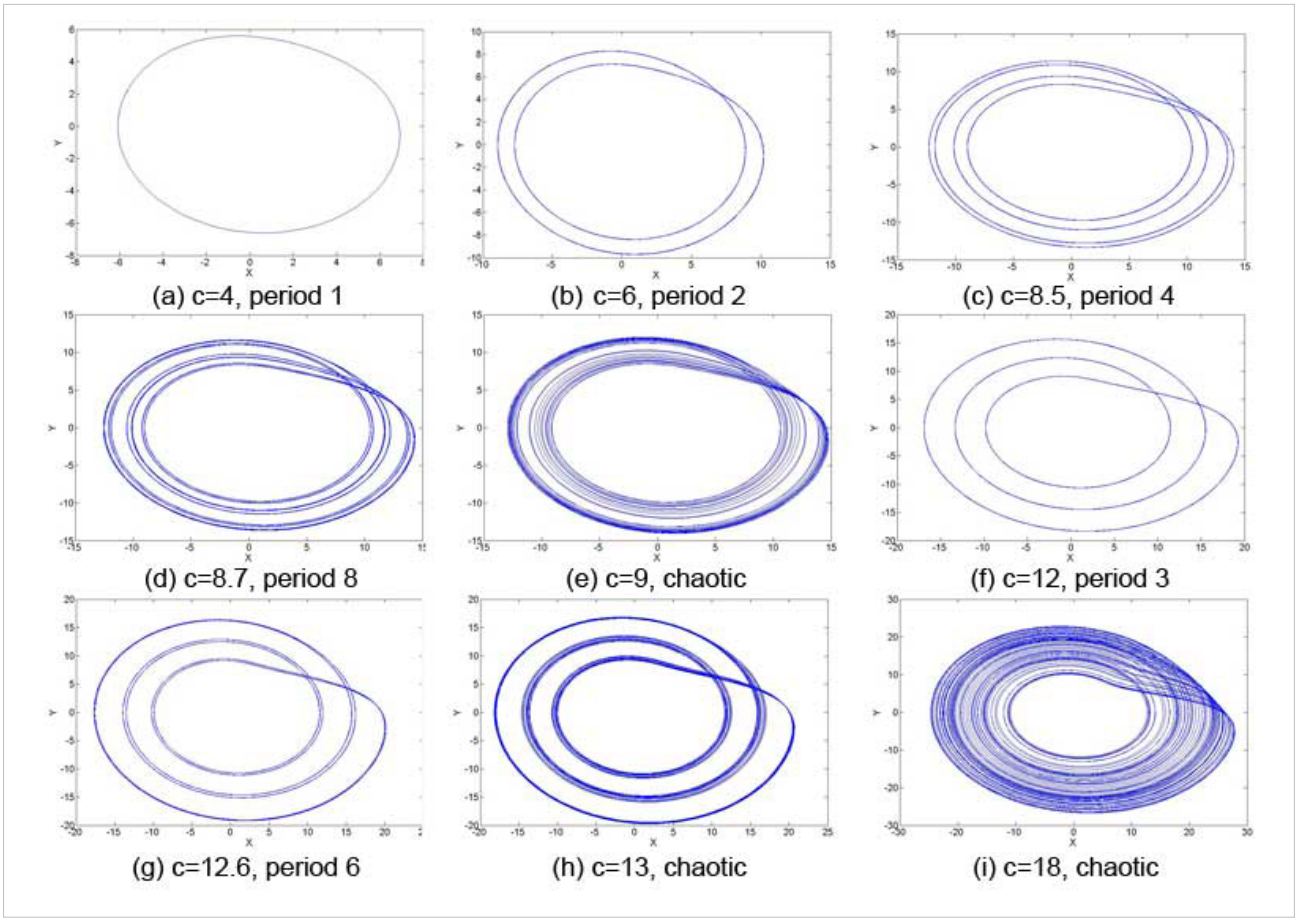
Varying c



Bifurcation diagram for the Rössler attractor for varying c

The traditional bifurcation diagram for the Rössler attractor is created by varying c with $a = b = .1$. This bifurcation diagram reveals that low values of c are periodic, but quickly become chaotic as c increases. This pattern repeats itself as c increases – there are sections of periodicity interspersed with periods of chaos, although the trend is towards higher order periodic orbits in the periodic sections as c increases. For example, the period one orbit only appears for values of c around 4 and is never found again in the bifurcation diagram. The same phenomena is seen with period three; until $c = 12$, period three orbits can be found, but thereafter, they do not appear.

A graphical illustration of the changing attractor over a range of c values illustrates the general behavior seen for all of these parameter analyses – the frequent transitions from ranges of relative stability and periodicity to completely chaotic and back again.



The above set of images illustrates the variations in the post-transient Rössler system as c is varied over a range of values. These images were generated with $a = b = .1$ (a) $c = 4$, periodic orbit. (b) $c = 6$, period-2 orbit. (c) $c = 8.5$, period-4 orbit. (d) $c = 8.7$, period-8 orbit. (e) $c = 9$, sparse chaotic attractor. (f) $c = 12$, period-3 orbit. (g) $c = 12.6$, period-6 orbit. (h) $c = 13$, sparse chaotic attractor. (i) $c = 18$, filled-in chaotic attractor.

Links to other topics

The banding evident in the Rössler attractor is similar to a Cantor set rotated about its midpoint. Additionally, the half-twist in the Rössler attractor makes it similar to a Möbius strip.

See also

- Lorenz attractor
- List of chaotic maps
- Chaos theory
- Dynamical system
- Fractals
- Otto Rössler

References

- E. N. Lorenz (1963). "Deterministic nonperiodic flow". *J. Atmos. Sci.* **20**: 130–141. doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2 [1].
- O. E. Rössler (1976). "An Equation for Continuous Chaos". *Physics Letters* **57A** (5): 397–398.
- O. E. Rössler (1979). "An Equation for Hyperchaos". *Physics Letters* **71A** (2,3): 155–157.
- Steven H. Strogatz (1994). *Nonlinear Dynamics and Chaos*. Perseus publishing.

External links

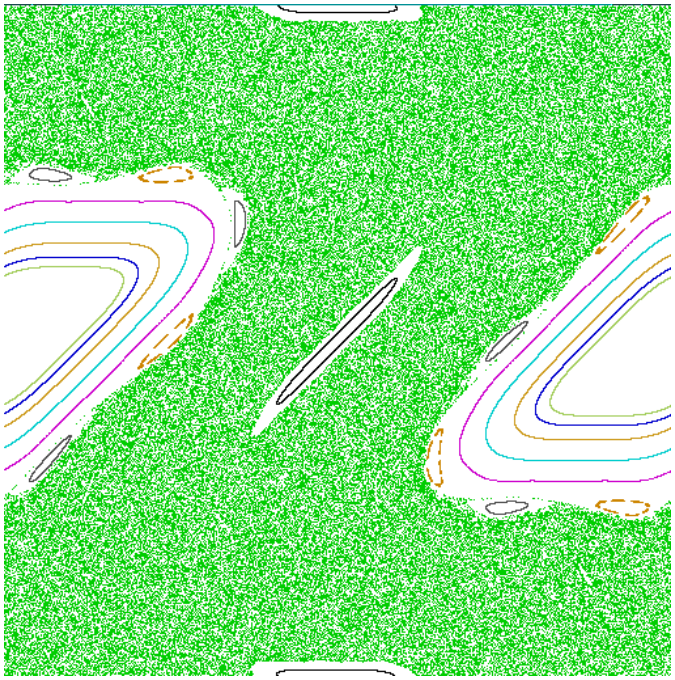
- Flash Animation using PovRay [2]
- Lorenz and Rössler attractors [3] - Java animation
- Java 3D interactive Rössler attractor [4]
- Rössler attractor in Scholarpedia [5]

References

[1] http://dx.doi.org/10.1175%2F1520-0469%281963%29020%3C0130%3ADNF%3E2.0.CO%3B2
[2] http://lagrange.physics.drexel.edu/flash/rossray
[3] http://to-campos.planetaclix.pt/fractal/lorenz_eng.html
[4] http://mrmartin.net/code/RosslerAttractor.html
[5] http://scholarpedia.org/article/Rossler_attractor

Standard map

The **Standard map** (also known as **Chirikov-Taylor map** or **Chirikov standard map**[1]) is an area-preserving chaotic map from a square with side 2π onto itself. It is defined by:



Example of the mapping of ten orbits of the Standard map for $K = 2.0$. The large green region is the main chaotic region of the map.

$$p_{n+1} = p_n + K sin(\theta_n)$$

$$\theta_{n+1} = \theta_n + p_{n+1}$$

where p_n and θ_n are taken modulo 2π. This map describes the motion of a simple mechanical system called a *kicked rotator*. This is made by a stick that is free of the gravitational force, which can rotate frictionless in a plane around an axis located in one of its tips, and which is periodically kicked on the other tip. The variables θ_n and p_n respectively determine the angular position of the stick and its angular momentum after the *n*-th kick. The constant *K* measures the intensity of the kicks.

Besides the kicked rotator, the standard map also describes other systems in the fields of mechanics of particles, accelerator physics, plasma physics, and solid state physics. However, this map is interesting from a fundamental point of view in physics and mathematics because it is a very simple model of a conservative system that displays hamiltonian chaos. It is therefore useful to study the development of chaos in this kind of system.

For $K = 0$ the map is linear and only periodic and quasiperiodic orbits are allowed. When plotted in phase space (the θ–*p* plane), periodic orbits appear as closed curves, and quasiperiodic orbits as necklaces of closed curves whose centers lie in another larger closed curve. Which type of orbit is observed depends on the map's initial conditions.

Nonlinearity of the map increases with *K*, and with it the possibility to observe chaotic dynamics for appropriate initial conditions. This is illustrated in the figure, which displays a collection of different orbits allowed to the standard map for a value of $K > 0$. Each orbit starts from a different initial condition, and different colors are used to distinguish the distinct orbits. All the orbits shown are periodic or quasiperiodic, with the exception of the green one that is chaotic and develops in a large region of phase space as an apparently random set of points.

History

The properties of chaos of the standard map were established by Boris Chirikov in 1969. See more details at Chirikov standard map [2].

Notes

[1] Scholarpedia entry (http://www.scholarpedia.org/article/Chirikov_standard_map)
[2] http://www.scholarpedia.org/article/Chirikov_standard_map

Books

- Lichtenberg, A.J. and Lieberman, M.A. (1992). *Regular and Chaotic Dynamics*. Springer, Berlin. ISBN 978-0-387-97745-4. Springer link (http://www.springer.com/math/analysis/book/978-0-387-97745-4)
- Ott, Edward (2002). *Chaos in Dynamical Systems*. Cambridge University Press New, York. ISBN 0-521-01084-5.
- Sprott, Julien Clinton (2003). *Chaos and Time-Series Analysis*. Oxford University Press. ISBN 0-19-850840-9.

External links

- Standard map (http://mathworld.wolfram.com/StandardMap.html) at MathWorld
- Chirikov standard map (http://www.scholarpedia.org/article/Chirikov_standard_map) at Scholarpedia (http://www.scholarpedia.org)
- Website dedicated to Boris Chirikov (http://www.quantware.ups-tlse.fr/chirikov/)
- Interactive Java Applet visualizing orbits of the Standard Map (http://complexity.xozzox.de/nonlinmappings.html), by Achim Luhn

Synchronization of chaos

Synchronization of chaos is a phenomenon that may occur when two, or more, chaotic oscillators are coupled, or when a chaotic oscillator drives another chaotic oscillator. Because of the butterfly effect, which causes the exponential divergence of the trajectories of two identical chaotic system started with nearly the same initial conditions, having two chaotic system evolving in synchrony might appear quite surprising. However, synchronization of coupled or driven chaotic oscillators is a phenomenon well established experimentally and reasonably understood theoretically.

It has been found that chaos synchronization is quite a rich phenomenon that may present a variety of forms. When two chaotic oscillators are considered, these include:

- **Identical synchronization**. This is a straightforward form of synchronization that may occur when two identical chaotic oscillators are mutually coupled, or when one of them drives the other. If $(x_1,x_2,,...,x_n)$ and $(x'_1, x'_2,...,x'_n)$ denote the set of dynamical variables that describe the state of the first and second oscillator, respectively, it is said that identical synchronization occurs when there is a set of initial conditions $[x_1(0), x_2(0),...,x_n(0)]$, $[x'_1(0), x'_2(0),...,x'_n(0)]$ such that, denoting the time by t, $|x'_i(t)-x_i((t)|\to 0$, for i=1,2,...,n, when $t\to\infty$. That means that for time large enough the dynamics of the two oscillators verifies $x'_i(t)=x_i(t)$, for i=1,2,...,n, in a good approximation. This is called the synchronized state in the sense of identical synchronization.
- **Generalized synchronization**. This type of synchronization occurs mainly when the coupled chaotic oscillators are different, although it has also been reported between identical oscillators. Given the dynamical variables $(x_1,x_2,,...,x_n)$ and $(y_1,y_2,,...,y_m)$ that determine the state of the oscillators, generalized synchronization occurs when there is a functional, Φ, such that, after a transitory evolution from appropriate initial conditions, it is $[y_1(t), y_2(t),...,y_m(t)]=\Phi[x_1(t), x_2(t),...,x_n(t)]$. This means that the dynamical state of one of the oscillators is completely determined by the state of the other. When the oscillators are mutually coupled this functional has to be invertible, if there is a drive-response configuration the drive determines the evolution of the response, and Φ does not need to be invertible. Identical synchronization is the particular case of generalized synchronization when Φ is the identity.
- **Phase synchronization**. This form of synchronization, which occurs when the oscillators coupled are not identical, is partial in the sense that, in the synchronized state, the amplitudes of the oscillator remain unsynchronized, and only their phases evolve in synchrony. Observation of phase synchronization requires a previous definition of the phase of a chaotic oscillator. In many practical cases, it is possible to find a plane in

phase space in which the projection of the trajectories of the oscillator follows a rotation around a well-defined center. If this is the case, the phase is defined by the angle, $\varphi(t)$, described by the segment joining the center of rotation and the projection of the trajectory point onto the plane. In other cases it is still possible to define a phase by means of techniques provided by the theory of signal processing, such as the Hilbert transform. In any case, if $\varphi_1(t)$ and $\varphi_2(t)$ denote the phases of the two coupled oscillators, synchronization of the phase is given by the relation $n\varphi_1(t)=m\varphi_2(t)$ with m and n whole numbers.

- **Anticipated and lag synchronization**. In these cases the synchronized state is characterized by a time interval τ such that the dynamical variables of the oscillators, $(x_1,x_2,,...,x_n)$ and $(x'_1, x'_2,...,x'_n)$, are related by $x'_i(t)=x_i(t+\tau)$; this means that the dynamics of one of the oscillators follows, or anticipates, the dynamics of the other. **Anticipated synchronization** may occur between chaotic oscillators whose dynamics is described by delay differential equations, coupled in a drive-response configuration. In this case, the response anticipates de dynamics of the drive. **Lag synchronization** may occur when the strength of the coupling between phase-synchronized oscillators is increased.
- **Amplitude envelope synchronization**. This is a mild form of synchronization that may appear between two weakly coupled chaotic oscillators. In this case, there is no correlation between phases nor amplitudes; instead, the oscillations of the two systems develop a periodic envelope that has the same frequency in the two systems. This has the same order of magnitude than the difference between the average frequencies of oscillation of the two chaotic oscillator. Often, amplitude envelope synchronization precedes phase synchronization in the sense that when the strength of the coupling between two amplitude envelope synchronized oscillators is increased, phase synchronization develops.

All these forms of synchronization share the property of asymptotic stability. This means that once the synchronized state has been reached, the effect of a small perturbation that destroys synchronization is rapidly damped, and synchronization is recovered again. Mathematically, asymptotic stability is characterized by a positive Lyapunov exponent of the system composed of the two oscillators, which becomes negative when chaotic synchronization is achieved.

Some chaotic systems allow even stronger control of chaos.

Books

- Pikovsky, A.; Rosemblum, M.; Kurths, J. (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press. ISBN 0-521-53352-X.
- González-Miranda, J. M. (2004). *Synchronization and Control of Chaos. An introduction for scientists and engineers*. Imperial College Press. ISBN 1-86094-488-4.

Complex Systems Methods and Modeling

Molecular dynamics

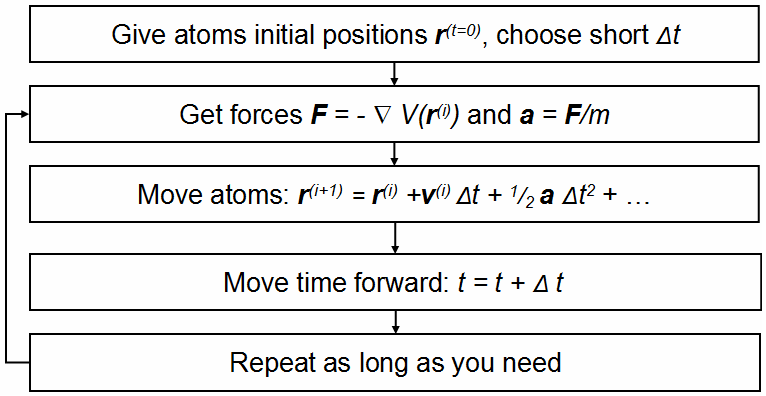
Molecular dynamics (**MD**) is a form of computer simulation in which atoms and molecules are allowed to interact for a period of time by approximations of known physics, giving a view of the motion of the atoms. Because molecular systems generally consist of a vast number of particles, it is impossible to find the properties of such complex systems analytically. When the number of bodies are more than two no analytical solutions can be found and result in chaotic motion (see n-body problem). MD simulation circumvents this problem by using numerical methods. It represents an interface between laboratory experiments and theory, and can be understood as a "virtual experiment". MD probes the relationship between molecular structure, movement and function. Molecular dynamics is a multidisciplinary method. Its laws and theories stem from mathematics, physics, and chemistry, and it employs algorithms from computer science and information theory. It was originally conceived within theoretical physics in the late 1950s[1] and early 1960s [2] , but is applied today mostly in materials science and modeling of biomolecules.

Before it became possible to simulate molecular dynamics with computers, some undertook the hard work of trying it with physical models such as macroscopic spheres. The idea was to arrange them to replicate the properties of a liquid. J.D. Bernal said, in 1962: "... I took a number of rubber balls and stuck them together with rods of a selection of different lengths ranging from 2.75 to 4 inches. I tried to do this in the first place as casually as possible, working in my own office, being interrupted every five minutes or so and not remembering what I had done before the interruption."[3] Fortunately, now computers keep track of bonds during a simulation.

Molecular dynamics is a specialized discipline of molecular modeling and computer simulation based on statistical mechanics; the main justification of the MD method is that statistical ensemble averages are equal to time averages of the system, known as the ergodic hypothesis. MD has also been termed "statistical mechanics by numbers" and "Laplace's vision of Newtonian mechanics" of predicting the future by animating nature's forces[4] [5] and allowing insight into molecular motion on an atomic scale. However, long MD simulations are mathematically ill-conditioned, generating cumulative errors in numerical integration that can be minimized with proper selection of algorithms and parameters, but not eliminated entirely. Furthermore, current potential functions are, in many cases, not sufficiently accurate to reproduce the dynamics of molecular systems, so the much more computationally demanding Ab Initio Molecular Dynamics method must be used. Nevertheless, molecular dynamics techniques allow detailed time and space resolution into representative behavior in phase space.

Areas of Application

There is a significant difference between the focus and methods used by chemists and physicists, and this is reflected in differences in the jargon used by the different fields. In chemistry and biophysics, the interaction between the particles is either described by a "force field" (**classical MD**), a quantum chemical model, or a mix between the two. These terms are not used in physics, where the interactions are usually described by the name of the theory or approximation being used and called the potential energy, or just "potential".



Highly simplified description of the molecular dynamics simulation algorithm. The simulation proceeds iteratively by alternatively calculating forces and solving the equations of motion based on the accelerations obtained from the new forces. In practise, almost all MD codes use much more complicated versions of the algorithm, including two steps (predictor and corrector) in solving the equations of motion and many additional steps for e.g. temperature and pressure control, analysis and output.

Beginning in theoretical physics, the method of MD gained popularity in materials science and since the 1970s also in biochemistry and biophysics. In chemistry, MD serves as an important tool in protein structure determination and refinement using experimental tools such as X-ray crystallography and NMR. It has also been applied with limited success as a method of refining protein structure predictions. In physics, MD is used to examine the dynamics of atomic-level phenomena that cannot be observed directly, such as thin film growth and ion-subplantation. It is also used to examine the physical properties of nanotechnological devices that have not or cannot yet be created.

In applied mathematics and theoretical physics, molecular dynamics is a part of the research realm of dynamical systems, ergodic theory and statistical mechanics in general. The concepts of energy conservation and molecular entropy come from thermodynamics. Some techniques to calculate conformational entropy such as principal components analysis come from information theory. Mathematical techniques such as the transfer operator become applicable when MD is seen as a Markov chain. Also, there is a large community of mathematicians working on volume preserving, symplectic integrators for more computationally efficient MD simulations.

MD can also be seen as a special case of the discrete element method (DEM) in which the particles have spherical shape (e.g. with the size of their van der Waals radii.) Some authors in the DEM community employ the term MD rather loosely, even when their simulations do not model actual molecules.

Design Constraints

Design of a molecular dynamics simulation should account for the available computational power. Simulation size (n=number of particles), timestep and total time duration must be selected so that the calculation can finish within a reasonable time period. However, the simulations should be long enough to be relevant to the time scales of the natural processes being studied. To make statistically valid conclusions from the simulations, the time span simulated should match the kinetics of the natural process. Otherwise, it is analogous to making conclusions about how a human walks from less than one footstep. Most scientific publications about the dynamics of proteins and DNA use data from simulations spanning nanoseconds (1E-9 s) to microseconds (1E-6 s). To obtain these simulations, several CPU-days to CPU-years are needed. Parallel algorithms allow the load to be distributed among CPUs; an example is the spatial decomposition in LAMMPS.

During a classical MD simulation, the most CPU intensive task is the evaluation of the potential (force field) as a function of the particles' internal coordinates. Within that energy evaluation, the most expensive one is the non-bonded or non-covalent part. In Big O notation, common molecular dynamics simulations scale by $O(n^2)$ if all pair-wise electrostatic and van der Waals interactions must be accounted for explicitly. This computational cost can be reduced by employing electrostatics methods such as Particle Mesh Ewald ($O(n\log(n))$) or good spherical cutoff techniques ($O(n)$).

Another factor that impacts total CPU time required by a simulation is the size of the integration timestep. This is the time length between evaluations of the potential. The timestep must be chosen small enough to avoid discretization errors (i.e. smaller than the fastest vibrational frequency in the system). Typical timesteps for classical MD are in the order of 1 femtosecond (1E-15 s). This value may be extended by using algorithms such as SHAKE, which fix the vibrations of the fastest atoms (e.g. hydrogens) into place. Multiple time scale methods have also been developed, which allow for extended times between updates of slower long-range forces.[6] [7] [8]

For simulating molecules in a solvent, a choice should be made between explicit solvent and implicit solvent. Explicit solvent particles (such as the TIP3P and SPC/E water models) must be calculated expensively by the force field, while implicit solvents use a mean-field approach. Using an explicit solvent is computationally expensive, requiring inclusion of about ten times more particles in the simulation. But the granularity and viscosity of explicit solvent is essential to reproduce certain properties of the solute molecules. This is especially important to reproduce kinetics.

In all kinds of molecular dynamics simulations, the simulation box size must be large enough to avoid boundary condition artifacts. Boundary conditions are often treated by choosing fixed values at the edges, or by employing periodic boundary conditions in which one side of the simulation loops back to the opposite side, mimicking a bulk phase.

Microcanonical ensemble (NVE)

In the **microcanonical**, or **NVE** ensemble, the system is isolated from changes in moles (N), volume (V) and energy (E). It corresponds to an adiabatic process with no heat exchange. A microcanonical molecular dynamics trajectory may be seen as an exchange of potential and kinetic energy, with total energy being conserved. For a system of N particles with coordinates X and velocities V, the following pair of first order differential equations may be written in Newton's notation as

$$F(X) = -\nabla U(X) = M\dot{V}(t)$$
$$V(t) = \dot{X}(t).$$

The potential energy function $U(X)$ of the system is a function of the particle coordinates X. It is referred to simply as the "potential" in Physics, or the "force field" in Chemistry. The first equation comes from Newton's laws; the force F acting on each particle in the system can be calculated as the negative gradient of $U(X)$.

For every timestep, each particle's position X and velocity V may be integrated with a symplectic method such as Verlet. The time evolution of X and V is called a trajectory. Given the initial positions (e.g. from theoretical knowledge) and velocities (e.g. randomized Gaussian), we can calculate all future (or past) positions and velocities.

One frequent source of confusion is the meaning of temperature in MD. Commonly we have experience with macroscopic temperatures, which involve a huge number of particles. But temperature is a statistical quantity. If there is a large enough number of atoms, statistical temperature can be estimated from the *instantaneous temperature*, which is found by equating the kinetic energy of the system to $nk_BT/2$ where n is the number of degrees of freedom of the system.

A temperature-related phenomenon arises due to the small number of atoms that are used in MD simulations. For example, consider simulating the growth of a copper film starting with a substrate containing 500 atoms and a deposition energy of 100 eV. In the real world, the 100 eV from the deposited atom would rapidly be transported through and shared among a large number of atoms (10^{10} or more) with no big change in temperature. When there are only 500 atoms, however, the substrate is almost immediately vaporized by the deposition. Something similar happens in biophysical simulations. The temperature of the system in NVE is naturally raised when macromolecules such as proteins undergo exothermic conformational changes and binding.

Canonical ensemble (NVT)

In the canonical ensemble, moles (N), volume (V) and temperature (T) are conserved. It is also sometimes called constant temperature molecular dynamics (CTMD). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat.

A variety of thermostat methods are available to add and remove energy from the boundaries of an MD system in a realistic way, approximating the canonical ensemble. Popular techniques to control temperature include the Nosé-Hoover thermostat, the Berendsen thermostat, and Langevin dynamics. Note that the Berendsen thermostat might introduce the flying ice cube effect, which leads to unphysical translations and rotations of the simulated system.

Isothermal-Isobaric (NPT) ensemble

In the isothermal-isobaric ensemble, moles (N), pressure (P) and temperature (T) are conserved. In addition to a thermostat, a barostat is needed. It corresponds most closely to laboratory conditions with a flask open to ambient temperature and pressure.

In the simulation of biological membranes, isotropic pressure control is not appropriate. For lipid bilayers, pressure control occurs under constant membrane area (NPAT) or constant surface tension "gamma" (NPγT).

Generalized ensembles

The replica exchange method is a generalized ensemble. It was originally created to deal with the slow dynamics of disordered spin systems. It is also called parallel tempering. The replica exchange MD (REMD) formulation [9] tries to overcome the multiple-minima problem by exchanging the temperature of non-interacting replicas of the system running at several temperatures.

Potentials in MD simulations

A molecular dynamics simulation requires the definition of a potential function, or a description of the terms by which the particles in the simulation will interact. In chemistry and biology this is usually referred to as a force field. Potentials may be defined at many levels of physical accuracy; those most commonly used in chemistry are based on molecular mechanics and embody a classical treatment of particle-particle interactions that can reproduce structural and conformational changes but usually cannot reproduce chemical reactions.

The reduction from a fully quantum description to a classical potential entails two main approximations. The first one is the Born-Oppenheimer approximation, which states that the dynamics of electrons is so fast that they can be considered to react instantaneously to the motion of their nuclei. As a consequence, they may be treated separately. The second one treats the nuclei, which are much heavier than electrons, as point particles that follow classical Newtonian dynamics. In classical molecular dynamics the effect of the electrons is approximated as a single potential energy surface, usually representing the ground state.

When finer levels of detail are required, potentials based on quantum mechanics are used; some techniques attempt to create hybrid classical/quantum potentials where the bulk of the system is treated classically but a small region is treated as a quantum system, usually undergoing a chemical transformation.

Empirical potentials

Empirical potentials used in chemistry are frequently called force fields, while those used in materials physics are called just empirical or analytical potentials.

Most force fields in chemistry are empirical and consist of a summation of bonded forces associated with chemical bonds, bond angles, and bond dihedrals, and non-bonded forces associated with van der Waals forces and electrostatic charge. Empirical potentials represent quantum-mechanical effects in a limited way through ad-hoc functional approximations. These potentials contain free parameters such as atomic charge, van der Waals parameters reflecting estimates of atomic radius, and equilibrium bond length, angle, and dihedral; these are obtained by fitting against detailed electronic calculations

(quantum chemical simulations) or experimental physical properties such as elastic constants, lattice parameters and spectroscopic measurements.

Because of the non-local nature of non-bonded interactions, they involve at least weak interactions between all particles in the system. Its calculation is normally the bottleneck in the speed of MD simulations. To lower the computational cost, force fields employ numerical approximations such as shifted cutoff radii, reaction field algorithms, particle mesh Ewald summation, or the newer Particle-Particle Particle Mesh (P3M).

Chemistry force fields commonly employ preset bonding arrangements (an exception being *ab-initio* dynamics), and thus are unable to model the process of chemical bond breaking and reactions explicitly. On the other hand, many of the potentials used in physics, such as those based on the bond order formalism can describe several different coordinations of a system and bond breaking. Examples of such potentials include the Brenner potential[10] for hydrocarbons and its further developments for the C-Si-H and C-O-H systems. The ReaxFF potential[11] can be considered a fully reactive hybrid between bond order potentials and chemistry force fields.

Pair potentials vs.many-body potentials

The potential functions representing the non-bonded energy are formulated as a sum over interactions between the particles of the system. The simplest choice, employed in many popular force fields, is the "pair potential", in which the total potential energy can be calculated from the sum of energy contributions between pairs of atoms. An example of such a pair potential is the non-bonded Lennard-Jones potential (also known as the 6-12 potential), used for calculating van der Waals forces.

$$U(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

Another example is the Born (ionic) model of the ionic lattice. The first term in the next equation is Coulomb's law for a pair of ions, the second term is the short-range repulsion explained by Pauli's exclusion principle and the final term is the dispersion interaction term. Usually, a simulation only includes the dipolar term, although sometimes the quadrupolar term is included as well.

$$U_{ij}(r_{ij}) = \sum \frac{z_i z_j}{4\pi\epsilon_0} \frac{1}{r_{ij}} + \sum A_l \exp \frac{-r_{ij}}{p_l} + \sum C_l r_{ij}^{-n_j} + \cdots$$

In many-body potentials, the potential energy includes the effects of three or more particles interacting with each other. In simulations with pairwise potentials, global interactions in the system also exist, but they occur only through pairwise terms. In many-body potentials, the potential energy cannot be found by a sum over pairs of atoms, as these interactions are calculated explicitly as a combination of higher-order terms. In the statistical view, the dependency between the variables cannot in general be expressed using only pairwise products of the degrees of freedom. For example, the Tersoff potential[12] , which was originally used to simulate carbon, silicon and germanium and has since been used for a wide range of other materials, involves a sum over groups of three atoms, with the angles between the atoms being an important factor in the potential. Other examples are the embedded-atom method (EAM)[13] and the Tight-Binding Second Moment Approximation (TBSMA) potentials[14] , where the electron density of states in the region of an atom is calculated from a sum of contributions from surrounding atoms, and the potential energy contribution is then a function of this sum.

Semi-empirical potentials

Semi-empirical potentials make use of the matrix representation from quantum mechanics. However, the values of the matrix elements are found through empirical formulae that estimate the degree of overlap of specific atomic orbitals. The matrix is then diagonalized to determine the occupancy of the different atomic orbitals, and empirical formulae are used once again to determine the energy contributions of the orbitals.

There are a wide variety of semi-empirical potentials, known as tight-binding potentials, which vary according to the atoms being modeled.

Polarizable potentials

Most classical force fields implicitly include the effect of polarizability, e.g. by scaling up the partial charges obtained from quantum chemical calculations. These partial charges are stationary with respect to the mass of the atom. But molecular dynamics simulations can explicitly model polarizability with the introduction of induced dipoles through different methods, such as Drude particles or fluctuating charges. This allows for a dynamic redistribution of charge between atoms which responds to the local chemical environment.

For many years, polarizable MD simulations have been touted as the next generation. For homogenous liquids such as water, increased accuracy has been achieved through the inclusion of polarizability.[15] Some promising results have also been achieved for proteins.[16] However, it is still uncertain how to best approximate polarizability in a simulation.

Ab-initio methods

In classical molecular dynamics, a single potential energy surface (usually the ground state) is represented in the force field. This is a consequence of the Born-Oppenheimer approximation. If excited states, chemical reactions or a more accurate representation is needed, electronic behavior can be obtained from first principles by using a quantum mechanical method, such as Density Functional Theory. This is known as Ab Initio Molecular Dynamics (AIMD). Due to the cost of treating the electronic degrees of freedom, the computational cost of this simulations is much higher than classical molecular dynamics. This implies that AIMD is limited to smaller systems and shorter periods of time.

Ab-initio quantum-mechanical methods may be used to calculate the potential energy of a system on the fly, as needed for conformations in a trajectory. This calculation is usually made in the close neighborhood of the reaction coordinate. Although various approximations may be used, these are based on theoretical considerations, not on empirical fitting. *Ab-Initio* calculations produce a vast amount of information that is not available from empirical methods, such as density of electronic states or other electronic properties. A significant advantage of using *ab-initio* methods is the ability to study reactions that involve breaking or formation of covalent bonds, which correspond to multiple electronic states.

A popular software for *ab-initio* molecular dynamics is the Car-Parrinello Molecular Dynamics (CPMD) package based on the density functional theory.

Hybrid QM/MM

QM (quantum-mechanical) methods are very powerful. However, they are computationally expensive, while the MM (classical or molecular mechanics) methods are fast but suffer from several limitations (require extensive parameterization; energy estimates obtained are not very accurate; cannot be used to simulate reactions where covalent bonds are broken/formed; and are limited in their abilities for providing accurate details regarding the chemical environment). A new class of method has emerged that combines the good points of QM (accuracy) and MM (speed) calculations. These methods are known as mixed or hybrid quantum-mechanical and molecular mechanics methods (hybrid QM/MM). The methodology for such techniques was introduced by Warshel and coworkers. In the recent years have been pioneered by several groups including: Arieh Warshel (University of Southern California), Weitao Yang (Duke University), Sharon Hammes-Schiffer (The Pennsylvania State University), Donald Truhlar and Jiali Gao (University of Minnesota) and Kenneth Merz (University of Florida).

The most important advantage of hybrid QM/MM methods is the speed. The cost of doing classical molecular dynamics (MM) in the most straightforward case scales $O(n^2)$, where N is the number of atoms in the system. This is mainly due to electrostatic interactions term (every particle interacts with every other particle). However, use of cutoff radius, periodic pair-list updates and more recently the variations of the particle-mesh Ewald's (PME) method has reduced this between O(N) to $O(n^2)$. In other words, if a system with twice many atoms is simulated then it would take between twice to four times as much computing power. On the other hand the simplest *ab-initio* calculations typically scale $O(n^3)$ or worse (Restricted Hartree-Fock calculations have been suggested to scale $\sim O(n^{2.7})$). To overcome the limitation, a small part of the system is treated quantum-mechanically (typically active-site of an enzyme) and the remaining system is treated classically.

In more sophisticated implementations, QM/MM methods exist to treat both light nuclei susceptible to quantum effects (such as hydrogens) and electronic states. This allows generation of hydrogen wave-functions (similar to electronic wave-functions). This methodology has been useful in investigating phenomenon such as hydrogen tunneling. One example where QM/MM methods have provided new discoveries is the calculation of hydride transfer in the enzyme liver alcohol dehydrogenase. In this case, tunneling is important for the hydrogen, as it determines the reaction rate.[17]

Coarse-graining and reduced representations

At the other end of the detail scale are coarse-grained and lattice models. Instead of explicitly representing every atom of the system, one uses "pseudo-atoms" to represent groups of atoms. MD simulations on very large systems may require such large computer resources that they cannot easily be studied by traditional all-atom methods. Similarly, simulations of processes on long timescales (beyond about 1 microsecond) are prohibitively expensive, because they require so many timesteps. In these cases, one can sometimes tackle the problem by using reduced representations, which are also called coarse-grained models.

Examples for coarse graining (CG) methods are discontinuous molecular dynamics (CG-DMD)[18] [19] and Go-models[20] . Coarse-graining is done sometimes taking larger pseudo-atoms. Such united atom approximations have been used in MD simulations of biological membranes. The aliphatic tails of lipids are represented by a few pseudo-atoms

by gathering 2-4 methylene groups into each pseudo-atom.

The parameterization of these very coarse-grained models must be done empirically, by matching the behavior of the model to appropriate experimental data or all-atom simulations. Ideally, these parameters should account for both enthalpic and entropic contributions to free energy in an implicit way. When coarse-graining is done at higher levels, the accuracy of the dynamic description may be less reliable. But very coarse-grained models have been used successfully to examine a wide range of questions in structural biology.

Examples of applications of coarse-graining in biophysics:

- protein folding studies are often carried out using a single (or a few) pseudo-atoms per amino acid;
- DNA supercoiling has been investigated using 1-3 pseudo-atoms per basepair, and at even lower resolution;
- Packaging of double-helical DNA into bacteriophage has been investigated with models where one pseudo-atom represents one turn (about 10 basepairs) of the double helix;
- RNA structure in the ribosome and other large systems has been modeled with one pseudo-atom per nucleotide.

The simplest form of coarse-graining is the "united atom" (sometimes called "extended atom") and was used in most early MD simulations of proteins, lipids and nucleic acids. For example, instead of treating all four atoms of a CH_3 methyl group explicitly (or all three atoms of CH_2 methylene group), one represents the whole group with a single pseudo-atom. This pseudo-atom must, of course, be properly parameterized so that its van der Waals interactions with other groups have the proper distance-dependence. Similar considerations apply to the bonds, angles, and torsions in which the pseudo-atom participates. In this kind of united atom representation, one typically eliminates all explicit hydrogen atoms except those that have the capability to participate in hydrogen bonds ("polar hydrogens"). An example of this is the Charmm 19 force-field.

The polar hydrogens are usually retained in the model, because proper treatment of hydrogen bonds requires a reasonably accurate description of the directionality and the electrostatic interactions between the donor and acceptor groups. A hydroxyl group, for example, can be both a hydrogen bond donor and a hydrogen bond acceptor, and it would be impossible to treat this with a single OH pseudo-atom. Note that about half the atoms in a protein or nucleic acid are nonpolar hydrogens, so the use of united atoms can provide a substantial savings in computer time.

Examples of applications

Molecular dynamics is used in many fields of science.

- First macromolecular MD simulation published (1977, Size: 500 atoms, Simulation Time: 9.2 ps=0.0092 ns, Program: CHARMM precursor) Protein: Bovine Pancreatic Trypsine Inhibitor. This is one of the best studied proteins in terms of folding and kinetics. Its simulation published in Nature magazine paved the way for understanding protein motion as essential in function and not just accessory.[21]
- MD is the standard method to treat collision cascades in the heat spike regime, i.e. the effects that energetic neutron and ion irradiation have on solids an solid surfaces.[22] [23]

The following two biophysical examples are not run-of-the-mill MD simulations. They illustrate almost heroic efforts to produce simulations of a system of very large size (a complete virus) and very long simulation times (500 microseconds):

- MD simulation of the complete satellite tobacco mosaic virus (**STMV**) (2006, Size: 1 million atoms, Simulation time: 50 ns, program: NAMD) This virus is a small, icosahedral plant virus which worsens the symptoms of infection by Tobacco Mosaic Virus (TMV). Molecular dynamics simulations were used to probe the mechanisms of viral assembly. The entire STMV particle consists of 60 identical copies of a single protein that make up the viral capsid (coating), and a 1063 nucleotide single stranded RNA genome. One key finding is that the capsid is very unstable when there is no RNA inside. The simulation would take a single 2006 desktop computer around 35 years to complete. It was thus done in many processors in parallel with continuous communication between them.[24]
- Folding Simulations of the Villin Headpiece in All-Atom Detail (2006, Size: 20,000 atoms; Simulation time: 500 μs = 500,000 ns, Program: folding@home) This simulation was run in 200,000 CPU's of participating personal computers around the world. These computers had the folding@home program installed, a large-scale distributed computing effort coordinated by Vijay Pande at Stanford University. The kinetic properties of the Villin Headpiece protein were probed by using many independent, short trajectories run by CPU's without continuous real-time communication. One technique employed was the Pfold value analysis, which measures the probability of folding before unfolding of a specific starting conformation. Pfold gives information about transition state structures and an ordering of conformations along the folding pathway. Each trajectory in a Pfold calculation can be relatively short, but many independent trajectories are needed.[25]

Molecular dynamics algorithms

Integrators

- Verlet-Stoermer integration
- Runge-Kutta integration
- Beeman's algorithm
- Gear predictor - corrector
- Constraint algorithms (for constrained systems)
- Symplectic integrator

Short-range interaction algorithms

- Cell lists
- Verlet list
- Bonded interactions

Long-range interaction algorithms

- Ewald summation
- Particle Mesh Ewald (PME)
- Particle-Particle Particle Mesh P3M
- Reaction Field Method

Parallelization strategies

- Domain decomposition method (Distribution of system data for parallel computing)
- Molecular Dynamics - Parallel Algorithms [26]

Major software for MD simulations

- Abalone (classical, implicit water)
- ABINIT (DFT)
- ACEMD [27] (running on NVIDIA GPUs: heavily optimized with CUDA)
- ADUN [28] (classical, P2P database for simulations)
- AMBER (classical)
- Ascalaph [29] (classical, GPU accelerated)
- CASTEP (DFT)
- CPMD (DFT)
- CP2K [30] (DFT)
- CHARMM (classical, the pioneer in MD simulation, extensive analysis tools)
- COSMOS [31] (classical and hybrid QM/MM, quantum-mechanical atomic charges with BPT)
- Desmond [32] (classical, parallelization with up to thousands of CPU's)
- DL_POLY [33] (classical)
- ESPResSo (classical, coarse-grained, parallel, extensible)
- Fireball [34] (tight-binding DFT)
- GROMACS (classical)
- GROMOS (classical)
- GULP (classical)
- Hippo [35] (classical)
- LAMMPS (classical, large-scale with spatial-decomposition of simulation domain for parallelism)
- MDynaMix (classical, parallel)
- MOLDY [36] (classical, parallel) latest release [37]
- Materials Studio [38] (Forcite MD using COMPASS, Dreiding, Universal, cvff and pcff forcefields in serial or parallel, QMERA (QM+MD), ONESTEP (DFT), etc.)
- MOSCITO (classical)
- NAMD (classical, parallelization with up to thousands of CPU's)
- NEWTON-X [39] (ab initio, surface-hopping dynamics)
- ProtoMol [40] (classical, extensible, includes multigrid electrostatics)
- PWscf (DFT)
- S/PHI/nX [41] (DFT)
- SIESTA (DFT)
- VASP (DFT)
- TINKER (classical)
- YASARA [42] (classical)
- ORAC [43] (classical)
- XMD (classical)

Related software

- VMD - MD simulation trajectories can be visualized and analyzed.
- PyMol - Molecular Visualization software written in python
- Packmol [44] Package for building starting configurations for MD in an automated fashion
- Sirius - Molecular modeling, analysis and visualization of MD trajectories
- esra [45] - Lightweight molecular modeling and analysis library (Java/Jython/Mathematica).
- Molecular Workbench [46] - Interactive molecular dynamics simulations on your desktop
- BOSS - MC in OPLS

Specialized hardware for MD simulations

- Anton - A specialized, massively parallel supercomputer designed to execute MD simulations.
- MDGRAPE - A special purpose system built for molecular dynamics simulations, especially protein structure prediction.

See also

- Molecular graphics
- Molecular modeling
- Computational chemistry
- Energy drift
- Force field in Chemistry
- Force field implementation
- Monte Carlo method
- Molecular Design software
- Molecular mechanics
- Molecular modeling on GPU
- Protein dynamics
- Implicit solvation
- Car-Parrinello method
- Symplectic numerical integration
- Software for molecular mechanics modeling
- Dynamical systems
- Theoretical chemistry
- Statistical mechanics
- Quantum chemistry
- Discrete element method
- List of nucleic acid simulation software

References

[1] Alder, B. J.; T. E. Wainwright (1959). "Studies in Molecular Dynamics. I. General Method". *J. Chem. Phys.* **31** (2): 459. doi: 10.1063/1.1730376 (http://dx.doi.org/10.1063/1.1730376).

[2] A. Rahman (1964). "Correlations in the Motion of Atoms in Liquid Argon". *Phys Rev* **136**: A405-A411. doi: 10.1103/PhysRev.136.A405 (http://dx.doi.org/10.1103/PhysRev.136.A405).

[3] Bernal, J.D. (1964). "The Bakerian lecture, 1962: The structure of liquids". *Proc. R. Soc.* **280**: 299–322. doi: 10.1098/rspa.1964.0147 (http://dx.doi.org/10.1098/rspa.1964.0147).

[4] Schlick, T. (1996). "Pursuing Laplace's Vision on Modern Computers". in J. P. Mesirov, K. Schulten and D. W. Sumners. *Mathematical Applications to Biomolecular Structure and Dynamics, IMA Volumes in Mathematics and Its Applications*. **82**. New York: Springer-Verlag. pp. 218–247. ISBN 978-0387948386.

[5] de Laplace, P. S. (1820) (in French). *Oeuveres Completes de Laplace, Theorie Analytique des Probabilites*. Paris, France: Gauthier-Villars.

[6] Streett WB, Tildesley DJ, Saville G (1978). "Multiple time-step methods in molecular dynamics". *Mol Phys* **35** (3): 639-648. doi: 10.1080/00268977800100471 (http://dx.doi.org/10.1080/00268977800100471).

[7] Tuckerman ME, Berne BJ, Martyna GJ (1991). "Molecular dynamics algorithm for multiple time scales: systems with long range forces". *J Chem Phys* **94** (10): 6811-6815.

[8] Tuckerman ME, Berne BJ, Martyna GJ (1992). "Reversible multiple time scale molecular dynamics". *J Chem Phys* **97** (3): 1990-2001. doi: 10.1063/1.463137 (http://dx.doi.org/10.1063/1.463137).

[9] Sugita, Yuji; Yuko Okamoto (1999). "Replica-exchange molecular dynamics method for protein folding". *Chem Phys Letters* **314**: 141–151. doi: 10.1016/S0009-2614(99)01123-9 (http://dx.doi.org/10.1016/S0009-2614(99)01123-9).

[10] Brenner, D. W. (1990). "Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films". *Phys. Rev. B* **42** (15): 9458. doi: 10.1103/PhysRevB.42.9458 (http://dx.doi.org/10.1103/PhysRevB.42.9458).

[11] van Duin, A.; Siddharth Dasgupta, Francois Lorant and William A. Goddard III (2001). *J. Phys. Chem. A* **105**: 9398.

[12] Tersoff, J. (1989). ""Modeling solid-state chemistry: Interatomic potentials for multicomponent systems". *Phys. Rev. B* **39**: 5566. doi: 10.1103/PhysRevB.39.5566 (http://dx.doi.org/10.1103/PhysRevB.39.5566).

[13] Daw, M. S.; S. M. Foiles and M. I. Baskes (1993). "The embedded-atom method: a review of theory and applications". *Mat. Sci. And Engr. Rep.* **9**: 251. doi: 10.1016/0920-2307(93)90001-U (http://dx.doi.org/10.1016/0920-2307(93)90001-U).

[14] Cleri, F.; V. Rosato (1993). "Tight-binding potentials for transition metals and alloys". *Phys. Rev. B* **48**: 22. doi: 10.1103/PhysRevB.48.22 (http://dx.doi.org/10.1103/PhysRevB.48.22).

[15] Lamoureux G, Harder E, Vorobyov IV, Roux B, MacKerell AD (2006). "A polarizable model of water for molecular dynamics simulations of biomolecules". *Chem Phys Lett* **418**: 245-249. doi: 10.1016/j.cplett.2005.10.135 (http://dx.doi.org/10.1016/j.cplett.2005.10.135).

[16] Patel, S.; MacKerell, Jr. AD; Brooks III, Charles L (2004). "CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model". *J Comput Chem* **25**: 1504–1514. doi: 10.1002/jcc.20077 (http://dx.doi.org/10.1002/jcc.20077).

[17] Billeter, SR; SP Webb, PK Agarwal, T Iordanov, S Hammes-Schiffer (2001). "Hydride Transfer in Liver Alcohol Dehydrogenase: Quantum Dynamics, Kinetic Isotope Effects, and Role of Enzyme Motion". *J Am Chem Soc* **123**: 11262–11272. doi: 10.1021/ja011384b (http://dx.doi.org/10.1021/ja011384b).

[18] Smith, A; CK Hall (2001). "Alpha-Helix Formation: Discontinuous Molecular Dynamics on an Intermediate-Resolution Protein Model". *Proteins* **44**: 344-360.

[19] Ding, F; JM Borreguero, SV Buldyrey, HE Stanley, NV Dokholyan (2003). "Mechanism for the alpha-helix to beta-hairpin transition". *J Am Chem Soc* **53**: 220–228. doi: 10.1002/prot.10468 (http://dx.doi.org/10.1002/prot.10468).

[20] Paci, E; M Vendruscolo, M Karplus (2002). "Validity of Go Models: Comparison with a Solvent-Shielded Empirical Energy Decomposition". *Biophys J* **83**: 3032–3038. doi: 10.1016/S0006-3495(02)75308-3 (http://dx.doi.org/10.1016/S0006-3495(02)75308-3).

[21] McCammon, J; JB Gelin, M Karplus (1977). "Dynamics of folded proteins". *Nature* **267**: 585–590. doi: 10.1038/267585a0 (http://dx.doi.org/10.1038/267585a0).

[22] Averback, R. S.; Diaz de la Rubia, T. (1998). "Displacement damage in irradiated metals and semiconductors". in H. Ehrenfest and F. Spaepen. *Solid State Physics*. **51**. New York: Academic Press. p. 281-402.

[23] R. Smith, ed (1997). *Atomic & ion collisions in solids and at surfaces: theory, simulation and applications*. Cambridge, UK: Cambridge University Press.

[24] Freddolino P, Arkhipov A, Larson SB, McPherson A, Schulten K. "Molecular dynamics simulation of the Satellite Tobacco Mosaic Virus (STMV)" (http://www.ks.uiuc.edu/Research/STMV/). *Theoretical and*

Computational Biophysics Group. University of Illinois at Urbana Champaign. .
[25] The Folding@Home Project (http://folding.stanford.edu/) and recent papers (http://folding.stanford.edu/papers.html) published using trajectories from it. Vijay Pande Group. Stanford University
[26] http://www.cs.sandia.gov/~sjplimp/md.html
[27] http://www.acellera.com/index.php?arg=acemd
[28] http://cbbl.imim.es/Adun
[29] http://www.agilemolecule.com/Products.html
[30] http://cp2k.berlios.de/
[31] http://www.cosmos-software.de/ce_intro.html
[32] http://www.DEShawResearch.com/resources.html
[33] http://www.ccp5.ac.uk/DL_POLY/
[34] http://fireball-dft.org
[35] http://www.biowerkzeug.com/
[36] http://www.ccp5.ac.uk/moldy/moldy.html
[37] http://ccpforge.cse.rl.ac.uk/frs/?group_id=34
[38] http://accelrys.com/products/materials-studio/
[39] http://www.univie.ac.at/newtonx/
[40] http://protomol.sourceforge.net/
[41] http://www.sphinxlib.de
[42] http://www.yasara.org
[43] http://www.chim.unifi.it/orac/
[44] http://www.ime.unicamp.br/~martinez/packmol
[45] http://esra.sourceforge.net/cgi-bin/index.cgi
[46] http://mw.concord.org/modeler/

General references

- M. P. Allen, D. J. Tildesley (1989) Computer simulation of liquids. Oxford University Press. ISBN 0-19-855645-4.
- J. A. McCammon, S. C. Harvey (1987) Dynamics of Proteins and Nucleic Acids. Cambridge University Press. ISBN 0521307503 (hardback).
- D. C. Rapaport (1996) The Art of Molecular Dynamics Simulation. ISBN 0-521-44561-2.
- Frenkel, Daan; Smit, Berend (2002) [2001]. *Understanding Molecular Simulation : from algorithms to applications*. San Diego, California: Academic Press. ISBN 0-12-267351-4.
- J. M. Haile (2001) Molecular Dynamics Simulation: Elementary Methods. ISBN 0-471-18439-X
- R. J. Sadus, *Molecular Simulation of Fluids: Theory, Algorithms and Object-Orientation*, 2002, ISBN 0-444-51082-6
- Oren M. Becker, Alexander D. Mackerell Jr, Benoît Roux, Masakatsu Watanabe (2001) Computational Biochemistry and Biophysics. Marcel Dekker. ISBN 0-8247-0455-X.
- Andrew Leach (2001) Molecular Modelling: Principles and Applications. (2nd Edition) Prentice Hall. ISBN 978-0582382107.
- Tamar Schlick (2002) Molecular Modeling and Simulation. Springer. ISBN 0-387-95404-X.
- William Graham Hoover (1991) Computational Statistical Mechanics, Elsevier, ISBN 0-444-88192-1.

External links

- The Blue Gene Project (http://researchweb.watson.ibm.com/bluegene/) (IBM)
- D. E. Shaw Research (http://deshawresearch.com/) (D. E. Shaw Research)
- Molecular Physics (http://www.tandf.co.uk/journals/titles/00268976.asp)
- Statistical mechanics of Nonequilibrium Liquids (http://www.phys.unsw.edu.au/~gary/book.html) Lecture Notes on non-equilibrium MD
- Introductory Lecture on Classical Molecular Dynamics (http://www.fz-juelich.de/nic-series/volume10/sutmann.pdf) by Dr. Godehard Sutmann, NIC, Forschungszentrum Jülich, Germany
- Introductory Lecture on Ab Initio Molecular Dynamics and Ab Initio Path Integrals (http://www.fz-juelich.de/nic-series/volume10/tuckerman2.pdf) by Mark E. Tuckerman, New York University, USA
- Introductory Lecture on Ab initio molecular dynamics: Theory and Implementation (http://www.fz-juelich.de/nic-series/Volume1/marx.pdf) by Dominik Marx, Ruhr-Universität Bochum and Jürg Hutter, Universität Zürich

Monte Carlo method

Monte Carlo methods are a class of computational algorithms that rely on repeated random sampling to compute their results. Monte Carlo methods are often used when simulating physical and mathematical systems. Because of their reliance on repeated computation and random or pseudo-random numbers, Monte Carlo methods are most suited to calculation by a computer. Monte Carlo methods tend to be used when it is unfeasible or impossible to compute an exact result with a deterministic algorithm.[1]

Monte Carlo simulation methods are especially useful in studying systems with a large number of coupled degrees of freedom, such as fluids, disordered materials, strongly coupled solids, and cellular structures (see cellular Potts model). More broadly, Monte Carlo methods are useful for modeling phenomena with significant uncertainty in inputs, such as the calculation of risk in business. These methods are also widely used in mathematics: a classic use is for the evaluation of definite integrals, particularly multidimensional integrals with complicated boundary conditions. It is a widely successful method in risk analysis when compared to alternative methods or human intuition. When Monte Carlo simulations have been applied in space exploration and oil exploration, actual observations of failures, cost overruns and schedule overruns are routinely better predicted by the simulations than by human intuition or alternative "soft" methods.[2]

The term "Monte Carlo method" was coined in the 1940s by physicists working on nuclear weapon projects in the Los Alamos National Laboratory.[3]

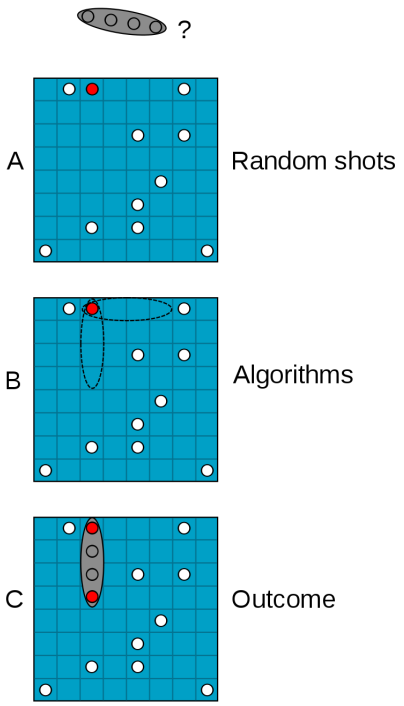
Overview

There is no single Monte Carlo method; instead, the term describes a large and widely-used class of approaches. However, these approaches tend to follow a particular pattern:

1. Define a domain of possible inputs.
2. Generate inputs randomly from the domain.
3. Perform a deterministic computation using the inputs.
4. Aggregate the results of the individual computations into the final result.

For example, the value of π can be approximated using a Monte Carlo method:

1. Draw a square on the ground, then inscribe a circle within it. From plane geometry, the ratio of the area of an inscribed circle to that of the surrounding square is π/4.
2. Uniformly scatter some objects of uniform size throughout the square. For example, grains of rice or sand.
3. Since the two areas are in the ratio π/4, the objects should fall in the areas in approximately the same ratio. Thus, counting the number of objects in the circle and dividing by the total number of objects in the square will yield an approximation for π/4. Multiplying the result by 4 will then yield an approximation for π itself.



The Monte Carlo method can be illustrated as a game of battleship. First a player makes some random shots. Next the player applies algorithms (i.e. a battleship is four dots in the vertical or horizontal direction). Finally based on the outcome of the random sampling and the algorithm the player can determine the likely locations of the other player's ships.

Notice how the π approximation follows the general pattern of Monte Carlo algorithms. First, we define a domain of inputs: in this case, it's the square which circumscribes our circle. Next, we generate inputs randomly (scatter individual grains within the square), then perform a computation on each input (test whether it falls within the circle). At the end, we aggregate the results into our final result, the approximation of π. Note, also, two other common properties of Monte Carlo methods: the computation's reliance on good random numbers, and its slow convergence to a better approximation as more data points are sampled. If grains are purposefully dropped into only, for example, the center of the circle, they will not be uniformly distributed, and so our approximation will be poor. An approximation will also be poor if only a few grains are randomly dropped into the whole square. Thus, the approximation of π will become more accurate both as the grains are dropped more uniformly and as more are dropped.

History

The name "Monte Carlo" was popularized by physics researchers Stanislaw Ulam, Enrico Fermi, John von Neumann, and Nicholas Metropolis, among others; the name is a reference to the Monte Carlo Casino in Monaco where Ulam's uncle would borrow money to gamble.[4] The use of randomness and the repetitive nature of the process are analogous to the activities conducted at a casino.

Random methods of computation and experimentation (generally considered forms of stochastic simulation) can be arguably traced back to the earliest pioneers of probability theory (see, e.g., Buffon's needle, and the work on small samples by William Sealy Gosset), but are more specifically traced to the pre-electronic computing era. The general difference usually described about a Monte Carlo form of simulation is that it systematically "inverts" the typical mode of simulation, treating deterministic problems by *first* finding a probabilistic analog (see Simulated annealing). Previous methods of simulation and statistical sampling generally did the opposite: using simulation to test a previously understood deterministic problem. Though examples of an "inverted" approach do exist historically, they were not considered a general method until the popularity of the Monte Carlo method spread.

Perhaps the most famous early use was by Enrico Fermi in 1930, when he used a random method to calculate the properties of the newly-discovered neutron. Monte Carlo methods were central to the simulations required for the Manhattan Project, though were severely limited by the computational tools at the time. Therefore, it was only after electronic computers were first built (from 1945 on) that Monte Carlo methods began to be studied in depth. In the 1950s they were used at Los Alamos for early work relating to the development of the hydrogen bomb, and became popularized in the fields of physics, physical chemistry, and operations research. The Rand Corporation and the U.S. Air Force were two of the major organizations responsible for funding and disseminating information on Monte Carlo methods during this time, and they began to find a wide application in many different fields.

Uses of Monte Carlo methods require large amounts of random numbers, and it was their use that spurred the development of pseudorandom number generators, which were far quicker to use than the tables of random numbers which had been previously used for statistical sampling.

Applications

As mentioned, Monte Carlo simulation methods are especially useful for modeling phenomena with significant uncertainty in inputs and in studying systems with a large number of coupled degrees of freedom. Specific areas of application include:

Physical sciences

Monte Carlo methods are very important in computational physics, physical chemistry, and related applied fields, and have diverse applications from complicated quantum chromodynamics calculations to designing heat shields and aerodynamic forms. The Monte Carlo method is widely used in statistical physics, in particular, Monte Carlo molecular modeling as an alternative for computational molecular dynamics; see Monte Carlo method in statistical physics. In experimental particle physics, these methods are used for

designing detectors, understanding their behavior and comparing experimental data to theory.

Monte Carlo methods are also used in the ensemble models that form the basis of modern weather forecasting operations.

Design and visuals

Monte Carlo methods have also proven efficient in solving coupled integral differential equations of radiation fields and energy transport, and thus these methods have been used in global illumination computations which produce photorealistic images of virtual 3D models, with applications in video games, architecture, design, computer generated films, special effects in cinema.

Finance and business

Monte Carlo methods in finance are often used to calculate the value of companies, to evaluate investments in projects at corporate level or to evaluate financial derivatives. The Monte Carlo method is intended for financial analysts who want to construct stochastic or probabilistic financial models as opposed to the traditional static and deterministic models. For its use in the insurance industry, see stochastic modelling.

Telecommunications

When planning a wireless network, design must be proved to work for a wide variety of scenarios that depend mainly on the number of users, their locations and the services they want to use. Monte Carlo methods are typically used to generate these users and their states. The network performance is then evaluated and, if results are not satisfactory, the network design goes through an optimization process.

Games

Monte Carlo methods have recently been applied in game playing related artificial intelligence theory. Most notably the game of Go has seen remarkably successful Monte Carlo algorithm based computer players. One of the main problems that this approach has in game playing is that it sometimes misses an isolated, very good move. These approaches are often strong strategically but weak tactically, as tactical decisions tend to rely on a small number of crucial moves which are easily missed by the randomly searching Monte Carlo algorithm.

Monte Carlo simulation versus “what if” scenarios

The opposite of Monte Carlo simulation might be considered deterministic modeling using single-point estimates. Each uncertain variable within a model is assigned a “best guess” estimate. Various combinations of each input variable are manually chosen (such as best case, worst case, and most likely case), and the results recorded for each so-called “what if” scenario. [5]

By contrast, Monte Carlo simulation considers random sampling of probability distribution functions as model inputs to produce hundreds or thousands of possible outcomes instead of a few discrete scenarios. The results provide probabilities of different outcomes occurring. [6] For example, a comparison of a spreadsheet cost construction model run using traditional “what if” scenarios, and then run again with Monte Carlo simulation and

Triangular probability distributions shows that the Monte Carlo analysis has a narrower range than the “what if” analysis. This is because the “what if” analysis gives equal weight to all scenarios.[7]

For an application, see quantifying uncertainty under corporate finance.

Use in mathematics

In general, Monte Carlo methods are used in mathematics to solve various problems by generating suitable random numbers and observing that fraction of the numbers obeying some property or properties. The method is useful for obtaining numerical solutions to problems which are too complicated to solve analytically. The most common application of the Monte Carlo method is Monte Carlo integration.

Integration

Deterministic methods of numerical integration operate by taking a number of evenly spaced samples from a function. In general, this works very well for functions of one variable. However, for functions of vectors, deterministic quadrature methods can be very inefficient. To numerically integrate a function of a two-dimensional vector, equally spaced grid points over a two-dimensional surface are required. For instance a 10x10 grid requires 100 points. If the vector has 100 dimensions, the same spacing on the grid would require 10^{100} points—far too many to be computed. 100 dimensions is by no means unreasonable, since in many physical problems, a "dimension" is equivalent to a degree of freedom. (See Curse of dimensionality.)

Monte Carlo methods provide a way out of this exponential time-increase. As long as the function in question is reasonably well-behaved, it can be estimated by randomly selecting points in 100-dimensional space, and taking some kind of average of the function values at these points. By the law of large numbers, this method will display $1/\sqrt{N}$ convergence—i.e. quadrupling the number of sampled points will halve the error, regardless of the number of dimensions.

A refinement of this method is to somehow make the points random, but more likely to come from regions of high contribution to the integral than from regions of low contribution. In other words, the points should be drawn from a distribution similar in form to the integrand. Understandably, doing this precisely is just as difficult as solving the integral in the first place, but there are approximate methods available: from simply making up an integrable function thought to be similar, to one of the adaptive routines discussed in the topics listed below.

A similar approach involves using low-discrepancy sequences instead—the quasi-Monte Carlo method. Quasi-Monte Carlo methods can often be more efficient at numerical integration because the sequence "fills" the area better in a sense and samples more of the most important points that can make the simulation converge to the desired solution more quickly.

Integration methods

- Direct sampling methods
  - Importance sampling
  - Stratified sampling
  - Recursive stratified sampling
  - VEGAS algorithm
- Random walk Monte Carlo including Markov chains
  - Metropolis-Hastings algorithm
- Gibbs sampling

Optimization

Another powerful and very popular application for random numbers in numerical simulation is in numerical optimization. These problems use functions of some often large-dimensional vector that are to be minimized (or maximized). Many problems can be phrased in this way: for example a computer chess program could be seen as trying to find the optimal set of, say, 10 moves which produces the best evaluation function at the end. The traveling salesman problem is another optimization problem. There are also applications to engineering design, such as multidisciplinary design optimization.

Most Monte Carlo optimization methods are based on random walks. Essentially, the program will move around a marker in multi-dimensional space, tending to move in directions which lead to a lower function, but sometimes moving against the gradient.

Optimization methods

- Evolution strategy
- Genetic algorithms
- Parallel tempering
- Simulated annealing
- Stochastic optimization
- Stochastic tunneling

Inverse problems

Probabilistic formulation of inverse problems leads to the definition of a probability distribution in the model space. This probability distribution combines *a priori* information with new information obtained by measuring some observable parameters (data). As, in the general case, the theory linking data with model parameters is nonlinear, the a posteriori probability in the model space may not be easy to describe (it may be multimodal, some moments may not be defined, etc.).

When analyzing an inverse problem, obtaining a maximum likelihood model is usually not sufficient, as we normally also wish to have information on the resolution power of the data. In the general case we may have a large number of model parameters, and an inspection of the marginal probability densities of interest may be impractical, or even useless. But it is possible to pseudorandomly generate a large collection of models according to the posterior probability distribution and to analyze and display the models in such a way that information on the relative likelihoods of model properties is conveyed to the spectator. This can be accomplished by means of an efficient Monte Carlo method, even in cases where no explicit formula for the a priori distribution is available.

The best-known importance sampling method, the Metropolis algorithm, can be generalized, and this gives a method that allows analysis of (possibly highly nonlinear) inverse problems with complex a priori information and data with an arbitrary noise distribution. For details, see Mosegaard and Tarantola (1995),[8] or Tarantola (2005).[9]

Computational mathematics

Monte Carlo methods are useful in many areas of computational mathematics, where a *lucky choice* can find the correct result. A classic example is Rabin's algorithm for primality testing: for any n which is not prime, a random x has at least a 75% chance of proving that n is not prime. Hence, if n is not prime, but x says that it might be, we have observed at most a 1-in-4 event. If 10 different random x say that "n is probably prime" when it is not, we have observed a one-in-a-million event. In general a Monte Carlo algorithm of this kind produces one correct answer with a guarantee ***n* is composite, and *x* proves it so**, but another one without, but with a guarantee of not getting this answer when it is wrong **too often** — in this case at most 25% of the time. See also Las Vegas algorithm for a related, but different, idea.

Monte Carlo and random numbers

Interestingly, Monte Carlo simulation methods do not always require truly random numbers to be useful — while for some applications, such as primality testing, unpredictability is vital (see Davenport (1995)).[10] Many of the most useful techniques use deterministic, pseudo-random sequences, making it easy to test and re-run simulations. The only quality usually necessary to make good simulations is for the pseudo-random sequence to appear "random enough" in a certain sense.

What this means depends on the application, but typically they should pass a series of statistical tests. Testing that the numbers are uniformly distributed or follow another desired distribution when a large enough number of elements of the sequence are considered is one of the simplest, and most common ones.

See also

General

- Auxiliary field Monte Carlo
- Bootstrapping (statistics)
- Demon algorithm
- Evolutionary Computation
- Las Vegas algorithm
- Markov chain
- Molecular dynamics
- Monte Carlo option model
- Monte Carlo integration
- Quasi-Monte Carlo method
- Random number generator
- Randomness
- Resampling (statistics)

Application areas

- Graphics, particularly for ray tracing; a version of the Metropolis-Hastings algorithm is also used for ray tracing where it is known as Metropolis light transport
- Modeling light transport in biological tissue
- Monte Carlo methods in finance
- Reliability engineering
- In simulated annealing for protein structure prediction
- In semiconductor device research, to model the transport of current carriers
- Environmental science, dealing with contaminant behavior
- Search And Rescue and Counter-Pollution. Models used to predict the drift of a life raft or movement of an oil slick at sea.
- In probabilistic design for simulating and understanding the effects of variability
- In physical chemistry, particularly for simulations involving atomic clusters
- In biomolecular simulations
- In polymer physics
  - Bond fluctuation model
- In computer science
  - Las Vegas algorithm
  - LURCH
  - Computer go
  - General Game Playing
- Modeling the movement of impurity atoms (or ions) in plasmas in existing and tokamaks (e.g.: DIVIMP).
- Nuclear and particle physics codes using the Monte Carlo method:
  - GEANT — CERN's simulation of high energy particles interacting with a detector.
  - CompHEP, PYTHIA — Monte-Carlo generators of particle collisions
  - MCNP(X) - LANL's radiation transport codes
  - MCU: universal computer code for simulation of particle transport (neutrons, photons, electrons) in three-dimensional systems by means of the Monte Carlo method
  - EGS — Stanford's simulation code for coupled transport of electrons and photons
  - PEREGRINE: LLNL's Monte Carlo tool for radiation therapy dose calculations
  - BEAMnrc — Monte Carlo code system for modeling radiotherapy sources (LINAC's)
  - PENELOPE — Monte Carlo for coupled transport of photons and electrons, with applications in radiotherapy
  - MONK — Serco Assurance's code for the calculation of k-effective of nuclear systems
- Modelling of foam and cellular structures
- Modeling of tissue morphogenesis
- Computation of holograms
- Phylogenetic analysis, i.e. Bayesian inference, Markov chain Monte Carlo

Other methods employing Monte Carlo

- Assorted random models, e.g. self-organised criticality
- Direct simulation Monte Carlo
- Dynamic Monte Carlo method
- Kinetic Monte Carlo
- Quantum Monte Carlo
- Quasi-Monte Carlo method using low-discrepancy sequences and self avoiding walks
- Semiconductor charge transport and the like
- Electron microscopy beam-sample interactions
- Stochastic optimization
- Cellular Potts model
- Markov chain Monte Carlo
- Cross-entropy method
- Applied information economics
- Monte Carlo localization

Notes

[1] Douglas Hubbard "How to Measure Anything: Finding the Value of Intangibles in Business" pg. 46, John Wiley & Sons, 2007

[2] Douglas Hubbard "The Failure of Risk Management: Why It's Broken and How to Fix It", John Wiley & Sons, 2009

[3] Nicholas Metropolis (1987), " The beginning of the Monte Carlo method (http://library.lanl.gov/la-pubs/00326866.pdf)", *Los Alamos Science* (1987 Special Issue dedicated to Stanislaw Ulam): 125–130,

[4] Douglas Hubbard "How to Measure Anything: Finding the Value of Intangibles in Business" pg. 46, John Wiley & Sons, 2007

[5] David Vose: “Risk Analysis, A Quantitative Guide,” Second Edition, p. 13, John Wiley & Sons, 2000.

[6] Ibid, p. 16

[7] Ibid, p. 17, showing graph

[8] http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/Papers_PDF/MonteCarlo_latex.pdf

[9] http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/SIAM/index.html

[10] Davenport, J. H.. "Primality testing revisited" (http://doi.acm.org/10.1145/143242.143290). doi: http://doi.acm.org/10.1145/143242.143290 (http://dx.doi.org/http://doi.acm.org/10.1145/143242.143290). . Retrieved on 2007-08-19.

References

- Metropolis, N.; Ulam, S. (1949). "The Monte Carlo Method". *Journal of the American Statistical Association* **44** (247): 335–341. doi: 10.2307/2280232 (http://dx.doi.org/10.2307/2280232).
- Metropolis, Nicholas; Rosenbluth, Arianna W.; Rosenbluth, Marshall N.; Teller, Augusta H.; Teller, Edward (1953). "Equation of State Calculations by Fast Computing Machines". *Journal of Chemical Physics* **21** (6): 1087. doi: 10.1063/1.1699114 (http://dx.doi.org/10.1063/1.1699114).
- Hammersley, J. M.; Handscomb, D. C. (1975). *Monte Carlo Methods*. London: Methuen. ISBN 0416523404.
- Kahneman, D.; Tversky, A. (1982). *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Gould, Harvey; Tobochnik, Jan (1988). *An Introduction to Computer Simulation Methods, Part 2, Applications to Physical Systems*. Reading: Addison-Wesley. ISBN 020116504X.

- Binder, Kurt (1995). *The Monte Carlo Method in Condensed Matter Physics*. New York: Springer. ISBN 0387543694.
- Berg, Bernd A. (2004). *Markov Chain Monte Carlo Simulations and Their Statistical Analysis (With Web-Based Fortran Code)*. Hackensack, NJ: World Scientific. ISBN 9812389350.
- Caflisch, R. E. (1998). *Monte Carlo and quasi-Monte Carlo methods*. Acta Numerica. **7**. Cambridge University Press. pp. 1–49.
- Doucet, Arnaud; Freitas, Nando de; Gordon, Neil (2001). *Sequential Monte Carlo methods in practice*. New York: Springer. ISBN 0387951466.
- Fishman, G. S. (1995). *Monte Carlo: Concepts, Algorithms, and Applications*. New York: Springer. ISBN 038794527X.
- MacKeown, P. Kevin (1997). *Stochastic Simulation in Physics*. New York: Springer. ISBN 9813083263.
- Robert, C. P.; Casella, G. (2004). *Monte Carlo Statistical Methods* (2nd ed.). New York: Springer. ISBN 0387212396.
- Rubinstein, R. Y.; Kroese, D. P. (2007). *Simulation and the Monte Carlo Method* (2nd ed.). New York: John Wiley & Sons. ISBN 9780470177938.
- Mosegaard, Klaus; Tarantola, Albert (1995). "Monte Carlo sampling of solutions to inverse problems". *J. Geophys. Res.* **100** (B7): 12431–12447. doi: 10.1029/94JB03097 (http://dx.doi.org/10.1029/94JB03097).
- Tarantola, Albert (2005). *Inverse Problem Theory (http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/SIAM/index.html)*. Philadelphia: Society for Industrial and Applied Mathematics. ISBN 0898715725. http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/SIAM/index.html.

External links

- Overview and reference list (http://mathworld.wolfram.com/MonteCarloMethod.html), Mathworld
- Introduction to Monte Carlo Methods (http://www.ipp.mpg.de/de/for/bereiche/stellarator/Comp_sci/CompScience/csep/csep1.phy.ornl.gov/mc/mc.html), Computational Science Education Project
- Overview of formulas used in Monte Carlo simulation (http://www.sitmo.com/eqcat/15), the Quant Equation Archive, at sitmo.com
- The Basics of Monte Carlo Simulations (http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html), University of Nebraska-Lincoln
- Introduction to Monte Carlo simulation (http://office.microsoft.com/en-us/assistance/HA011118931033.aspx) (for Excel), Wayne L. Winston
- Monte Carlo Methods - Overview and Concept (http://www.brighton-webs.co.uk/montecarlo/concept.asp), brighton-webs.co.uk
- Molecular Monte Carlo Intro (http://www.cooper.edu/engineering/chemechem/monte.html), Cooper Union
- Monte Carlo techniques applied in physics (http://homepages.ed.ac.uk/s0095122/Applet1-page.htm)
- MonteCarlo Simulation in Finance (http://www.global-derivatives.com/maths/k-o.php), global-derivatives.com
- Approximation of π with the Monte Carlo Method (http://twt.mpei.ac.ru/MAS/Worksheets/approxpi.mcd)

- Risk Analysis in Investment Appraisal (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=265905), The Application of Monte Carlo Methodology in Project Appraisal, Savvakis C. Savvides
- Probabilistic Assessment of Structures using the Monte Carlo method (http://en.wikiversity.org/wiki/Probabilistic_Assessment_of_Structures), Wikiuniversity paper for students of Structural Engineering
- A very intuitive and comprehensive introduction to Quasi-Monte Carlo methods (http://www.puc-rio.br/marco.ind/quasi_mc.html)
- Pricing using Monte Carlo simulation (http://knol.google.com/k/giancarlo-vercellino/pricing-using-monte-carlo-simulation/11d5i2rgd9gn5/3#), a practical example, Prof. Giancarlo Vercellino

Software

- The BUGS project (http://www.mrc-bsu.cam.ac.uk/bugs/) (including WinBUGS and OpenBUGS)
- Monte Carlo Simulation, Resampling, Bootstrap tool (http://www.statistics101.net)
- YASAI: Yet Another Simulation Add-In (http://yasai.rutgers.edu/) - Free Monte Carlo Simulation Add-In for Excel created by Rutgers University

Quantum chemistry

Quantum chemistry is a branch of theoretical chemistry, which applies quantum mechanics and quantum field theory to address issues and problems in chemistry. The description of the electronic behavior of atoms and molecules as pertaining to their reactivity is one of the applications of quantum chemistry. Quantum chemistry lies on the border between chemistry and physics, and significant contributions have been made by scientists from both fields. It has a strong and active overlap with the field of atomic physics and molecular physics, as well as physical chemistry.

Quantum chemistry mathematically describes the fundamental behavior of matter at the molecular scale.[1] It is, in principle, possible to describe all chemical systems using this theory. In practice, only the simplest chemical systems may realistically be investigated in purely quantum mechanical terms, and approximations must be made for most practical purposes (e.g., Hartree-Fock, post Hartree-Fock or Density functional theory, see computational chemistry for more details). Hence a detailed understanding of quantum mechanics is not necessary for most chemistry, as the important implications of the theory (principally the orbital approximation) can be understood and applied in simpler terms.

In quantum mechanics the Hamiltonian, or the physical state, of a particle can be expressed as the sum of two operators, one corresponding to kinetic energy and the other to potential energy. The Hamiltonian in the Schrödinger wave equation used in quantum chemistry does not contain terms for the spin of the electron.

Solutions of the Schrödinger equation for the hydrogen atom gives the form of the wave function for atomic orbitals, and the relative energy of the various orbitals. The orbital approximation can be used to understand the other atoms e.g. helium, lithium and carbon.

History

The **history of quantum chemistry** essentially began with the 1838 discovery of cathode rays by Michael Faraday, the 1859 statement of the black body radiation problem by Gustav Kirchhoff, the 1877 suggestion by Ludwig Boltzmann that the energy states of a physical system could be discrete, and the 1900 quantum hypothesis by Max Planck that any energy radiating atomic system can theoretically be divided into a number of discrete energy elements ε such that each of these energy elements is proportional to the frequency ν with which they each individually radiate energy, as defined by the following formula:

$$\epsilon = h\nu$$

where h is a numerical value called Planck's Constant. Then, in 1905, to explain the photoelectric effect (1839), i.e., that shining light on certain materials can function to eject electrons from the material, Albert Einstein postulated, based on Planck's quantum hypothesis, that light itself consists of individual quantum particles, which later came to be called photons (1926). In the years to follow, this theoretical basis slowly began to be applied to chemical structure, reactivity, and bonding.

Electronic structure

The first step in solving a quantum chemical problem is usually solving the Schrödinger equation (or Dirac equation in relativistic quantum chemistry) with the electronic molecular Hamiltonian. This is called determining the **electronic structure** of the molecule. It can be said that the electronic structure of a molecule or crystal implies essentially its chemical properties.

Wave model

The foundation of quantum mechanics and quantum chemistry is the **wave model**, in which the atom is a small, dense, positively charged nucleus surrounded by electrons. Unlike the earlier Bohr model of the atom, however, the wave model describes electrons as "clouds" moving in orbitals, and their positions are represented by probability distributions rather than discrete points. The strength of this model lies in its predictive power. Specifically, it predicts the pattern of chemically similar elements found in the periodic table. The wave model is so named because electrons exhibit properties (such as interference) traditionally associated with waves. See wave-particle duality.

Valence bond

Although the mathematical basis of quantum chemistry had been laid by Schrödinger in 1926, it is generally accepted that the first true calculation in quantum chemistry was that of the German physicists Walter Heitler and Fritz London on the hydrogen (H_2) molecule in 1927. Heitler and London's method was extended by the American theoretical physicist John C. Slater and the American theoretical chemist Linus Pauling to become the **Valence-Bond (VB)** [or **Heitler-London-Slater-Pauling (HLSP)**] method. In this method, attention is primarily devoted to the pairwise interactions between atoms, and this method therefore correlates closely with classical chemists' drawings of bonds.

Molecular orbital

An alternative approach was developed in 1929 by Friedrich Hund and Robert S. Mulliken, in which electrons are described by mathematical functions delocalized over an entire molecule. The **Hund-Mulliken** approach or **molecular orbital (MO) method** is less intuitive to chemists, but has turned out capable of predicting spectroscopic properties better than the VB method. This approach is the conceptional basis of the **Hartree-Fock method** and further post Hartree-Fock methods.

Density functional theory

The **Thomas-Fermi model** was developed independently by Thomas and Fermi in 1927. This was the first attempt to describe many-electron systems on the basis of electronic density instead of wave functions, although it was not very successful in the treatment of entire molecules. The method did provide the basis for what is now known as **density functional theory**. Though this method is less developed than post Hartree-Fock methods, its lower computational requirements allow it to tackle larger polyatomic molecules and even macromolecules, which has made it the most used method in computational chemistry at present.

Chemical dynamics

A further step can consist of solving the Schrödinger equation with the total molecular Hamiltonian in order to study the motion of molecules. Direct solution of the Schrödinger equation is called *quantum molecular dynamics*, within the semiclassical approximation *semiclassical molecular dynamics*, and within the classical mechanics framework *molecular dynamics (MD)*. Statistical approaches, using for example Monte Carlo methods, are also possible.

Adiabatic chemical dynamics

Main article: Adiabatic formalism or Born-Oppenheimer approximation

In **adiabatic dynamics**, interatomic interactions are represented by single scalar potentials called potential energy surfaces. This is the Born-Oppenheimer approximation introduced by Born and Oppenheimer in 1927. Pioneering applications of this in chemistry were performed by Rice and Ramsperger in 1927 and Kassel in 1928, and generalized into the RRKM theory in 1952 by Marcus who took the transition state theory developed by Eyring in 1935 into account. These methods enable simple estimates of unimolecular reaction rates from a few characteristics of the potential surface.

Non-adiabatic chemical dynamics

Non-adiabatic dynamics consists of taking the interaction between several coupled potential energy surface (corresponding to different electronic quantum states of the molecule). The coupling terms are called **vibronic couplings**. The pioneering work in this field was done by Stueckelberg, Landau, and Zener in the 1930s, in their work on what is now known as the Landau-Zener transition. Their formula allows the transition probability between two diabatic potential curves in the neighborhood of an avoided crossing to be calculated.

Quantum chemistry and quantum field theory

The application of quantum field theory (QFT) to chemical systems and theories has become increasingly common in the modern physical sciences. One of the first and most fundamentally explicit appearances of this is seen in the theory of the photomagneton. In this system, plasmas, which are ubiquitous in both physics and chemistry, are studied in order to determine the basic quantization of the underlying bosonic field. However, quantum field theory is of interest in many fields of chemistry, including: nuclear chemistry, astrochemistry, sonochemistry, and quantum hydrodynamics. Field theoretic methods have also been critical in developing the ab initio Effective Hamiltonian theory of semi-empirical pi-electron methods.

See also

- Atomic physics
- Computational chemistry
- Condensed matter physics
- International Academy of Quantum Molecular Science
- Physical chemistry
- Quantum chemistry computer programs
- Quantum electrochemistry
- QMC@Home
- Theoretical physics

Further reading

- Pauling, L. (1954). *General Chemistry*. Dover Publications. ISBN 0-486-65622-5.
- Pauling, L., and Wilson, E. B. *Introduction to Quantum Mechanics with Applications to Chemistry* (Dover Publications) ISBN 0-486-64871-0
- Atkins, P.W. *Physical Chemistry* (Oxford University Press) ISBN 0-19-879285-9
- McWeeny, R. *Coulson's Valence* (Oxford Science Publications) ISBN 0-19-855144-4
- Landau, L.D. and Lifshitz, E.M. *Quantum Mechanics:Non-relativistic Theory* (Course of Theoretical Physics vol.3) (Pergamon Press)
- Bernard Pullman and Alberte Pullman. 1963. *Quantum Biochemistry.*, New York and London: Academic Press.
- Eric R. Scerri, The Periodic Table: Its Story and Its Significance, Oxford University Press, 2006. Considers the extent to which chemistry and especially the periodic system has been reduced to quantum mechanics. ISBN 0-19-530573-6.
- Simon, Z. 1976. *Quantum Biochemistry and Specific Interactions.*, Taylor & Francis; ISBN-13: 978-0856260872 and ISBN 0-85-6260878 .

References

[1] "Quantum Chemistry" (http://cmm.cit.nih.gov/modeling/guide_documents/quantum_mechanics_document.html). *The NIH Guide to Molecular Modeling*. National Institutes of Health. . Retrieved on 2007-09-08.

External links

- The Sherrill Group - Notes (http://vergil.chemistry.gatech.edu/notes/index.html)
- ChemViz Curriculum Support Resources (http://www.shodor.org/chemviz/)
- Early ideas in the history of quantum chemistry (http://www.quantum-chemistry-history.com/)

Nobel lectures by quantum chemists

- Walter Kohn's Nobel lecture (http://nobelprize.org/chemistry/laureates/1998/kohn-lecture.html)
- Rudolph Marcus' Nobel lecture (http://nobelprize.org/chemistry/laureates/1992/marcus-lecture.html)
- Robert Mulliken's Nobel lecture (http://nobelprize.org/chemistry/laureates/1966/mulliken-lecture.html)
- Linus Pauling's Nobel lecture (http://nobelprize.org/chemistry/laureates/1954/pauling-lecture.html)
- John Pople's Nobel lecture (http://nobelprize.org/chemistry/laureates/1998/pople-lecture.html)

Quantum Monte Carlo

Electronic structure methods
Tight binding
Nearly-free electron model
Hartree–Fock
Modern valence bond
Generalized valence bond
Møller–Plesset perturbation theory
Configuration interaction
Coupled cluster
Multi-configurational self-consistent field
Density functional theory
Quantum chemistry composite methods
Quantum Monte Carlo
k·p perturbation theory
Muffin-tin approximation
LCAO method

Quantum Monte Carlo is a large class of computer algorithms that simulate quantum systems with the idea of solving the many-body problem. They use, in one way or another, the Monte Carlo method to handle the many-dimensional integrals that arise. Quantum Monte Carlo allows a direct representation of many-body effects in the wavefunction, at the cost of statistical uncertainty that can be reduced with more simulation time. For bosons, there exist numerically exact and polynomial-scaling algorithms. For fermions, there exist very good approximations and numerically exact exponentially scaling quantum Monte Carlo algorithms, but none that are both.

Background

In principle, any physical system can be described by the many-body Schrödinger equation as long as the constituent particles are not moving "too" fast; that is, they are not moving near the speed of light. This includes the electrons in almost every material in the world, so if we could solve the Schrödinger equation, we could predict the behavior of any electronic system, which has important applications in fields from computers to biology. This also includes the nuclei in Bose–Einstein condensate and superfluids such as liquid helium. The difficulty is that the Schrödinger equation involves a function of three times the number of particles and is difficult to solve even using parallel computing technology in a reasonable amount of time (less than 2 years). Traditionally, theorists have approximated the many-body wave function as an antisymmetric function of one-body orbitals, as shown concisely at this link.[1] This kind of formulation either limits the possible wave functions, as in the case of the Hartree-Fock (HF) approximation, or converges very slowly, as in configuration interaction. One of the reasons for the difficulty with an HF initial estimate (ground state seed, also known as Slater determinant) is that it is very difficult to model the electronic and nuclear cusps in the wavefunction. However, one does not generally model at this point of the approximation. As two particles approach each other, the wavefunction has exactly known derivatives.

Quantum Monte Carlo is a way around these problems because it allows us to model a many-body wavefunction of our choice directly. Specifically, we can use a Hartree-Fock approximation as our starting point but then multiplying it by any symmetric function, of which Jastrow functions are typical, designed to enforce the cusp conditions. Most methods aim at computing the ground-state wavefunction of the system, with the exception of path integral Monte Carlo and finite-temperature auxiliary field Monte Carlo, which calculate the density matrix.

There are several quantum Monte Carlo methods, each of which uses Monte Carlo in different ways to solve the many-body problem:

Quantum Monte Carlo methods

- Variational Monte Carlo : A good place to start; it is commonly used in many sorts of quantum problems.
- Diffusion Monte Carlo : The most common high-accuracy method for electrons (that is, chemical problems), since it comes quite close to the exact ground-state energy fairly efficiently. Also used for simulating the quantum behavior of atoms, etc.
- Path integral Monte Carlo : Finite-temperature technique mostly applied to bosons where temperature is very important, especially superfluid helium.

- Auxiliary field Monte Carlo : Usually applied to lattice problems, although there has been recent work on applying it to electrons in chemical systems.
- Reptation Monte Carlo : Recent zero-temperature method related to path integral Monte Carlo, with applications similar to diffusion Monte Carlo but with some different tradeoffs.
- Gaussian quantum Monte Carlo

See also

- Stochastic Green Function (SGF) algorithm
- Monte Carlo method
- QMC@Home
- Quantum chemistry
- Density matrix renormalization group
- Time-evolving block decimation
- Metropolis algorithm
- Wavefunction optimization

References

[1] http://www.attaccalite.altervista.org/PhDThesis/html/node9.html

- Hammond, B.J.; W.A. Lester & P.J. Reynolds (1994) (in English). *Monte Carlo Methods in Ab Initio Quantum Chemistry (http://www.worldscibooks.com/chemistry/1170.html)*. Singapore: World Scientific. ISBN 981-02-0321-7. OCLC 29594695 (http://worldcat.org/oclc/29594695). http://www.worldscibooks.com/chemistry/1170.html. Retrieved on 2007-01-18.
- Nightingale, M.P.; Umrigar, Cyrus J., ed (1999) (in English). *Quantum Monte Carlo Methods in Physics and Chemistry (http://www.springer.com/west/0-7923-5552-0)*. Springer. ISBN 978-0-7923-5552-6. http://www.springer.com/west/0-7923-5552-0. Retrieved on 2007-01-18.
- W. M. C. Foulkes; L. Mitáš, R. J. Needs and G. Rajagopal (5 January 2001). " Quantum Monte Carlo simulations of solids (http://link.aps.org/abstract/RMP/v73/p33)" (in English) (abstract). *Rev. Mod. Phys.* **73**: 33–83. doi: 10.1103/RevModPhys.73.33 (http://dx.doi.org/10.1103/RevModPhys.73.33). http://link.aps.org/abstract/RMP/v73/p33. Retrieved on 2007-01-18.
- Raimundo R. dos Santos (2003). " Introduction to Quantum Monte Carlo simulations for fermionic systems (http://arxiv.org/abs/cond-mat/0303551v1)" (in English) (fulltext). *Braz. J. Phys.* **33**: 36. http://arxiv.org/abs/cond-mat/0303551v1. Retrieved on 2007-01-18.

External links

- QMCWIKI (http://www.qmcwiki.org/)
- Joint DEMOCRITOS-ICTP School on Continuum Quantum Monte Carlo Methods (http://cdsagenda5.ictp.trieste.it/full_display.php?ida=a0332&fid=)
- FreeScience Library -> Quantum Monte Carlo (http://freescience.info/books.php?id=35)
- UIUC 2007 Summer School on Computational Materials Science: Quantum Monte Carlo from Minerals and Materials to Molecules (http://www.mcc.uiuc.edu/summerschool/2007/qmc/)
- Quantum Monte Carlo in the Apuan Alps V (http://www.vallico.net/tti/tti.html) - international workshop, Vallico Sotto, Tuscany, 25 July-1 August 2009 (Click PUBLIC EVENTS) - Announcement (http://www.vallico.net/tti/qmcitaa_09/announcement.html), Poster (http://www.tcm.phy.cam.ac.uk/~mdt26/tti2/poster/tti_c_poster_2009.png)
- Quantum Monte Carlo and the CASINO program IV (http://www.vallico.net/tti/tti.html) - summer school, Vallico Sotto, Tuscany, 2-9 August 2009 (Click PUBLIC EVENTS) - Announcement (http://www.vallico.net/tti/qmcatcp_09/announcement.html), Poster (http://www.tcm.phy.cam.ac.uk/~mdt26/tti2/poster/tti_ss_poster_2009.png)

DNA Dynamics

DNA Molecular dynamics modeling involves simulations of DNA molecular geometry and topology changes with time as a result of both intra- and inter- molecular interactions of DNA. Whereas molecular models of Deoxyribonucleic acid (DNA) molecules such as closely packed spheres (CPK models) made of plastic or metal wires for 'skeletal models' are useful representations of static DNA structures, their usefulness is very limited for representing complex DNA dynamics. Computer molecular modeling allows both animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*.

An old standing dynamic problem is how DNA "self-replication" takes place in living cells that should involve transient uncoiling of supercoiled DNA fibers. Although DNA consists of relatively rigid, very large elongated biopolymer molecules called "fibers" or chains its molecular structure *in vivo* undergoes dynamic configuration changes that involve dynamically attached water molecules, ions or proteins/enzymes. Supercoiling, packing with histones in chromosome structures, and other such supramolecular aspects also involve *in vivo* DNA topology which is even more complex than DNA molecular geometry, thus turning molecular modeling of DNA dynamics into a series of challenging problems for biophysical chemists, molecular biologists and biotechnologists. Thus, DNA exists in multiple stable geometries (called conformational isomerism) and has a rather large number of configurational, quantum states which are close to each other in energy on the potential energy surface of the DNA molecule.

Such varying molecular geometries can also be computed, at least in principle, by employing *ab initio* quantum chemistry methods that can attain high accuracy for small molecules, although claims that acceptable accuracy can be also achieved for

polynuclelotides, as well as DNA conformations, were recently made on the basis of VCD spectral data. Such quantum geometries define an important class of *ab initio* molecular models of DNA whose exploration has barely started especially in connection with results obtained by VCD in solutions. More detailed comparisons with such *ab initio* quantum computations are in principle obtainable through 2D-FT NMR spectroscopy and relaxation studies of polynucleotide solutions or specifically labeled DNA, as for example with deuterium labels.

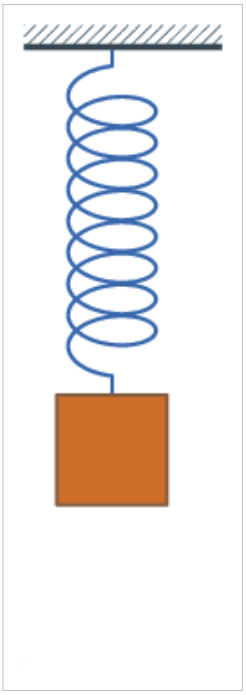
Importance of DNA molecular structure and dynamics modeling for Genomics and beyond

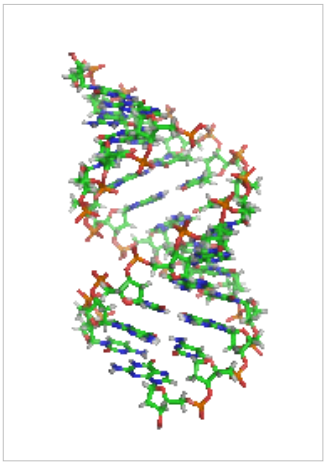
From the very early stages of structural studies of DNA by X-ray diffraction and biochemical means, molecular models such as the Watson-Crick double-helix model were successfully employed to solve the 'puzzle' of DNA structure, and also find how the latter relates to its key functions in living cells. The first high quality X-ray diffraction patterns of A-DNA were reported by Rosalind Franklin and Raymond Gosling in 1953[1] . The first reports of a double-helix molecular model of B-DNA structure were made by Watson and Crick in 1953[2] [3] . Then Maurice F. Wilkins, A. Stokes and H.R. Wilson, reported the first X-ray patterns of *in vivo* B-DNA in partially oriented salmon sperm heads [4] . The development of the first correct double-helix molecular model of DNA by Crick and Watson may not have been possible without the biochemical evidence for the nucleotide base-pairing ([A---T]; [C---G]), or Chargaff's rules[5] [6] [7] [8] [9] [10] . Although such initial studies of DNA structures with the help of molecular models were essentially static, their consequences for explaining the *in vivo* functions of DNA were significant in the areas of protein biosynthesis and the quasi-universality of the genetic code. Epigenetic transformation studies of DNA *in vivo* were however much slower to develop in spite of their importance for embryology, morphogenesis and cancer research. Such chemical dynamics and biochemical reactions of DNA are much more complex than the molecular dynamics of DNA physical interactions with water, ions and proteins/enzymes in living cells.

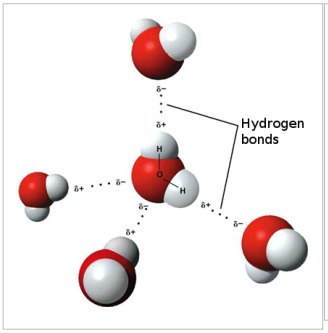
Animated DNA molecular models and hydrogen-bonding

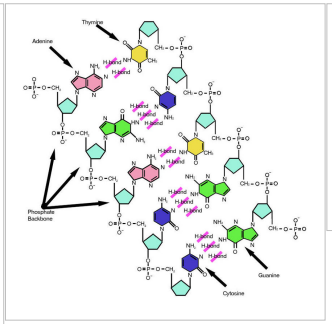
Animated molecular models allow one to visually explore the three-dimensional (3D) structure of DNA. The first DNA model is a space-filling, or CPK, model of the DNA double-helix whereas the third is an animated wire, or skeletal type, molecular model of DNA. The last two DNA molecular models in this series depict quadruplex DNA [11] that may be involved in certain cancers[12] [13] . The first CPK model in the second row is a molecular model of hydrogen bonds between water molecules in ice that are broadly similar to those found in DNA; the hydrogen bonding dynamics and proton exchange is however very different by many orders of magnitude between the two systems of fully hydrated DNA and water molecules in ice. Thus, the DNA dynamics is complex, involving nanosecond and several tens of picosecond time scales, whereas that of liquid ice is on the picosecond time scale, and that of proton exchange in ice is on the millisecond time scale; the proton exchange rates in DNA and attached proteins may vary from picosecond to nanosecond, minutes or years, depending on the exact locations of the exchanged protons in the large biopolymers. The simple harmonic oscillator 'vibration' in the third, animated image of the next gallery is only an oversimplified dynamic representation of the longitudinal vibrations of the DNA intertwined helices which were found to be anharmonic rather than harmonic as

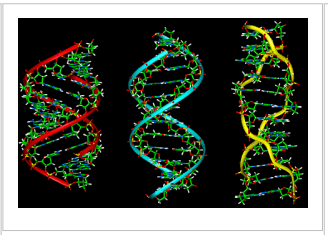
often assumed in quantum dynamic simulations of DNA.

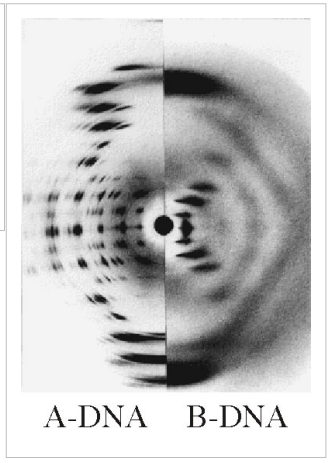


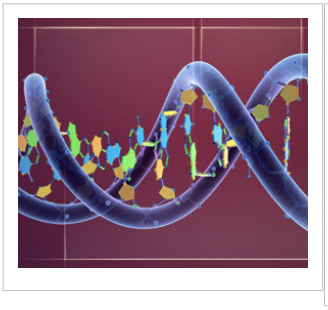


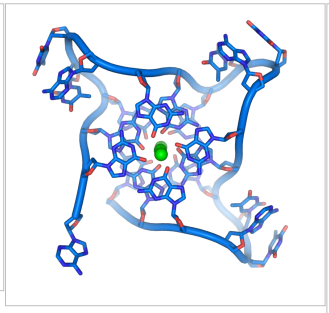


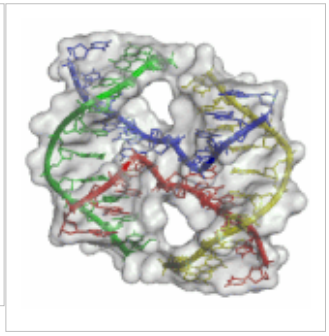










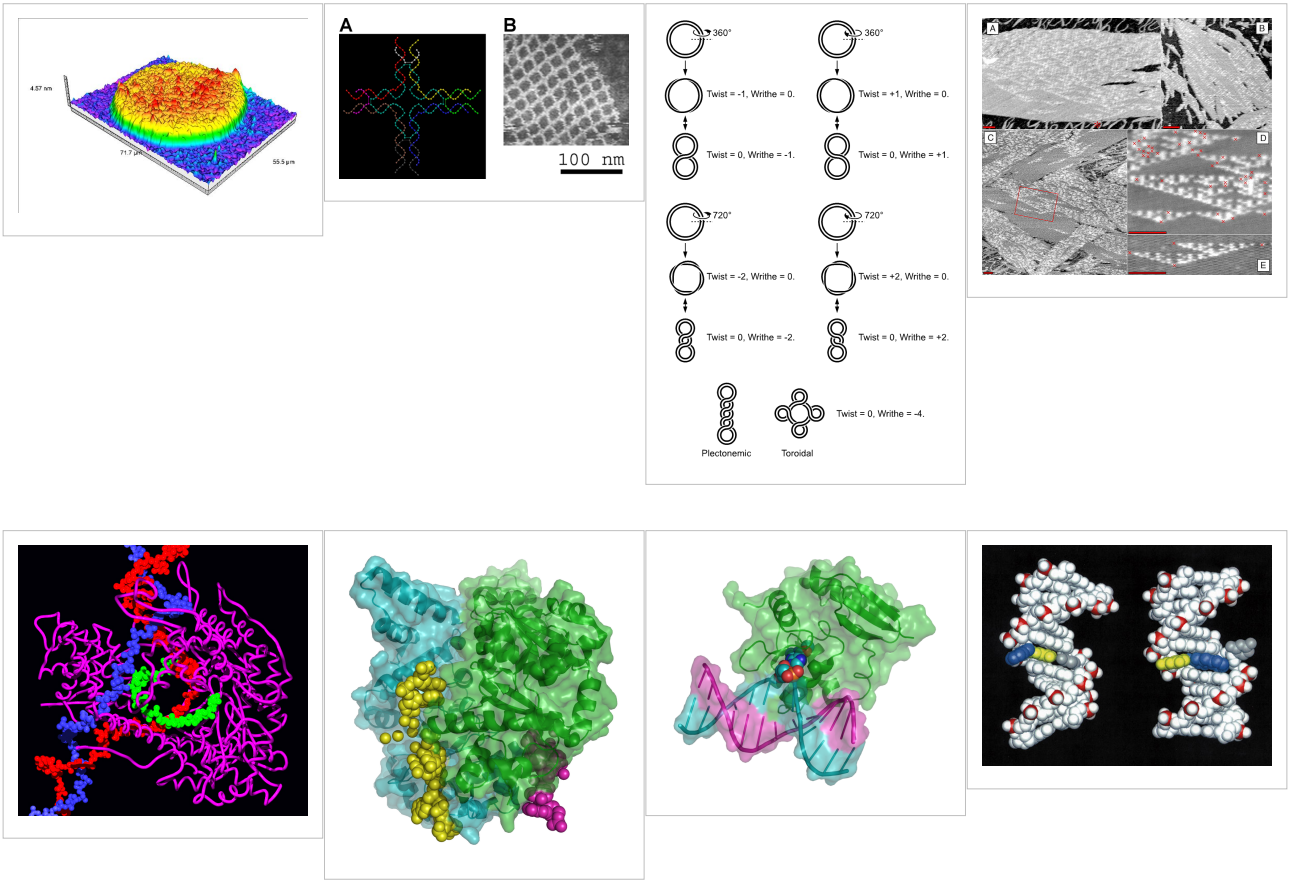


Human Genomics and Biotechnology Applications of DNA Molecular Modeling

The following two galleries of images illustrate various uses of DNA molecular modeling in Genomics and Biotechnology research applications from DNA repair to PCR and DNA nanostructures; each slide contains its own explanation and/or details. The first slide presents an overview of DNA applications, including DNA molecular models, with emphasis on Genomics and Biotechnology.

Applications of DNA molecular dynamics computations

- *First row* images present a DNA biochip and DNA nanostructures designed for DNA computing and other dynamic applications of DNA nanotechnology; last image in this row is of DNA arrays that display a representation of the Sierpinski gasket on their surfaces.
- *Second row*: the first two images show computer molecular models of RNA polymerase, followed by that of an E. coli, bacterial DNA primase template suggesting very complex dynamics at the interfaces between the enzymes and the DNA template; the fourth image illustrates in a computed molecular model the mutagenic, chemical interaction of a potent carcinogen molecule with DNA, and the last image shows the different interactions of specific fluorescence labels with DNA in human and orangoutan chromosomes.



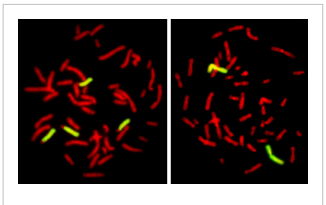
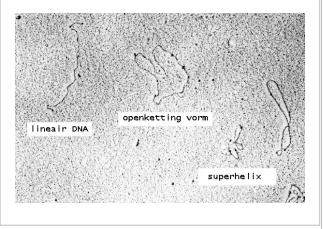
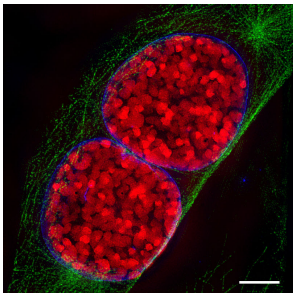
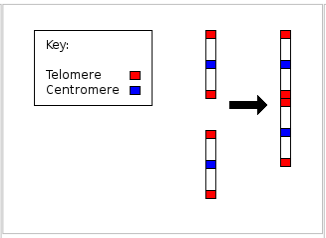


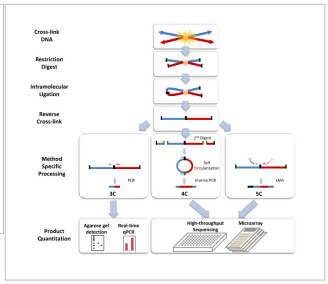
Image Gallery: *DNA Applications and Technologies at various scales in Biotechnology and Genomics research*

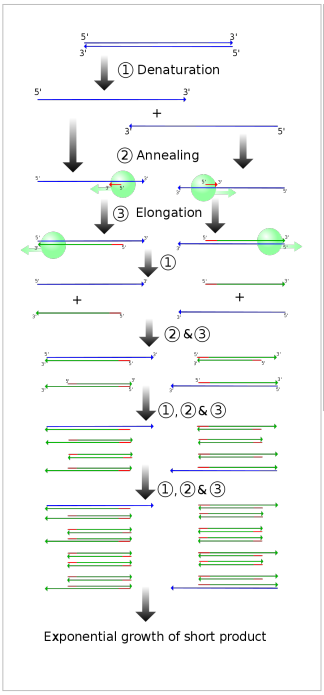
The first figure is an actual electron micrograph of a DNA fiber bundle, presumably of a single plasmid, bacterial DNA loop.

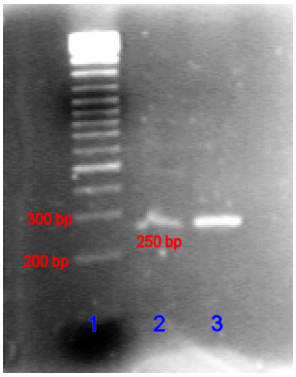


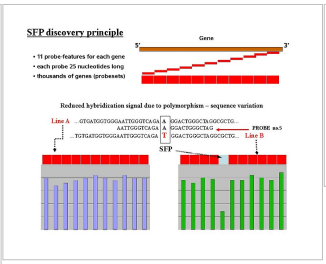


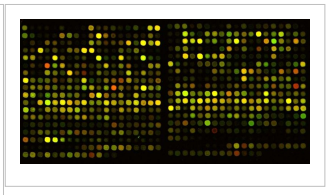


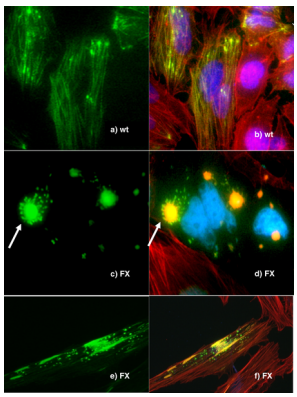


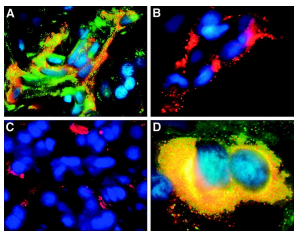


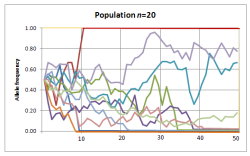


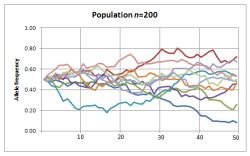


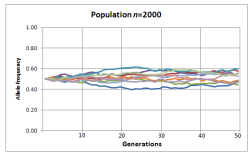


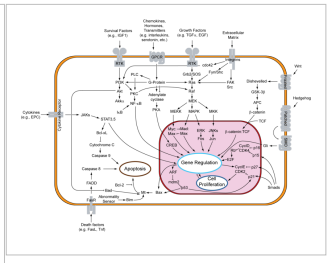


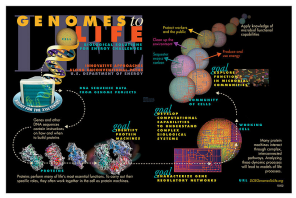










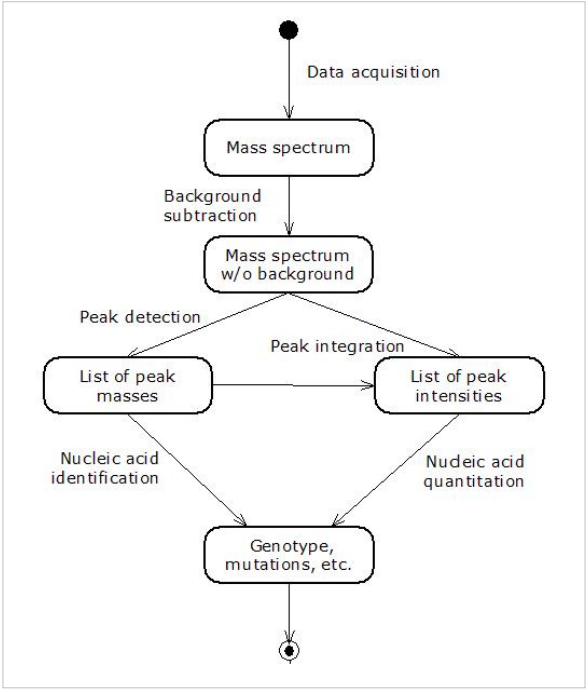


Databases for Genomics, DNA Dynamics and Sequencing

Genomic and structural databases

- CBS Genome Atlas Database [14] — contains examples of base skews.[15]
- The Z curve database of genomes — a 3-dimensional visualization and analysis tool of genomes [16][17] .
- DNA and other nucleic acids' molecular models: Coordinate files of nucleic acids molecular structure models in PDB and CIF formats [18]

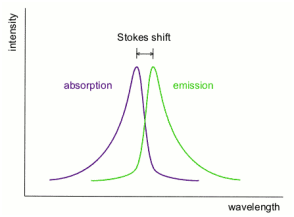
Mass spectrometry--Maldi informatics

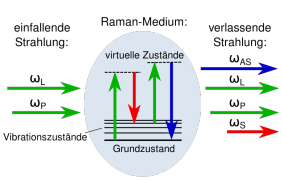


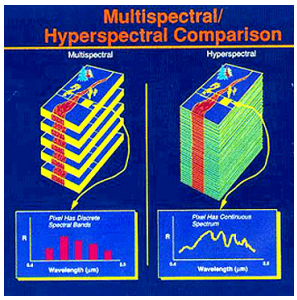
DNA Dynamics Data from Spectroscopy

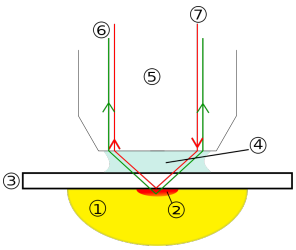
- FT-NMR[19] [20]
  - NMR Atlas--database [21]
  - mmcif downloadable coordinate files of nucleic acids in solution from 2D-FT NMR data [22]
  - NMR constraints files for NAs in PDB format [23]
- NMR microscopy[24]
- Vibrational circular dichroism (VCD)
- Microwave spectroscopy
- FT-IR
- FT-NIR[25] [26] [27]
- Spectral, Hyperspectral, and Chemical imaging)[28] [29] [30] [31] [32] [33] [34] .
- Raman spectroscopy/microscopy[35] and CARS[36] .
- Fluorescence correlation spectroscopy[37] [38] [39] [40] [41] [42] [43] [44] , Fluorescence cross-correlation spectroscopy and FRET[45] [46] [47] .
- Confocal microscopy[48]

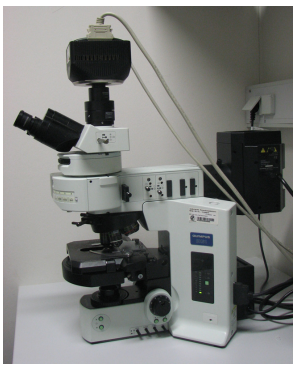
Gallery: CARS (Raman spectroscopy), Fluorescence confocal microscopy, and Hyperspectral imaging

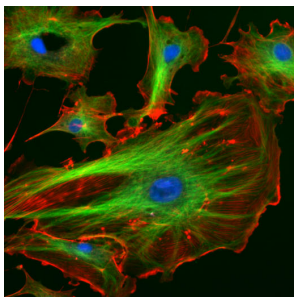


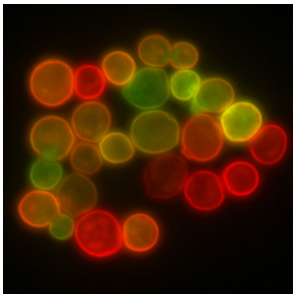


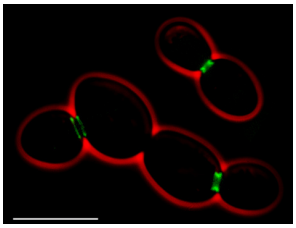


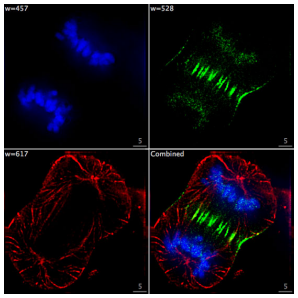


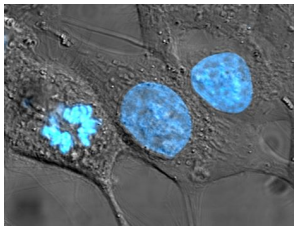


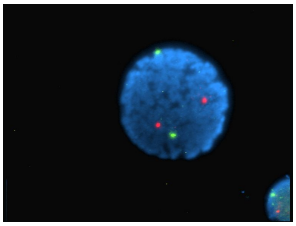


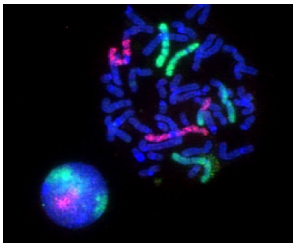


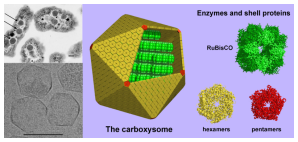












X-ray microscopy

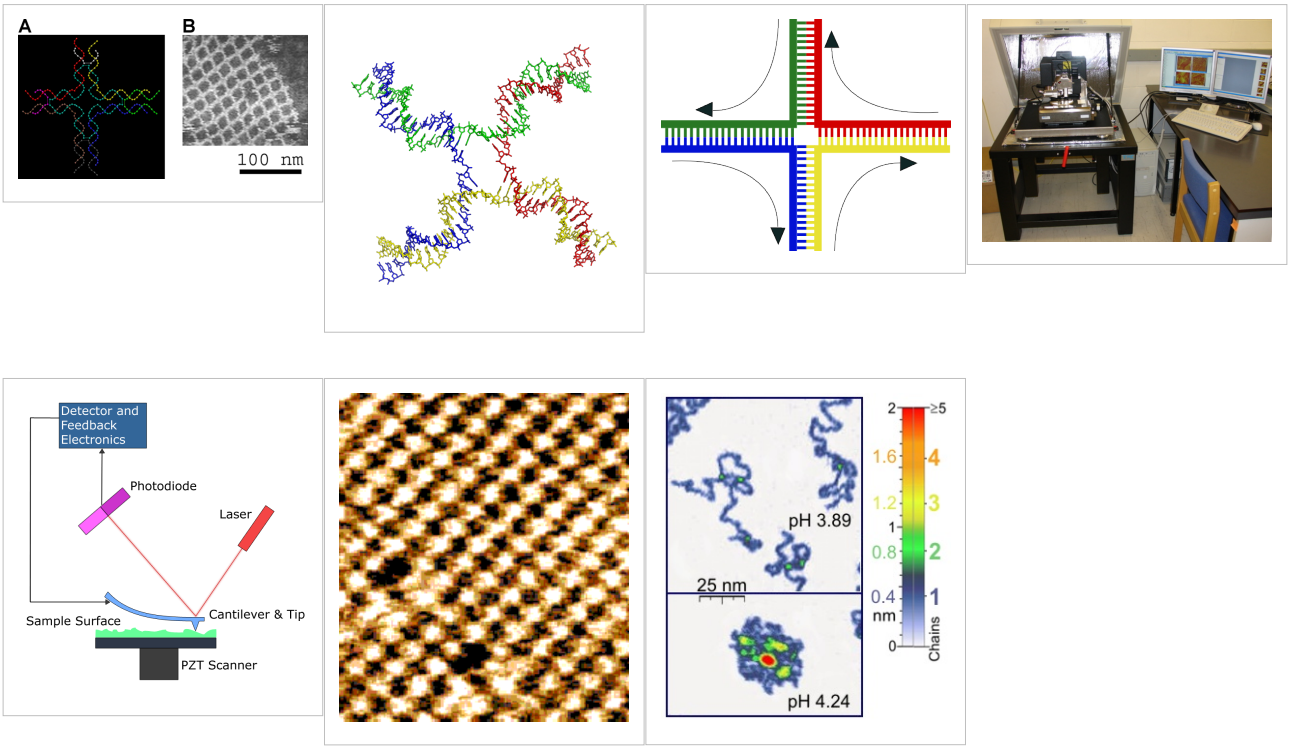
- Application of X-ray microscopy in the analysis of living hydrated cells [49]

Atomic Force Microscopy (AFM)

Two-dimensional DNA junction arrays have been visualized by Atomic Force Microscopy (AFM)[50] . Other imaging resources for AFM/Scanning probe microscopy(SPM) can be freely accessed at:

- How SPM Works [51]
- SPM Image Gallery - AFM STM SEM MFM NSOM and more. [52]

Gallery of AFM Images of DNA Nanostructures



Notes

[1] Franklin, R.E. and Gosling, R.G. recd.6 March 1953. Acta Cryst. (1953). 6, 673 The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content Acta Cryst. (1953). and 6, 678 The Structure of Sodium Thymonucleate Fibres II. The Cylindrically Symmetrical Patterson Function.

[2] Watson, J.D; Crick F.H.C. 1953a. Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid., Nature 171(4356):737-738.

[3] Watson, J.D; Crick F.H.C. 1953b. The Structure of DNA., Cold Spring Harbor Symposia on Quantitative Biology 18:123-131.

[4] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). " Molecular Structure of Deoxypentose Nucleic Acids (http://www.nature.com/nature/dna50/wilkins.pdf)" (PDF). *Nature* **171**: 738–740. doi: 10.1038/171738a0 (http://dx.doi.org/10.1038/171738a0). PMID 13054693. .

[5] Elson D, Chargaff E (1952). "On the deoxyribonucleic acid content of sea urchin gametes". *Experientia* **8** (4): 143-145.

[6] Chargaff E, Lipshitz R, Green C (1952). "Composition of the deoxypentose nucleic acids of four genera of sea-urchin". *J Biol Chem* **195** (1): 155-160. PMID 14938364.

[7] Chargaff E, Lipshitz R, Green C, Hodes ME (1951). "The composition of the deoxyribonucleic acid of salmon sperm". *J Biol Chem* **192** (1): 223-230. PMID 14917668.

[8] Chargaff E (1951). "Some recent studies on the composition and structure of nucleic acids". *J Cell Physiol Suppl* **38** (Suppl).
[9] Magasanik B, Vischer E, Doniger R, Elson D, Chargaff E (1950). "The separation and estimation of ribonucleotides in minute quantities". *J Biol Chem* **186** (1): 37-50. PMID 14778802.
[10] Chargaff E (1950). "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation". *Experientia* **6** (6): 201-209.
[11] http://ndbserver.rutgers.edu/atlas/xray/structures/U/ud0017/ud0017.html
[12] http://www.phy.cam.ac.uk/research/bss/molbiophysics.php
[13] http://planetphysics.org/encyclopedia/TheoreticalBiophysics.html
[14] http://www.cbs.dtu.dk/services/GenomeAtlas/
[15] Hallin PF, David Ussery D (2004). "CBS Genome Atlas Database: A dynamic storage for bioinformatic results and DNA sequence data". *Bioinformatics* **20**: 3682-3686.
[16] http://tubic.tju.edu.cn/zcurve/
[17] Zhang CT, Zhang R, Ou HY (2003). "The Z curve database: a graphic representation of genome sequences". Bioinformatics 19 (5): 593-599. doi:10.1093/bioinformatics/btg041
[18] http://ndbserver.rutgers.edu/ftp/NDB/models/
[19] (http://www.jonathanpmiller.com/Karplus.html)- obtaining dihedral angles from ^{3}J coupling constants
[20] (http://www.spectroscopynow.com/FCKeditor/UserFiles/File/specNOW/HTML files/General_Karplus_Calculator.htm) Another Javascript-like NMR coupling constant to dihedral
[21] http://ndbserver.rutgers.edu/atlas/nmr/index.html
[22] http://ndbserver.rutgers.edu/ftp/NDB/coordinates/na-nmr-mmcif/
[23] http://ndbserver.rutgers.edu/ftp/NDB/nmr-restraints/
[24] Lee, S. C. et al., (2001). One Micrometer Resolution NMR Microscopy. *J. Magn. Res.*, **150**: 207-213.
[25] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[26] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[27] Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
[28] http://www.imaging.net/chemical-imaging/ Chemical imaging
[29] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. Microscopy Today, Volume 12, No. 6, 11/2004.
[30] D.S. Mantus and G. H. Morrison. 1991. Chemical imaging in biology and medicine using ion microscopy., *Microchimica Acta*, **104**, (1-6) January 1991, doi: 10.1007/BF01245536
[31] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[32] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[33] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No.1-2, 2007.
[34] Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.(June 2004).,I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin q-bio/0406047 (http://arxiv.org/abs/q-bio/0406047)
[35] Chemical Imaging Without Dyeing (http://witec.de/en/download/Raman/ImagingMicroscopy04.pdf)
[36] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
[37] Eigen, M., Rigler, M. Sorting single molecules: application to diagnostics and evolutionary biotechnology,(1994) *Proc. Natl. Acad. Sci. USA*, 91,5740-5747.
[38] Rigler, M. Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology,(1995) J. Biotechnol., 41,177-186.
[39] Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.

[40] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[41] Oehlenschläger F., Schwille P. and Eigen M. (1996). Detection of HIV-1 RNA by nucleic acid sequence-based amplification combined with fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. USA* **93**:1281.
[42] Bagatolli, L.A., and Gratton, E. (2000). Two-photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. Biophys J., 78:290-305.
[43] Schwille, P., Haupts, U., Maiti, S., and Webb. W.(1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. Biophysical Journal, 77(10):2251-2265.
[44] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[45] FRET description (http://dwb.unl.edu/Teacher/NSF/C08/C08Links/pps99.cryst.bbk.ac.uk/projects/gmocz/fret.htm)
[46] doi:10.1016/S0959-440X(00)00190-1 (http://dx.doi.org/10.1016/S0959-440X(00)00190-1)Recent advances in FRET: distance determination in protein–DNA complexes. *Current Opinion in Structural Biology* **2001**, *11(2)*, 201-207
[47] http://www.fretimaging.org/mcnamaraintro.html FRET imaging introduction
[48] Eigen, M., and Rigler, R. (1994). Sorting single molecules: Applications to diagnostics and evolutionary biotechnology, Proc. Natl. Acad. Sci. USA 91:5740.
[49] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938
[50] Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi: 10.1021/ja9900398 (http://dx.doi.org/10.1021/ja9900398). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
[51] http://www.parkafm.com/New_html/resources/01general.php
[52] http://www.rhk-tech.com/results/showcase.php

References

- Sir Lawrence Bragg, FRS. *The Crystalline State, A General survey.* London: G. Bells and Sons, Ltd., vols. 1 and 2., 1966., 2024 pages.
- F. Bessel, *Untersuchung des Theils der planetarischen Störungen*, Berlin Abhandlungen (1824), article 14.
- Cantor, C. R. and Schimmel, P.R. *Biophysical Chemistry, Parts I and II.*, San Franscisco: W.H. Freeman and Co. 1980. 1,800 pages.
- Eigen, M., and Rigler, R. (1994). Sorting single molecules: Applications to diagnostics and evolutionary biotechnology, *Proc. Natl. Acad. Sci. USA* **91**:5740.
- Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
- Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.
- *Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.*(June 2004) I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin., q-bio/0406047.
- Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004. I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004).
- Voet, D. and J.G. Voet. *Biochemistry*, 2nd Edn., New York, Toronto, Singapore: John Wiley & Sons, Inc., 1995, ISBN 0-471-58651-X., 1361 pages.

- Watson, G. N. *A Treatise on the Theory of Bessel Functions.*, (1995) Cambridge University Press. ISBN 0-521-48391-3.
- Watson, James D. and Francis H.C. Crick. A structure for Deoxyribose Nucleic Acid (http://www.nature.com/nature/dna50/watsoncrick.pdf) (PDF). *Nature* 171, 737–738, 25 April 1953.
- Watson, James D. *Molecular Biology of the Gene.* New York and Amsterdam: W.A. Benjamin, Inc. 1965., 494 pages.
- Wentworth, W.E. *Physical Chemistry. A short course.*, Malden (Mass.): Blackwell Science, Inc. 2000.
- Herbert R. Wilson, FRS. *Diffraction of X-rays by proteins, Nucleic Acids and Viruses.*, London: Edward Arnold (Publishers) Ltd. 1966.
- Kurt Wuthrich. *NMR of Proteins and Nucleic Acids.*, New York, Brisbane,Chicester, Toronto, Singapore: J. Wiley & Sons. 1986., 292 pages.
- Robinson, Bruche H.; Seeman, Nadrian C. (August 1987). "The Design of a Biochip: A Self-Assembling Molecular-Scale Memory Device". *Protein Engineering* **1** (4): 295–300. ISSN 0269-2139 (http://worldcat.org/issn/0269-2139). Link (http://peds.oxfordjournals.org/cgi/content/abstract/1/4/295)
- Rothemund, Paul W. K.; Ekani-Nkodo, Axel; Papadakis, Nick; Kumar, Ashish; Fygenson, Deborah Kuchnir & Winfree, Erik (22 December 2004). "Design and Characterization of Programmable DNA Nanotubes". *Journal of the American Chemical Society* **126** (50): 16344–16352. doi: 10.1021/ja044319l (http://dx.doi.org/10.1021/ja044319l). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
- Keren, K.; Kinneret Keren, Rotem S. Berman, Evgeny Buchstab, Uri Sivan, Erez Braun (November 2003). " DNA-Templated Carbon Nanotube Field-Effect Transistor (http://www.sciencemag.org/cgi/content/abstract/sci;302/5649/1380)". *Science* **302** (6549): 1380–1382. doi: 10.1126/science.1091022 (http://dx.doi.org/10.1126/science.1091022). ISSN 1095-9203 (http://worldcat.org/issn/1095-9203). http://www.sciencemag.org/cgi/content/abstract/sci;302/5649/1380.
- Zheng, Jiwen; Constantinou, Pamela E.; Micheel, Christine; Alivisatos, A. Paul; Kiehl, Richard A. & Seeman Nadrian C. (2006). "2D Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs". *Nano Letters* **6**: 1502–1504. doi: 10.1021/nl060994c (http://dx.doi.org/10.1021/nl060994c). ISSN 1530-6984 (http://worldcat.org/issn/1530-6984).
- Cohen, Justin D.; Sadowski, John P.; Dervan, Peter B. (2007). "Addressing Single Molecules on DNA Nanostructures". *Angewandte Chemie* **46** (42): 7956–7959. doi: 10.1002/anie.200702767 (http://dx.doi.org/10.1002/anie.200702767). ISSN 0570-0833 (http://worldcat.org/issn/0570-0833).
- Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi: 10.1021/ja9900398 (http://dx.doi.org/10.1021/ja9900398). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
- Constantinou, Pamela E.; Wang, Tong; Kopatsch, Jens; Israel, Lisa B.; Zhang, Xiaoping; Ding, Baoquan; Sherman, William B.; Wang, Xing; Zheng, Jianping; Sha, Ruojie & Seeman, Nadrian C. (2006). "Double cohesion in structural DNA nanotechnology". *Organic and Biomolecular Chemistry* **4**: 3414–3419. doi: 10.1039/b605212f (http://dx.doi.org/10.1039/b605212f).

See also

- DNA
- Molecular modeling of DNA
- Genomics
- Signal transduction
- Transcriptomics
- Interactomics
- Biotechnology
- Molecular graphics
- Quantum computing
- MAYA-II
- DNA computing
- DNA structure
- Molecular structure
- Molecular dynamics
- Molecular topology
- DNA topology
- DNA, the Genome and Interactome
- Molecular structure
- Molecular geometry fluctuations
- Molecular interactions
- Molecular topology
- Hydrogen bonding
- Hydrophobic interactions
- DNA dynamics and conformations
- DNA Conformational isomerism
- 2D-FT NMRI and Spectroscopy
- Paracrystalline lattices/Paracrystals
- NMR Spectroscopy
- VCD or Vibrational circular dichroism
- Microwave spectroscopy
- Two-dimensional IR spectroscopy
- FRET and FCS- Fluorescence correlation spectroscopy
- Fluorescence cross-correlation spectroscopy (FCCS)
- Spectral imaging
- Hyperspectral imaging
- Chemical imaging
- NMR microscopy
- X-ray scattering
- Neutron scattering
- Crystallography
- Crystal lattices
- Molecular geometry
- Nanostructure
- DNA nanotechnology
- Imaging
- Sirius visualization software

- Atomic force microscopy
- X-ray microscopy
- Liquid crystals
- Glasses
- QMC@Home
- Sir Lawrence Bragg, FRS
- Sir John Randall
- Francis Crick
- Manfred Eigen
- Felix Bloch
- Paul Lauterbur
- Maurice Wilkins
- Herbert Wilson, FRS
- Alex Stokes

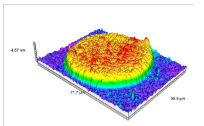
External links

- DNAlive: a web interface to compute DNA physical properties (http://mmb.pcb.ub.es/DNAlive). Also allows cross-linking of the results with the UCSC Genome browser and DNA dynamics.
- Application of X-ray microscopy in analysis of living hydrated cells (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938)
- DiProDB: Dinucleotide Property Database (http://diprodb.fli-leibniz.de). The database is designed to collect and analyse thermodynamic, structural and other dinucleotide properties.
- DNA the Double Helix Game (http://nobelprize.org/educational_games/medicine/dna_double_helix/) From the official Nobel Prize web site
- MDDNA: Structural Bioinformatics of DNA (http://humphry.chem.wesleyan.edu:8080/MDDNA/)
- Double Helix 1953–2003 (http://www.ncbe.reading.ac.uk/DNA50/) National Centre for Biotechnology Education
- DNA under electron microscope (http://www.fidelitysystems.com/Unlinked_DNA.html)
- Further details of mathematical and molecular analysis of DNA structure based on X-ray data (http://planetphysics.org/encyclopedia/BesselFunctionsApplicationsToDiffractionByHelicalStructures.html)
- Bessel functions corresponding to Fourier transforms of atomic or molecular helices. (http://planetphysics.org/?op=getobj&from=objects&name=BesselFunctionsAndTheirApplicationsToDiffractionByHelicalStructures)
- Characterization in nanotechnology some pdfs (http://nanocharacterization.sitesled.com/)
- An overview of STM/AFM/SNOM principles with educative videos (http://www.ntmdt.ru/SPM-Techniques/Principles/)
- SPM Image Gallery - AFM STM SEM MFM NSOM and More (http://www.rhk-tech.com/results/showcase.php)
- How SPM Works (http://www.parkafm.com/New_html/resources/01general.php)
- U.S. National DNA Day (http://www.genome.gov/10506367) — watch videos and participate in real-time discussions with scientists.

- The Secret Life of DNA - DNA Music compositions (http://www.tjmitchell.com/stuart/dna.html)
- Ascalaph DNA (http://www.agilemolecule.com/Ascalaph/Ascalaph_DNA.html) — Commercial software for DNA modeling

Molecular models of DNA

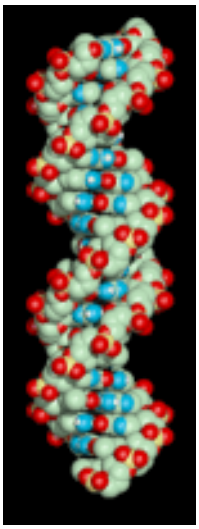
Molecular models of DNA structures are representations of the molecular geometry and topology of Deoxyribonucleic acid (DNA) molecules using one of several means, such as: closely packed spheres (CPK models) made of plastic, metal wires for 'skeletal models', graphic computations and animations by computers, artistic rendering, and so on, with the aim of simplifying and presenting the essential, physical and chemical, properties of DNA molecular structures either *in vivo* or *in vitro*. Computer molecular models also allow animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*. Thus, an old standing dynamic problem is how DNA "self-replication" takes place in living cells that should involve transient uncoiling of supercoiled DNA fibers. Although DNA consists of relatively rigid, very large elongated biopolymer molecules called "fibers" or chains (that are made of repeating nucleotide units of four basic types, attached to deoxyribose and phosphate groups), its molecular structure *in vivo* undergoes dynamic configuration changes that involve dynamically attached water molecules and ions. Supercoiling, packing with histones in chromosome structures, and other such supramolecular aspects also involve *in vivo* DNA topology which is even more complex than DNA molecular geometry, thus turning molecular modeling of DNA into an especially challenging problem for both molecular biologists and biotechnologists. Like other large molecules and biopolymers, DNA often exists in multiple stable geometries (that is, it exhibits conformational isomerism) and configurational, quantum states which are close to each other in energy on the potential energy surface of the DNA molecule. Such geometries can also be computed, at least in principle, by employing *ab initio* quantum chemistry methods that have high accuracy for small molecules. Such quantum geometries define an important class of *ab initio* molecular models of DNA whose exploration has barely started.



DNA computing biochip:3D

In an interesting twist of roles, the DNA molecule itself was proposed to be utilized for quantum computing. Both DNA nanostructures as well as DNA 'computing' biochips have been built (see biochip image at right).

The more advanced, computer-based molecular models of DNA involve molecular dynamics simulations as well as quantum mechanical computations of vibro-rotations, delocalized molecular orbitals (MOs), electric dipole moments, hydrogen-bonding, and so on.



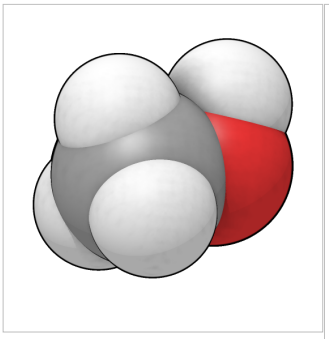
Spinning DNA generic model.

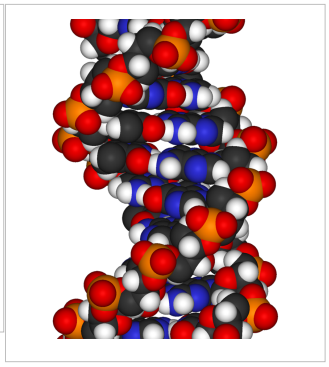
Importance

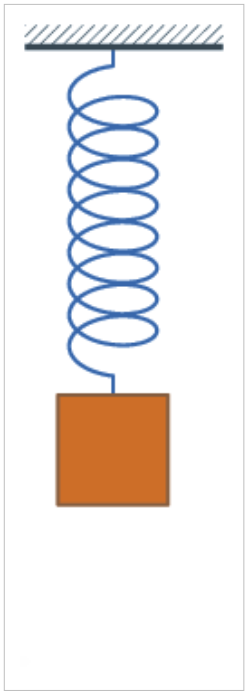
From the very early stages of structural studies of DNA by X-ray diffraction and biochemical means, molecular models such as the Watson-Crick double-helix model were successfully employed to solve the 'puzzle' of DNA structure, and also find how the latter relates to its key functions in living cells. The first high quality X-ray diffraction patterns of A-DNA were reported by Rosalind Franklin and Raymond Gosling in 1953[1] . The first calculations of the Fourier transform of an atomic helix were reported one year earlier by Cochran, Crick and Vand [2] , and were followed in 1953 by the computation of the Fourier transform of a coiled-coil by Crick[3] . The first reports of a double-helix molecular model of B-DNA structure were made by Watson and Crick in 1953[4] [5] . Last-but-not-least, Maurice F. Wilkins, A. Stokes and H.R. Wilson, reported the first X-ray patterns of *in vivo* B-DNA in partially oriented salmon sperm heads [6] . The development of the first correct double-helix molecular model of DNA by Crick and Watson may not have been possible without the biochemical evidence for the nucleotide base-pairing ([A---T]; [C---G]), or Chargaff's rules[7] [8] [9] [10] [11] [12] .

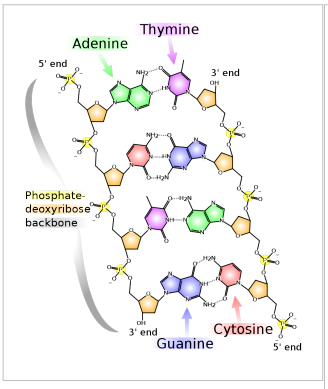
Examples of DNA molecular models

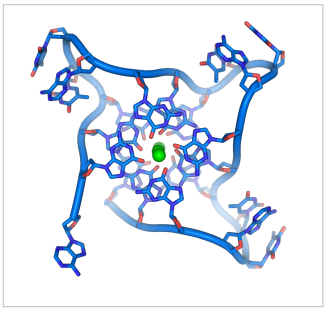
Animated molecular models allow one to visually explore the three-dimensional (3D) structure of DNA. The first DNA model is a space-filling, or CPK, model of the DNA double-helix whereas the third is an animated wire, or skeletal type, molecular model of DNA. The last two DNA molecular models in this series depict quadruplex DNA [13] that may be involved in certain cancers[14] [15] . The last figure on this panel is a molecular model of hydrogen bonds between water molecules in ice that are similar to those found in DNA.

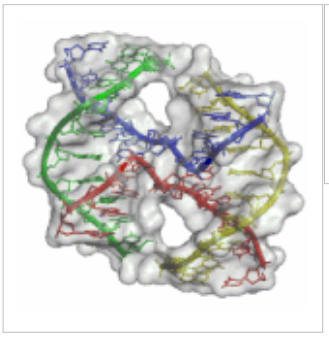


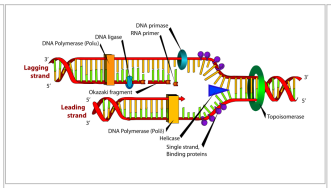


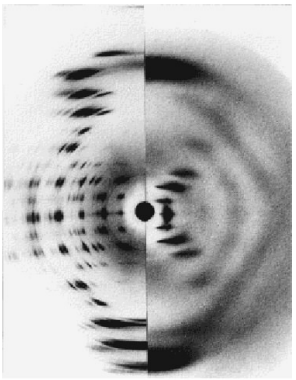



Thymine
Adenine
5' end
3' end
Phosphate-
deoxyribose
backbone
3' end
Guanine
Cytosine
5' end

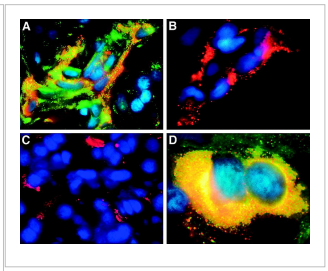


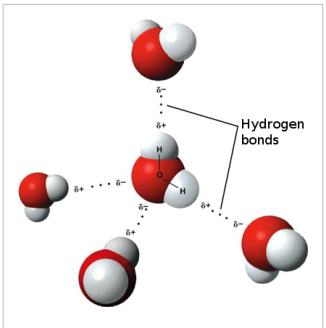



DNA primase
RNA primer
DNA ligase
DNA Polymerase (Pol I)
Lagging strand
Okazaki fragment
Leading strand
Topoisomerase
DNA Polymerase (Pol I)
Helicase
Single strand. Binding proteins

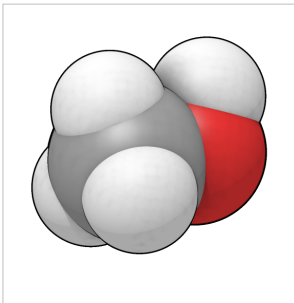


A-DNA B-DNA


A
B
C
D



- Spacefilling model or CPK model - a molecule is represented by overlapping spheres representing the atoms.

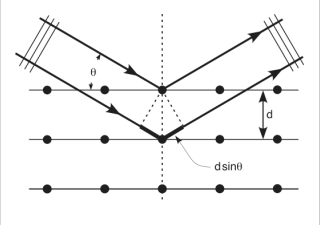


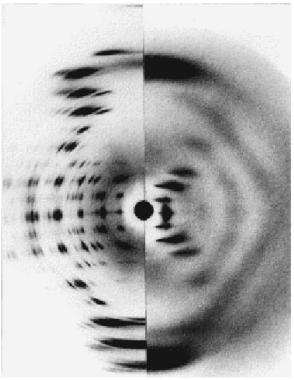
Images for DNA Structure Determination from X-Ray Patterns

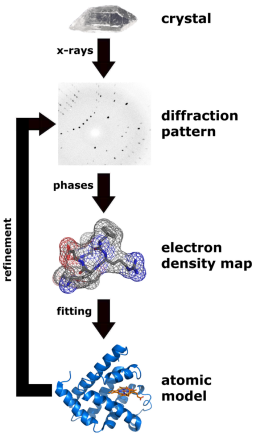
The following images illustrate both the principles and the main steps involved in generating structural information from X-ray diffraction studies of oriented DNA fibers with the help of molecular models of DNA that are combined with crystallographic and mathematical analysis of the X-ray patterns. From left to right the gallery of images shows:

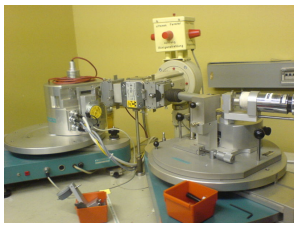
- *First row*:
- 1. Constructive X-ray interference, or diffraction, following Bragg's Law of X-ray "*reflection by the crystal planes*";
- 2. A comparison of A-DNA (crystalline) and highly hydrated B-DNA (paracrystalline) X-ray diffraction, and respectively, X-ray scattering patterns (courtesy of Dr. Herbert R. Wilson, FRS- see refs. list);
- 3. Purified DNA precipitated in a water jug;
- 4. The major steps involved in DNA structure determination by X-ray crystallography showing the important role played by molecular models of DNA structure in this iterative, structure--determination process;
- *Second row*:
- 5. Photo of a modern X-ray diffractometer employed for recording X-ray patterns of DNA with major components: X-ray source, goniometer, sample holder, X-ray detector and/or plate holder;
- 6. Illustrated animation of an X-ray goniometer;
- 7. X-ray detector at the SLAC synchrotron facility;
- 8. Neutron scattering facility at ISIS in UK;
- *Third and fourth rows*: *Molecular models of DNA structure at various scales*; figure #11 is an actual electron micrograph of a DNA fiber bundle, presumably of a single

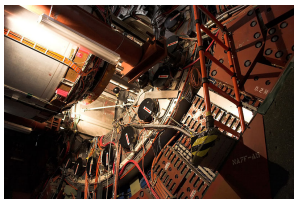
bacterial chromosome loop.

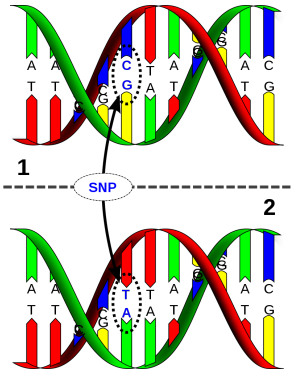


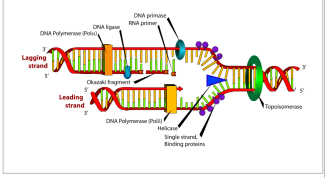


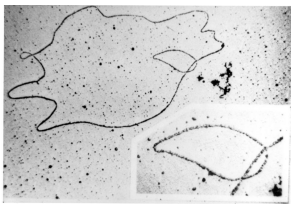


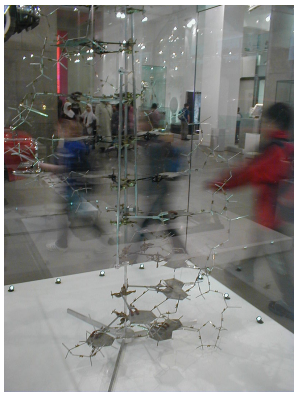


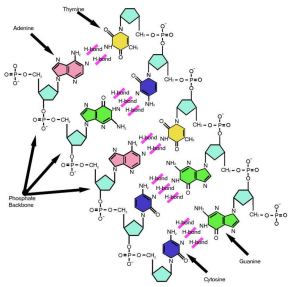


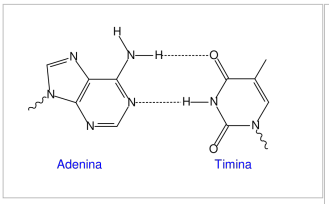


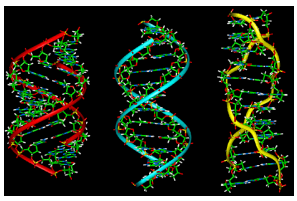


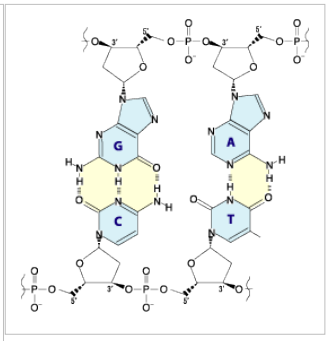


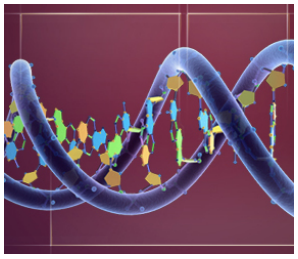


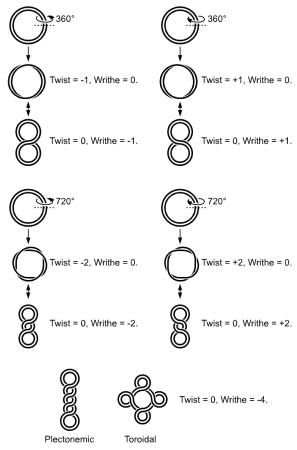


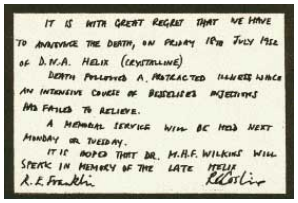






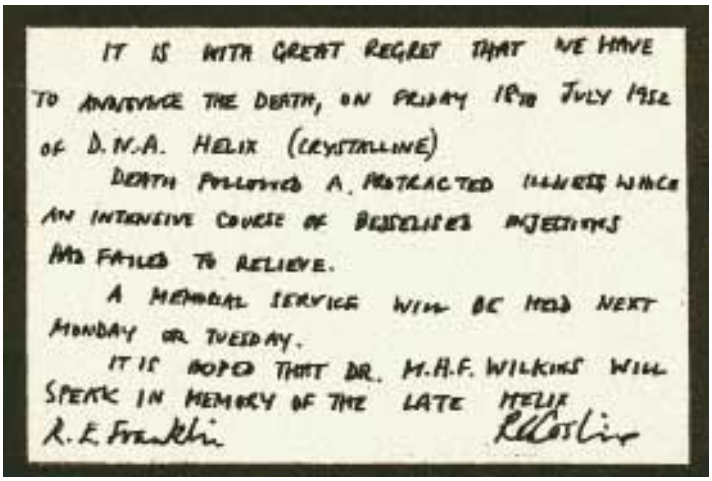






Paracrystalline lattice models of B-DNA structures

A paracrystalline lattice, or paracrystal, is a molecular or atomic lattice with significant amounts (e.g., larger than a few percent) of partial disordering of molecular arranegements. Limiting cases of the paracrystal model are nanostructures, such as glasses, liquids, etc., that may possess only local ordering and no global order. Liquid crystals also have paracrystalline rather than crystalline structures.



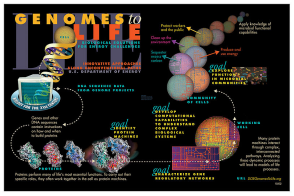
DNA Helix controversy in 1952

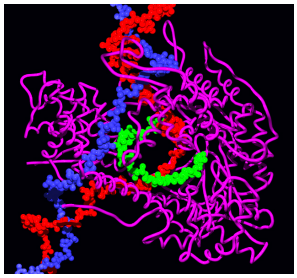
Highly hydrated B-DNA occurs naturally in living cells in such a paracrystalline state, which is a dynamic one in spite of the relatively rigid DNA double-helix stabilized by parallel hydrogen bonds between the nucleotide base-pairs in the two complementary, helical DNA chains (see figures). For simplicity most DNA molecular models ommit both water and ions dynamically bound to B-DNA, and are thus less useful for understanding the dynamic behaviors of B-DNA *in vivo*. The physical and mathematical analysis of X-ray[16] [17] and spectroscopic data for paracrystalline B-DNA is therefore much more complicated than that of crystalline, A-DNA X-ray diffraction patterns. The paracrystal model is also important for DNA technological applications such as DNA nanotechnology. Novel techniques that combine X-ray diffraction of DNA with X-ray microscopy in hydrated living cells are now also being developed (see, for example, "Application of X-ray microscopy in the analysis of living hydrated cells" [18]).

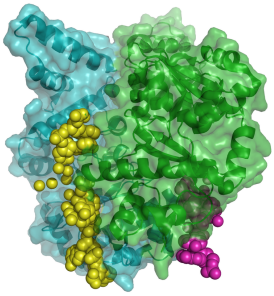
Genomic and Biotechnology Applications of DNA molecular modeling

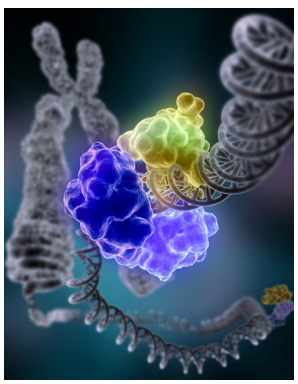
The following gallery of images illustrates various uses of DNA molecular modeling in Genomics and Biotechnology research applications from DNA repair to PCR and DNA nanostructures; each slide contains its own explanation and/or details. The first slide presents an overview of DNA applications, including DNA molecular models, with emphasis on Genomics and Biotechnology.

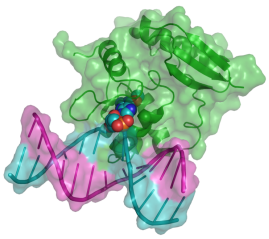
Gallery: DNA Molecular modeling applications

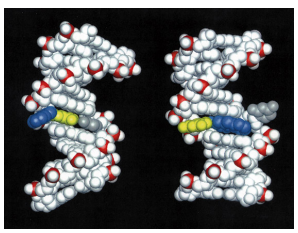


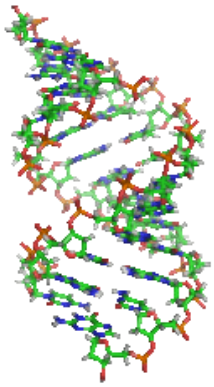


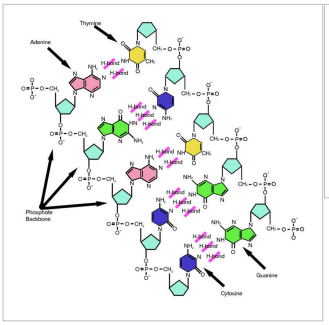


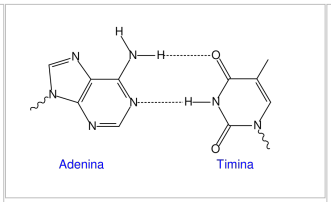


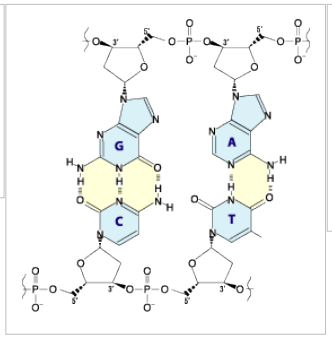


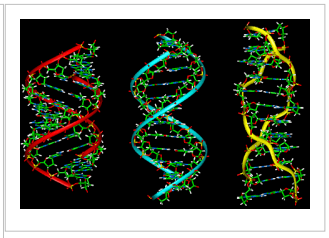


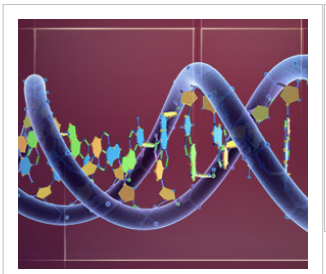


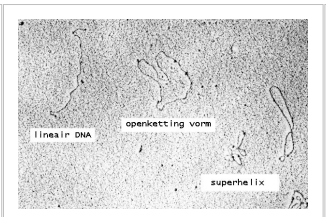

Thymine
Adenine
Phosphate Backbone
Guanine
Cytosine

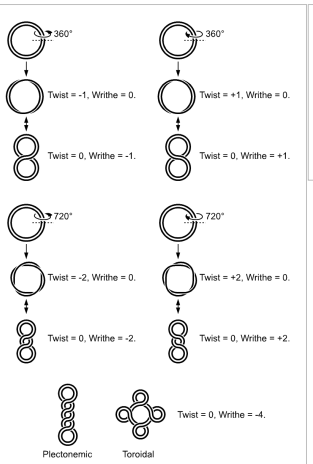

Adenina
Timina

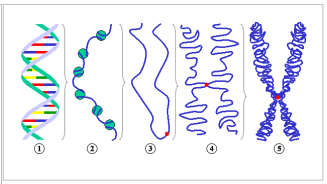

G
C
A
T

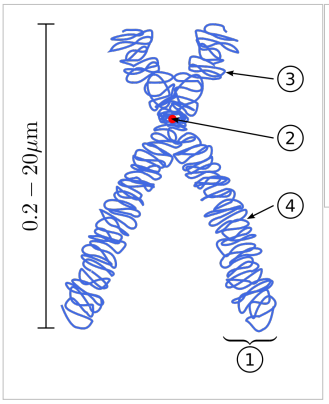


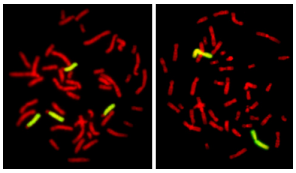


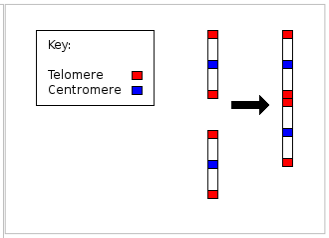

linear DNA
openkettling vorm
superhelix

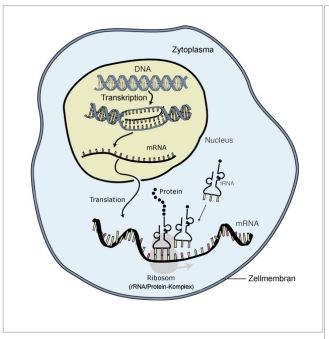

360°
Twist = -1, Writhe = 0.
Twist = 0, Writhe = -1.
360°
Twist = +1, Writhe = 0.
Twist = 0, Writhe = +1.
720°
Twist = -2, Writhe = 0.
Twist = 0, Writhe = -2.
720°
Twist = +2, Writhe = 0.
Twist = 0, Writhe = +2.
Twist = 0, Writhe = -4.
Plectonemic
Toroidal

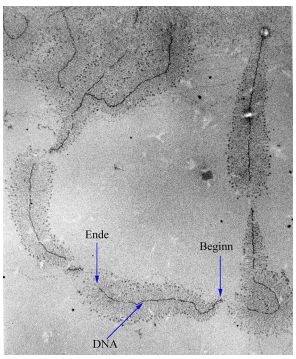

1
2
3
4
5

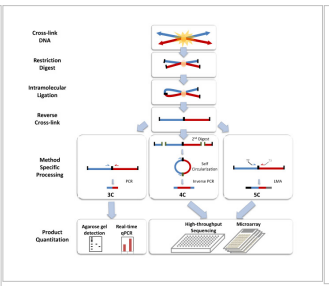

0.2 – 20µm
1
2
3
4

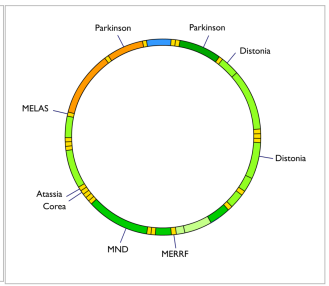


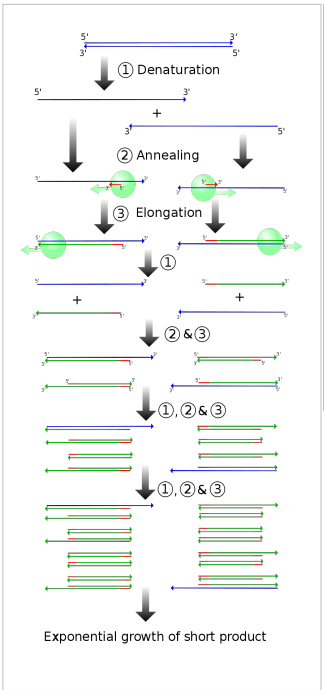

Key:
Telomere
Centromere

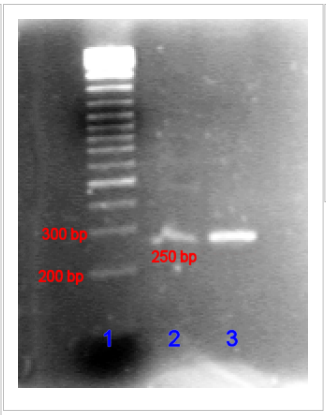

Zytoplasma
DNA
Transkription
Nucleus
mRNA
Translation
Protein
tRNA
mRNA
Ribosom
(rRNA/Protein-Komplex)
Zellmembran

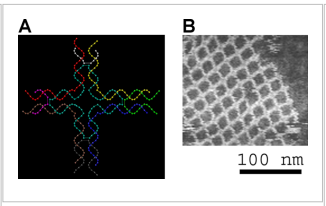

Ende
Beginn
DNA

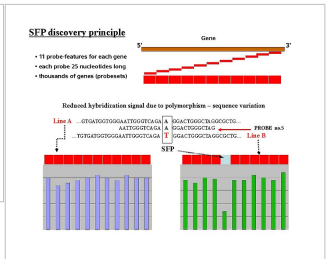

Cross-link DNA
Restriction Digest
Intramolecular Ligation
Reverse Cross-link
Method Specific Processing
3C
4C
5C
Product Quantitation

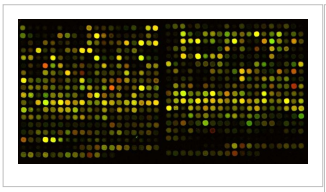

Parkinson
Parkinson
Distonia
MELAS
Distonia
Atassia
Corea
MND
MERRF

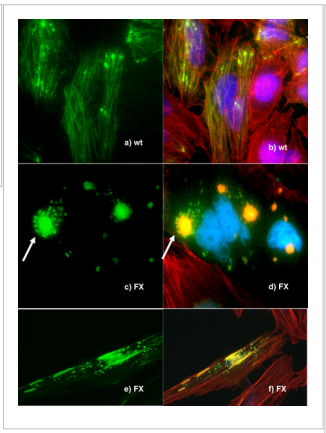


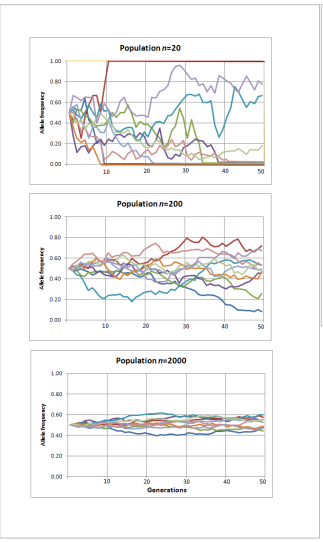












Databases for DNA molecular models and sequences

X-ray diffraction

- NDB ID: UD0017 Database [19]
- X-ray Atlas -database [20]
- PDB files of coordinates for nucleic acid structures from X-ray diffraction by NA (incl. DNA) crystals [21]
- Structure factors dowloadable files in CIF format [22]

Neutron scattering

- ISIS neutron source
- ISIS pulsed neutron source:A world centre for science with neutrons & muons at Harwell, near Oxford, UK. [23]

X-ray microscopy

- Application of X-ray microscopy in the analysis of living hydrated cells [24]

Electron microscopy

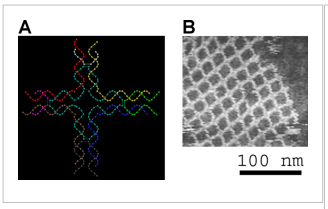
- DNA under electron microscope [25]

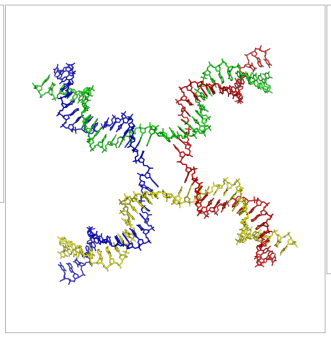
Atomic Force Microscopy (AFM)

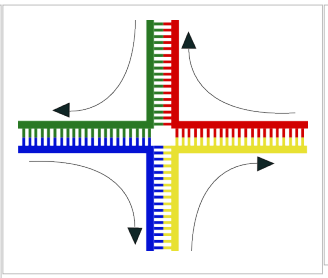
Two-dimensional DNA junction arrays have been visualized by Atomic Force Microscopy (AFM)[26] . Other imaging resources for AFM/Scanning probe microscopy(SPM) can be freely accessed at:

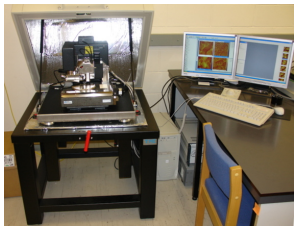
- How SPM Works [27]
- SPM Image Gallery - AFM STM SEM MFM NSOM and more. [28]

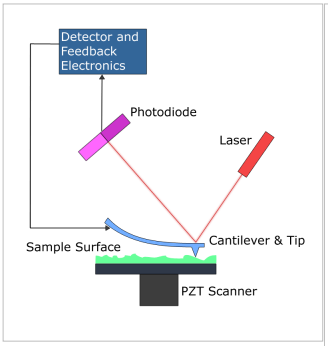
Gallery of AFM Images

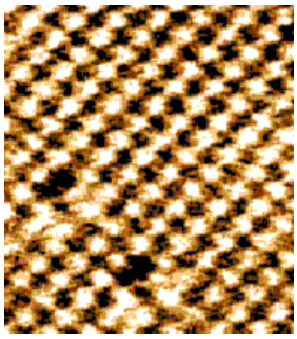


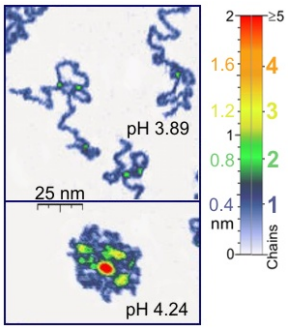


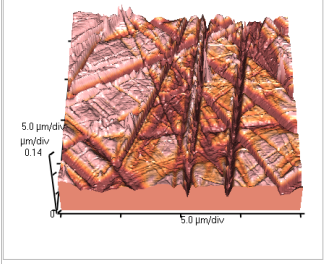




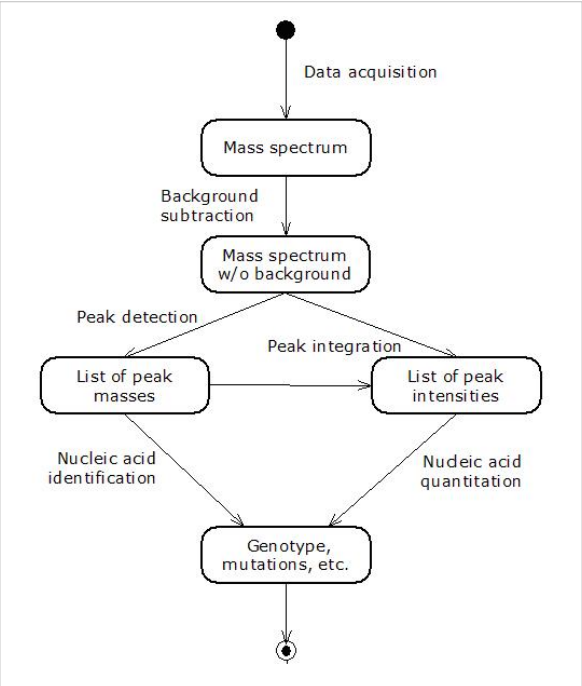








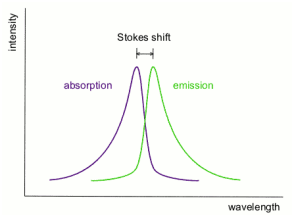
Mass spectrometry--Maldi informatics

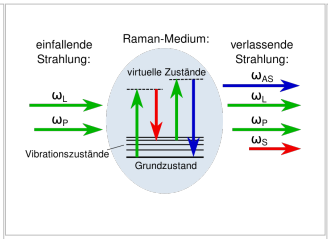


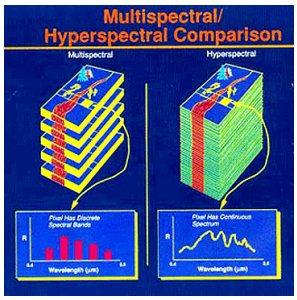
Spectroscopy

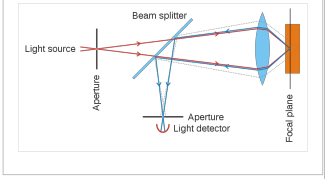
- Vibrational circular dichroism (VCD)
- FT-NMR[29] [30]
  - NMR Atlas--database [31]
  - mmcif downloadable coordinate files of nucleic acids in solution from 2D-FT NMR data [32]
  - NMR constraints files for NAs in PDB format [33]
- NMR microscopy[34]
- Microwave spectroscopy
- FT-IR
- FT-NIR[35] [36] [37]
- Spectral, Hyperspectral, and Chemical imaging)[38] [39] [40] [41] [42] [43] [44] .
- Raman spectroscopy/microscopy[45] and CARS[46] .
- Fluorescence correlation spectroscopy[47] [48] [49] [50] [51] [52] [53] [54] , Fluorescence cross-correlation spectroscopy and FRET[55] [56] [57] .
- Confocal microscopy[58]

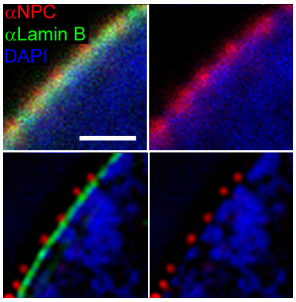
Gallery: CARS (Raman spectroscopy), Fluorescence confocal microscopy, and Hyperspectral imaging

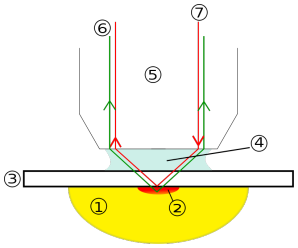


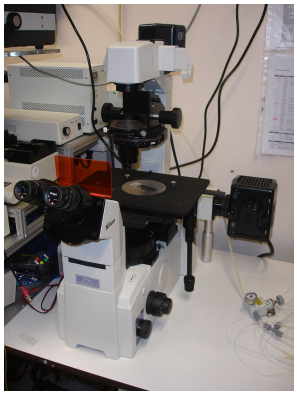


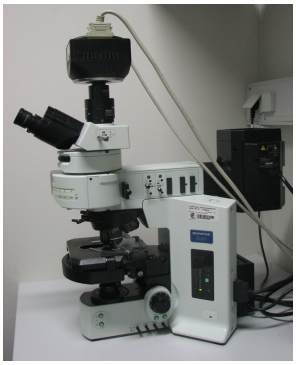


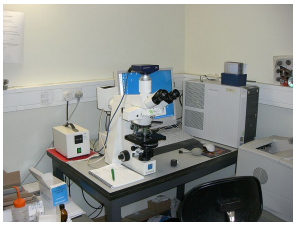


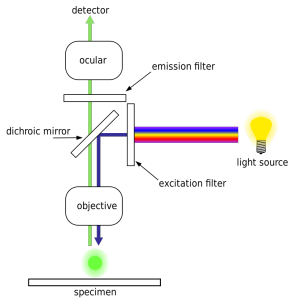


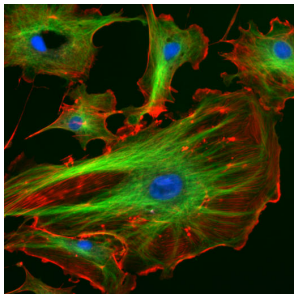


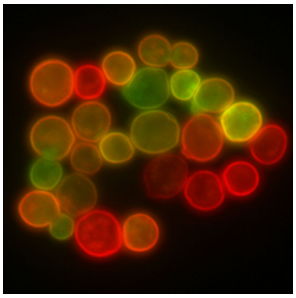


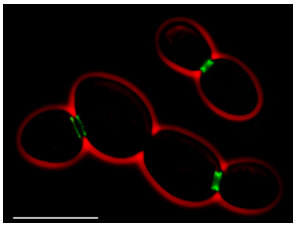


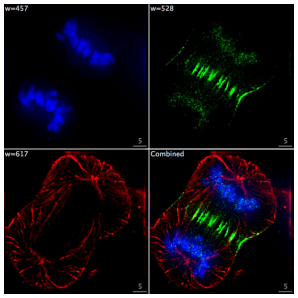


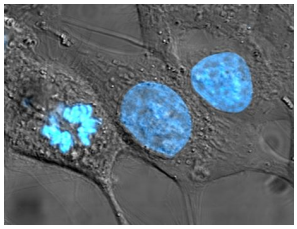


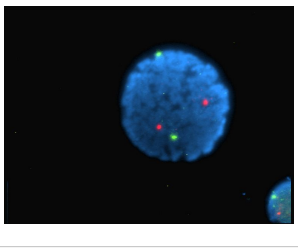


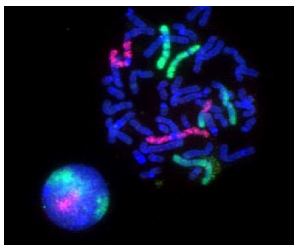


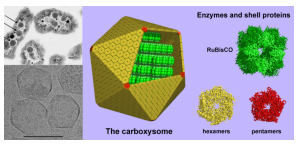












Genomic and structural databases

- CBS Genome Atlas Database [59] — contains examples of base skews.[60]
- The Z curve database of genomes — a 3-dimensional visualization and analysis tool of genomes [61][62] .
- DNA and other nucleic acids' molecular models: Coordinate files of nucleic acids molecular structure models in PDB and CIF formats [63]

Notes

[1] Franklin, R.E. and Gosling, R.G. recd.6 March 1953. Acta Cryst. (1953). 6, 673 The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content Acta Cryst. (1953). and 6, 678 The Structure of Sodium Thymonucleate Fibres II. The Cylindrically Symmetrical Patterson Function.

[2] Cochran, W., Crick, F.H.C. and Vand V. 1952. The Structure of Synthetic Polypeptides. 1. The Transform of Atoms on a Helix. Acta Cryst. 5(5):581-586.

[3] Crick, F.H.C. 1953a. The Fourier Transform of a Coiled-Coil., Acta Crystallographica 6(8-9):685-689.

[4] Watson, J.D; Crick F.H.C. 1953a. Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid., Nature 171(4356):737-738.

[5] Watson, J.D; Crick F.H.C. 1953b. The Structure of DNA., Cold Spring Harbor Symposia on Qunatitative Biology 18:123-131.

[6] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). " Molecular Structure of Deoxypentose Nucleic Acids (http://www.nature.com/nature/dna50/wilkins.pdf)" (PDF). *Nature* **171**: 738–740. doi: 10.1038/171738a0 (http://dx.doi.org/10.1038/171738a0). PMID 13054693. .

[7] Elson D, Chargaff E (1952). "On the deoxyribonucleic acid content of sea urchin gametes". *Experientia* **8** (4): 143-145.

[8] Chargaff E, Lipshitz R, Green C (1952). "Composition of the deoxypentose nucleic acids of four genera of sea-urchin". *J Biol Chem* **195** (1): 155-160. PMID 14938364.

[9] Chargaff E, Lipshitz R, Green C, Hodes ME (1951). "The composition of the deoxyribonucleic acid of salmon sperm". *J Biol Chem* **192** (1): 223-230. PMID 14917668.

[10] Chargaff E (1951). "Some recent studies on the composition and structure of nucleic acids". *J Cell Physiol Suppl* **38** (Suppl).

[11] Magasanik B, Vischer E, Doniger R, Elson D, Chargaff E (1950). "The separation and estimation of ribonucleotides in minute quantities". *J Biol Chem* **186** (1): 37-50. PMID 14778802.

[12] Chargaff E (1950). "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation". *Experientia* **6** (6): 201-209.

[13] http://ndbserver.rutgers.edu/atlas/xray/structures/U/ud0017/ud0017.html

[14] http://www.phy.cam.ac.uk/research/bss/molbiophysics.php

[15] http://planetphysics.org/encyclopedia/TheoreticalBiophysics.html

[16] Hosemann R., Bagchi R.N., *Direct analysis of diffraction by matter*, North-Holland Publs., Amsterdam – New York, 1962.

[17] Baianu, I.C. (1978). "X-ray scattering by partially disordered membrane systems.". *Acta Cryst.,* **A34** (5): 751–753. doi: 10.1107/S0567739478001540 (http://dx.doi.org/10.1107/S0567739478001540).

[18] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938

[19] http://ndbserver.rutgers.edu/atlas/xray/structures/U/ud0017/ud0017.html

[20] http://ndbserver.rutgers.edu/atlas/xray/index.html

[21] http://ndbserver.rutgers.edu/ftp/NDB/coordinates/na-biol/

[22] http://ndbserver.rutgers.edu/ftp/NDB/structure-factors/
[23] http://www.isis.rl.ac.uk/
[24] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938
[25] http://www.fidelitysystems.com/Unlinked_DNA.html
[26] Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi: 10.1021/ja9900398 (http://dx.doi.org/10.1021/ja9900398). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
[27] http://www.parkafm.com/New_html/resources/01general.php
[28] http://www.rhk-tech.com/results/showcase.php
[29] (http://www.jonathanpmiller.com/Karplus.html)- obtaining dihedral angles from ^{3}J coupling constants
[30] (http://www.spectroscopynow.com/FCKeditor/UserFiles/File/specNOW/HTML files/General_Karplus_Calculator.htm) Another Javascript-like NMR coupling constant to dihedral
[31] http://ndbserver.rutgers.edu/atlas/nmr/index.html
[32] http://ndbserver.rutgers.edu/ftp/NDB/coordinates/na-nmr-mmcif/
[33] http://ndbserver.rutgers.edu/ftp/NDB/nmr-restraints/
[34] Lee, S. C. et al., (2001). One Micrometer Resolution NMR Microscopy. *J. Magn. Res.*, **150**: 207-213.
[35] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[36] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[37] Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
[38] http://www.imaging.net/chemical-imaging/ Chemical imaging
[39] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. Microscopy Today, Volume 12, No. 6, 11/2004.
[40] D.S. Mantus and G. H. Morrison. 1991. Chemical imaging in biology and medicine using ion microscopy., *Microchimica Acta*, **104**, (1-6) January 1991, doi: 10.1007/BF01245536
[41] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[42] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[43] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No.1-2, 2007.
[44] Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.(June 2004).,I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin q-bio/0406047 (http://arxiv.org/abs/q-bio/0406047)
[45] Chemical Imaging Without Dyeing (http://witec.de/en/download/Raman/ImagingMicroscopy04.pdf)
[46] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
[47] Eigen, M., Rigler, M. Sorting single molecules: application to diagnostics and evolutionary biotechnology,(1994) *Proc. Natl. Acad. Sci. USA*, 91,5740-5747.
[48] Rigler, M. Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology,(1995) J. Biotechnol., 41,177-186.
[49] Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.
[50] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004) (http://arxiv.org/abs/q-bio/0407006)
[51] Oehlenschläger F., Schwille P. and Eigen M. (1996). Detection of HIV-1 RNA by nucleic acid sequence-based amplification combined with fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. USA* **93**:1281.

[52] Bagatolli, L.A., and Gratton, E. (2000). Two-photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. Biophys J., 78:290-305.
[53] Schwille, P., Haupts, U., Maiti, S., and Webb. W.(1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. Biophysical Journal, 77(10):2251-2265.
[54] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy,Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
[55] FRET description (http://dwb.unl.edu/Teacher/NSF/C08/C08Links/pps99.cryst.bbk.ac.uk/projects/gmocz/fret.htm)
[56] doi:10.1016/S0959-440X(00)00190-1 (http://dx.doi.org/10.1016/S0959-440X(00)00190-1)Recent advances in FRET: distance determination in protein–DNA complexes. *Current Opinion in Structural Biology* **2001**, *11(2)*, 201-207
[57] http://www.fretimaging.org/mcnamaraintro.html FRET imaging introduction
[58] Eigen, M., and Rigler, R. (1994). Sorting single molecules: Applications to diagnostics and evolutionary biotechnology, Proc. Natl. Acad. Sci. USA 91:5740.
[59] http://www.cbs.dtu.dk/services/GenomeAtlas/
[60] Hallin PF, David Ussery D (2004). "CBS Genome Atlas Database: A dynamic storage for bioinformatic results and DNA sequence data". *Bioinformatics* **20**: 3682-3686.
[61] http://tubic.tju.edu.cn/zcurve/
[62] Zhang CT, Zhang R, Ou HY (2003). "The Z curve database: a graphic representation of genome sequences". Bioinformatics 19 (5): 593-599. doi:10.1093/bioinformatics/btg041
[63] http://ndbserver.rutgers.edu/ftp/NDB/models/

References

- *Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology*.(June 2004) I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin., q-bio/0406047.
- F. Bessel, *Untersuchung des Theils der planetarischen Störungen*, Berlin Abhandlungen (1824), article 14.
- Sir Lawrence Bragg, FRS. *The Crystalline State, A General survey.* London: G. Bells and Sons, Ltd., vols. 1 and 2., 1966., 2024 pages.
- Cantor, C. R. and Schimmel, P.R. *Biophysical Chemistry, Parts I and II.*, San Franscisco: W.H. Freeman and Co. 1980. 1,800 pages.
- Eigen, M., and Rigler, R. (1994). Sorting single molecules: Applications to diagnostics and evolutionary biotechnology, *Proc. Natl. Acad. Sci. USA* **91**:5740.
- Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
- Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.
- Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004. I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004).
- Voet, D. and J.G. Voet. *Biochemistry*, 2nd Edn., New York, Toronto, Singapore: John Wiley & Sons, Inc., 1995, ISBN: 0-471-58651-X., 1361 pages.
- Watson, G. N. *A Treatise on the Theory of Bessel Functions.*, (1995) Cambridge University Press. ISBN 0-521-48391-3.
- Watson, James D. and Francis H.C. Crick. A structure for Deoxyribose Nucleic Acid (http://www.nature.com/nature/dna50/watsoncrick.pdf) (PDF). *Nature* 171, 737–738, 25 April 1953.

- Watson, James D. *Molecular Biology of the Gene.* New York and Amsterdam: W.A. Benjamin, Inc. 1965., 494 pages.
- Wentworth, W.E. *Physical Chemistry. A short course.*, Malden (Mass.): Blackwell Science, Inc. 2000.
- Herbert R. Wilson, FRS. *Diffraction of X-rays by proteins, Nucleic Acids and Viruses.*, London: Edward Arnold (Publishers) Ltd. 1966.
- Kurt Wuthrich. *NMR of Proteins and Nucleic Acids.*, New York, Brisbane,Chicester, Toronto, Singapore: J. Wiley & Sons. 1986., 292 pages.
- Robinson, Bruche H.; Seeman, Nadrian C. (August 1987). "The Design of a Biochip: A Self-Assembling Molecular-Scale Memory Device". *Protein Engineering* **1** (4): 295–300. ISSN 0269-2139 (http://worldcat.org/issn/0269-2139). Link (http://peds.oxfordjournals.org/cgi/content/abstract/1/4/295)
- Rothemund, Paul W. K.; Ekani-Nkodo, Axel; Papadakis, Nick; Kumar, Ashish; Fygenson, Deborah Kuchnir & Winfree, Erik (22 December 2004). "Design and Characterization of Programmable DNA Nanotubes". *Journal of the American Chemical Society* **126** (50): 16344–16352. doi: 10.1021/ja044319l (http://dx.doi.org/10.1021/ja044319l). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
- Keren, K.; Kinneret Keren, Rotem S. Berman, Evgeny Buchstab, Uri Sivan, Erez Braun (November 2003). " DNA-Templated Carbon Nanotube Field-Effect Transistor (http://www.sciencemag.org/cgi/content/abstract/sci;302/5649/1380)". *Science* **302** (6549): 1380–1382. doi: 10.1126/science.1091022 (http://dx.doi.org/10.1126/science.1091022). ISSN 1095-9203 (http://worldcat.org/issn/1095-9203). http://www.sciencemag.org/cgi/content/abstract/sci;302/5649/1380.
- Zheng, Jiwen; Constantinou, Pamela E.; Micheel, Christine; Alivisatos, A. Paul; Kiehl, Richard A. & Seeman Nadrian C. (2006). "2D Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs". *Nano Letters* **6**: 1502–1504. doi: 10.1021/nl060994c (http://dx.doi.org/10.1021/nl060994c). ISSN 1530-6984 (http://worldcat.org/issn/1530-6984).
- Cohen, Justin D.; Sadowski, John P.; Dervan, Peter B. (2007). "Addressing Single Molecules on DNA Nanostructures". *Angewandte Chemie* **46** (42): 7956–7959. doi: 10.1002/anie.200702767 (http://dx.doi.org/10.1002/anie.200702767). ISSN 0570-0833 (http://worldcat.org/issn/0570-0833).
- Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi: 10.1021/ja9900398 (http://dx.doi.org/10.1021/ja9900398). ISSN 0002-7863 (http://worldcat.org/issn/0002-7863).
- Constantinou, Pamela E.; Wang, Tong; Kopatsch, Jens; Israel, Lisa B.; Zhang, Xiaoping; Ding, Baoquan; Sherman, William B.; Wang, Xing; Zheng, Jianping; Sha, Ruojie & Seeman, Nadrian C. (2006). "Double cohesion in structural DNA nanotechnology". *Organic and Biomolecular Chemistry* **4**: 3414–3419. doi: 10.1039/b605212f (http://dx.doi.org/10.1039/b605212f).

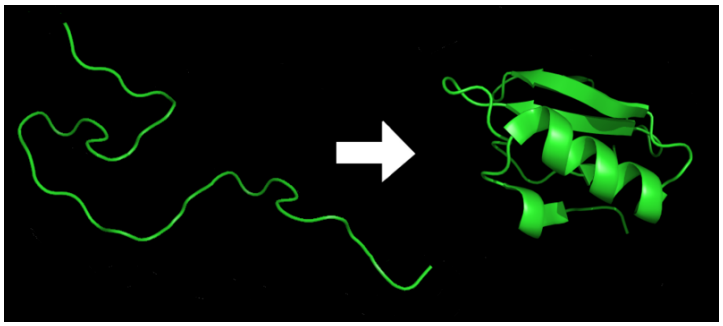
See also

- DNA
- Molecular graphics
- DNA structure
- DNA Dynamics
- X-ray scattering
- Neutron scattering
- Crystallography
- Crystal lattices
- Paracrystalline lattices/Paracrystals
- 2D-FT NMRI and Spectroscopy
- NMR Spectroscopy
- Microwave spectroscopy
- Two-dimensional IR spectroscopy
- Spectral imaging
- Hyperspectral imaging
- Chemical imaging
- NMR microscopy
- VCD or Vibrational circular dichroism
- FRET and FCS- Fluorescence correlation spectroscopy
- Fluorescence cross-correlation spectroscopy (FCCS)
- Molecular structure
- Molecular geometry
- Molecular topology
- DNA topology
- Sirius visualization software
- Nanostructure
- DNA nanotechnology
- Imaging
- Atomic force microscopy
- X-ray microscopy
- Liquid crystal
- Glasses
- QMC@Home
- Sir Lawrence Bragg, FRS
- Sir John Randall
- James Watson
- Francis Crick
- Maurice Wilkins
- Herbert Wilson, FRS
- Alex Stokes

External links

- DNA the Double Helix Game (http://nobelprize.org/educational_games/medicine/dna_double_helix/) From the official Nobel Prize web site
- MDDNA: Structural Bioinformatics of DNA (http://humphry.chem.wesleyan.edu:8080/MDDNA/)
- Double Helix 1953–2003 (http://www.ncbe.reading.ac.uk/DNA50/) National Centre for Biotechnology Education
- DNA under electron microscope (http://www.fidelitysystems.com/Unlinked_DNA.html)
- Ascalaph DNA (http://www.agilemolecule.com/Ascalaph/Ascalaph_DNA.html) — Commercial software for DNA modeling
- DNAlive: a web interface to compute DNA physical properties (http://mmb.pcb.ub.es/DNAlive). Also allows cross-linking of the results with the UCSC Genome browser and DNA dynamics.
- DiProDB: Dinucleotide Property Database (http://diprodb.fli-leibniz.de). The database is designed to collect and analyse thermodynamic, structural and other dinucleotide properties.
- Further details of mathematical and molecular analysis of DNA structure based on X-ray data (http://planetphysics.org/encyclopedia/BesselFunctionsApplicationsToDiffractionByHelicalStructures.html)
- Bessel functions corresponding to Fourier transforms of atomic or molecular helices. (http://planetphysics.org/?op=getobj&from=objects&name=BesselFunctionsAndTheirApplicationsToDiffractionByHelicalStructures)
- Application of X-ray microscopy in analysis of living hydrated cells (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938)
- Characterization in nanotechnology some pdfs (http://nanocharacterization.sitesled.com/)
- overview of STM/AFM/SNOM principles with educative videos (http://www.ntmdt.ru/SPM-Techniques/Principles/)
- SPM Image Gallery - AFM STM SEM MFM NSOM and More (http://www.rhk-tech.com/results/showcase.php)
- How SPM Works (http://www.parkafm.com/New_html/resources/01general.php)
- U.S. National DNA Day (http://www.genome.gov/10506367) — watch videos and participate in real-time discusssions with scientists.
- The Secret Life of DNA - DNA Music compositions (http://www.tjmitchell.com/stuart/dna.html)

Protein folding



Protein before and after folding.

Protein folding is the physical process by which a polypeptide folds into its characteristic and functional three-dimensional structure from random coil.[1] Each protein exists as an unfolded polypeptide or random coil when translated from a sequence of mRNA to a linear chain of amino acids. This polypeptide lacks any developed three-dimensional structure (the left hand side of the neighboring figure). However amino acids interact with each other to produce a well-defined three dimensional structure, the folded protein (the right hand side of the figure), known as the native state. The resulting three-dimensional structure is determined by the amino acid sequence.[2] .

For many proteins the correct three dimensional structure is essential to function.[3] Failure to fold into the intended shape usually produces inactive proteins with different properties including toxic prions. Several neurodegenerative and other diseases are believed to result from the accumulation of *misfolded* (incorrectly folded) proteins.[4]

Known facts about the process

The relationship between folding and amino acid sequence

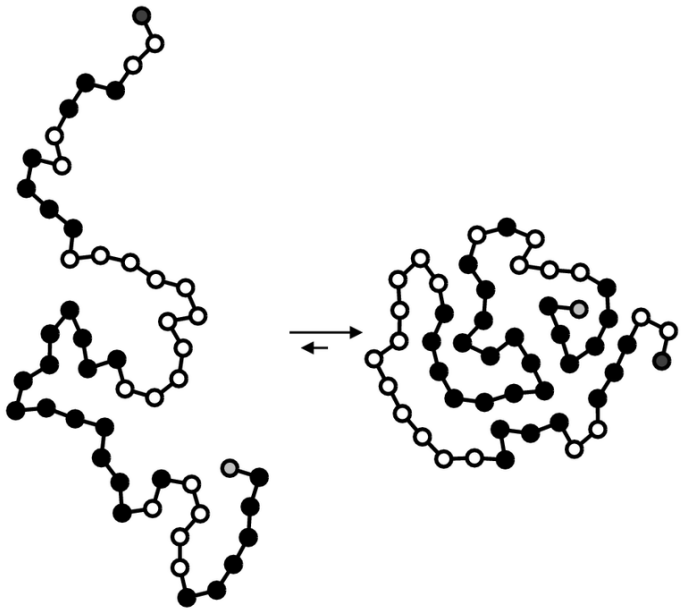


Illustration of the main driving force behind protein structure formation. In the compact fold (to the right), the hydrophobic amino acids (shown as black spheres) are in general shielded from the solvent.

The amino-acid sequence (or primary structure) of a protein defines its native conformation. A protein molecule folds spontaneously during or after synthesis. While these macromolecules may be regarded as "folding themselves", the process also depends on the solvent (water or lipid bilayer),[5] the concentration of salts, the temperature, and the presence of molecular chaperones.

Folded proteins usually have a hydrophobic core in which side chain packing stabilizes the folded state, and charged or polar side chains occupy the solvent-exposed surface where they interact with surrounding water. Minimizing the number of hydrophobic side-chains exposed to water is an important driving force behind the folding process,[6] . Formation of intramolecular hydrogen bonds provides another important contribution to protein stability.[7] The strength of hydrogen bonds depends on their environment, thus H-bonds enveloped in a hydrophobic core contribute more than H-bonds exposed to the aqueous environment to the stability of the native state.[8]

The process of folding *in vivo* often begins co-translationally, so that the N-terminus of the protein begins to fold while the C-terminal portion of the protein is still being synthesized by the ribosome. Specialized proteins called chaperones assist in the folding of other proteins.[9] A well studied example is the bacterial GroEL system, which assists in the folding of globular proteins. In eukaryotic organisms chaperones are known as heat shock proteins. Although most globular proteins are able to assume their native state unassisted, chaperone-assisted folding is often necessary in the crowded intracellular environment to prevent aggregation; chaperones are also used to prevent misfolding and aggregation which may occur as a consequence of exposure to heat or other changes in the cellular environment.

For the most part, scientists have been able to study many identical molecules folding together *en masse*. At the coarsest level, it appears that in transitioning to the native state, a given amino acid sequence takes on roughly the same route and proceeds through roughly the same intermediates and transition states. Often folding involves first the establishment of regular secondary and supersecondary structures, particularly alpha helices and beta sheets, and afterwards tertiary structure. Formation of quaternary structure usually involves the "assembly" or "coassembly" of subunits that have already folded. The regular alpha helix and beta sheet structures fold rapidly because they are stabilized by intramolecular hydrogen bonds, as was first characterized by Linus Pauling. Protein folding may involve covalent bonding in the form of disulfide bridges formed

between two cysteine residues or the formation of metal clusters. Shortly before settling into their more energetically favourable native conformation, molecules may pass through an intermediate "molten globule" state.

The essential fact of folding, however, remains that the amino acid sequence of each protein contains the information that specifies both the native structure and the pathway to attain that state. This is not to say that nearly identical amino acid sequences always fold similarly.[10] Conformations differ based on environmental factors as well; similar proteins fold differently based on where they are found. Folding is a spontaneous process independent of energy inputs from nucleoside triphosphates. The passage of the folded state is mainly guided by hydrophobic interactions, formation of intramolecular hydrogen bonds, and van der Waals forces, and it is opposed by conformational entropy.

Disruption of the native state

Under some conditions proteins will not fold into their biochemically functional forms. Temperatures above or below the range that cells tend to live in will cause thermally unstable proteins to unfold or "denature" (this is why boiling makes an egg white turn opaque). High concentrations of solutes, extremes of pH, mechanical forces, and the presence of chemical denaturants can do the same. A fully denatured protein lacks both tertiary and secondary structure, and exists as a so-called random coil. Under certain conditions some proteins can refold; however, in many cases denaturation is irreversible.[11] Cells sometimes protect their proteins against the denaturing influence of heat with enzymes known as chaperones or heat shock proteins, which assist other proteins both in folding and in remaining folded. Some proteins never fold in cells at all except with the assistance of chaperone molecules, which either isolate individual proteins so that their folding is not interrupted by interactions with other proteins or help to unfold misfolded proteins, giving them a second chance to refold properly. This function is crucial to prevent the risk of precipitation into insoluble amorphous aggregates.

Incorrect protein folding and neurodegenerative disease

Aggregated proteins are associated with prion-related illnesses such as Creutzfeldt-Jakob disease, bovine spongiform encephalopathy (mad cow disease), amyloid-related illnesses such as Alzheimer's Disease and familial amyloid cardiomyopathy or polyneuropathy, as well as intracytoplasmic aggregation diseases such as Huntington's and Parkinson's disease. These age onset degenerative diseases are associated with the multimerization of misfolded proteins into insoluble, extracellular aggregates and/or intracellular inclusions including cross-beta sheet amyloid fibrils; it is not clear whether the aggregates are the cause or merely a reflection of the loss of protein homeostasis, the balance between synthesis, folding, aggregation and protein turnover. Misfolding and excessive degradation instead of folding and function leads to a number of proteopathy diseases such as antitrypsin-associated Emphysema, cystic fibrosis and the lysosomal storage diseases, where loss of function is the origin of the disorder. While protein replacement therapy has historically been used to correct the latter disorders, an emerging approach is to use pharmaceutical chaperones to fold mutated proteins to render them functional. Chris Dobson, Jeffery W. Kelly, Dennis Selkoe, Stanley Prusiner, Peter T. Lansbury, William E. Balch, Richard I. Morimoto, Susan L. Lindquist and Byron C. Caughey have all contributed to this emerging understanding of protein-misfolding diseases.

Kinetics and the Levinthal Paradox

The duration of the folding process varies dramatically depending on the protein of interest. When studied outside the cell, the slowest folding proteins require many minutes or hours to fold primarily due to proline isomerization, and must pass through a number of intermediate states, like checkpoints, before the process is complete.[12] On the other hand, very small single-domain proteins with lengths of up to a hundred amino acids typically fold in a single step.[13] Time scales of milliseconds are the norm and the very fastest known protein folding reactions are complete within a few microseconds.[14]

The Levinthal paradox[15] observes that if a protein were to fold by sequentially sampling all possible conformations, it would take an astronomical amount of time to do so, even if the conformations were sampled at a rapid rate (on the nanosecond or picosecond scale). Based upon the observation that proteins fold much faster than this, Levinthal then proposed that a random conformational search does not occur, and the protein must, therefore, fold through a series of meta-stable intermediate states.

Techniques for studying protein folding

Circular Dichroism

Circular dichroism is one of the most general and basic tools to study protein folding. Circular dichroism spectroscopy measures the absorption of circularly polarized light. In proteins, structures such as alpha helicies and beta sheets are chiral, and thus absorb such light. The absorption of this light acts as a marker of the degree of foldedness of the protein ensemble. This technique can be used to measure equilibrium unfolding of the protein by measuring the change in this absorption as a function of denaturant concentration or temperature. A denaturant melt measures the free energy of unfolding as well as the protein's m value, or denaturant dependence. A temperature melt measures the melting temperature (T_m) of the protein. This type of spectroscopy can also be combined with fast-mixing devices, such as stopped flow, to measure protein folding kinetics and to generate chevron plots.

Vibrational circular dichroism of proteins

The more recent developments of vibrational circular dichroism (VCD) techniques for proteins, currently involving Fourier transform (FFT) instruments, provide powerful means for determining protein conformations in solution even for very large protein molecules. Such VCD studies of proteins are often combined with X-ray diffraction of protein crystals, FT-IR data for protein solutions in heavy water (D_2O), or *ab initio* quantum computations to provide unambiguous structural assignments that are unobtainable from CD.

Modern studies of folding with high time resolution

The study of protein folding has been greatly advanced in recent years by the development of fast, time-resolved techniques. These are experimental methods for rapidly triggering the folding of a sample of unfolded protein, and then observing the resulting dynamics. Fast techniques in widespread use include neutron scattering[16] , ultrafast mixing of solutions, photochemical methods, and laser temperature jump spectroscopy. Among the many scientists who have contributed to the development of these techniques are Jeremy Cook, Heinrich Roder, Harry Gray, Martin Gruebele, Brian Dyer, William Eaton, Sheena Radford,

Chris Dobson, Sir Alan R. Fersht and Bengt Nölting.

Energy landscape theory of protein folding

The protein folding phenomenon was largely an experimental endeavor until the formulation of energy landscape theory by Joseph Bryngelson and Peter Wolynes in the late 1980s and early 1990s. This approach introduced the principle of minimal frustration, which asserts that evolution has selected the amino acid sequences of natural proteins so that interactions between side chains largely favor the molecule's acquisition of the folded state. Interactions that do not favor folding are selected against, although some residual *frustration* is expected to exist. A consequence of these evolutionarily selected sequences is that proteins are generally thought to have globally "funneled energy landscapes" (coined by José Onuchic[reference needed]) that are largely directed towards the native state. This "folding funnel" landscape allows the protein to fold to the native state through any of a large number of pathways and intermediates, rather than being restricted to a single mechanism. The theory is supported by both computational simulations of model proteins and numerous experimental studies, and it has been used to improve methods for protein structure prediction and design.

Computational prediction of protein tertiary structure

De novo or *ab initio* techniques for computational protein structure prediction is related to, but strictly distinct from, studies involving protein folding. Molecular Dynamics (MD) is an important tool for studying protein folding and dynamics in silico. Because of computational cost, ab initio MD folding simulations with explicit water are limited to peptides and very small proteins. MD simulations of larger proteins remain restricted to dynamics of the experimental structure or its high-temperature unfolding. In order to simulate long time folding processes (beyond about 1 microsecond), like folding of small-size proteins (about 50 residues) or larger, some approximations or simplifications in protein models need to be introduced. An approach using reduced protein representation (pseudo-atoms representing groups of atoms are defined) and statistical potential is not only useful in protein structure prediction, but is also capable of reproducing the folding pathways.[17]

There are distributed computing projects which use idle CPU time of personal computers to solve problems such as protein folding or prediction of protein structure. People can run these programs on their computer or PlayStation 3 to support them. See links below (for example Folding@Home) to get information about how to participate in these projects.

Experimental techniques of protein structure determination

Folded structures of proteins are routinely determined by X-ray crystallography and NMR.

See also

- Anfinsen's dogma
- Chevron plot
- Denaturation (biochemistry)
- Denaturation midpoint
- Downhill folding
- Equilibrium unfolding
- Folding (chemistry)

- Folding@Home
- Foldit computer game
- Levinthal paradox
- Protein design
- Protein dynamics
- Protein structure prediction
- Protein structure prediction software
- Rosetta@Home
- Software for molecular mechanics modeling

References

[1] Alberts, Bruce; Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walters (2002). "The Shape and Structure of Proteins" (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=mboc4[book]+AND+372270[uid]&rid=mboc4.section.388). *Molecular Biology of the Cell; Fourth Edition*. New York and London: Garland Science. ISBN 0-8153-3218-1.

[2] Anfinsen C (1972). "The formation and stabilization of protein structure". *Biochem. J.* **128** (4): 737–49. PMID 4565129.

[3] Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; Web content by Neil D. Clarke (2002). "3. Protein Structure and Function" (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=stryer[book]+AND+215168[uid]&rid=stryer.chapter.280). *Biochemistry*. San Francisco: W.H. Freeman. ISBN 0-7167-4684-0.

[4] "Science of Folding@Home" (http://folding.stanford.edu/science.html). July 18, 2005. . Retrieved on 2007-04-22.

[5] van den Berg B, Wain R, Dobson CM, Ellis RJ (August 2000). " Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell (http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=306593)". *Embo J.* **19** (15): 3870–5. doi: 10.1093/emboj/19.15.3870 (http://dx.doi.org/10.1093/emboj/19.15.3870). PMID 10921869.

[6] Pace C, Shirley B, McNutt M, Gajiwala K (01 Jan 1996). " Forces contributing to the conformational stability of proteins (http://www.fasebj.org/cgi/reprint/10/1/75)". *Faseb J.* **10** (1): 75–83. PMID 8566551. .

[7] Rose G, Fleming P, Banavar J, Maritan A (2006). "A backbone-based theory of protein folding". *Proc. Natl. Acad. Sci. U.S.A.* **103** (45): 16623–33. doi: 10.1073/pnas.0606843103 (http://dx.doi.org/10.1073/pnas.0606843103). PMID 17075053.

[8] Deechongkit S, Nguyen H, Dawson PE, Gruebele M, Kelly JW (2004). "Context Dependent Contributions of Backbone H-Bonding to β-Sheet Folding Energetics". *Nature* **403** (45): 101–5. doi: 10.1073/pnas.0606843103 (http://dx.doi.org/10.1073/pnas.0606843103). PMID 17075053.

[9] Lee S, Tsai F (2005). " Molecular chaperones in protein quality control (http://www.jbmb.or.kr/fulltext/jbmb/view.php?vol=38&page=259)". *J. Biochem. Mol. Biol.* **38** (3): 259–65. PMID 15943899. .

[10] Alexander PA, He Y, Chen Y, Orban J, Bryan PN. (2007). " The design and characterization of two proteins with 88% sequence identity but different structure and function (http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1906725)". *Proc Natl Acad Sci U S A.* **104** (29): 11963–8. doi: 10.1073/pnas.0700922104 (http://dx.doi.org/10.1073/pnas.0700922104). PMID 17609385.

[11] Shortle D (01 Jan 1996). " The denatured state (the other half of the folding equation) and its role in protein stability (http://www.fasebj.org/cgi/reprint/10/1/27)". *Faseb J.* **10** (1): 27–34. PMID 8566543. .

[12] Kim PS, Baldwin RL (1990). "Intermediates in the folding reactions of small proteins". *Annu. Rev. Biochem.* **59**: 631–60. doi: 10.1146/annurev.bi.59.070190.003215 (http://dx.doi.org/10.1146/annurev.bi.59.070190.003215). PMID 2197986.

[13] Jackson SE (August 1998). " How do small single-domain proteins fold? (http://biomednet.com/elecref/13590278003R0081)". *Fold Des* **3** (4): R81–91. doi: 10.1016/S1359-0278(98)00033-9 (http://dx.doi.org/10.1016/S1359-0278(98)00033-9). PMID 9710577. .

[14] Kubelka J, Hofrichter J, Eaton WA (February 2004). "The protein folding 'speed limit'". *Curr. Opin. Struct. Biol.* **14** (1): 76–88. doi: 10.1016/j.sbi.2004.01.013 (http://dx.doi.org/10.1016/j.sbi.2004.01.013). PMID 15102453.

[15] C. Levinthal (1968). " Are there pathways for protein folding? (http://www.biochem.wisc.edu/courses/biochem704/Reading/Levinthal1968.pdf)". *J. Chim. Phys.* **65**: 44–5. .

[16] Bu Z, Cook J,Callaway DJE (2001). "Dynamic regimes and correlated structural dynamics in native and denatured alpha-lactalbuminC". *J Mol Biol* **312** (4): 865-873. doi: 10.1006/jmbi.2001.5006 (http://dx.doi.org/

10.1006/jmbi.2001.5006).
[17] Kmiecik S and Kolinski A (2007). "Characterization of protein-folding pathways by reduced-space modeling". *Proc. Natl. Acad. Sci. U.S.A.* **104** (30): 12330–5. doi: 10.1073/pnas.0702265104 (http://dx.doi.org/10.1073/pnas.0702265104). PMID 17636132.

External links

- FoldIt - Folding Protein Game (http://fold.it/portal/info/science)
- Folding@Home (http://www.stanford.edu/group/pandegroup/folding/about.html)
- Rosetta@Home (http://boinc.bakerlab.org/rosetta)

Article Sources and Contributors

System *Source*: http://en.wikipedia.org/w/index.php?oldid=293926258 *Contributors*: .:Ajvol:., 16@r, 64200a, 7195Prof, AbsolutDan, Adam78, Aesopos, Aimak, Alansohn, Alerante, Alink, Altenmann, Amalas, Amayoral, AnakngAraw, Ancheta Wis, Anclation, AndonicO, AndriuZ, Antandrus, Anwar saadat, AprilSKelly, Arendedwinter, Arthur Rubin, Ashley Thomas, Bajjoblaster, Blazotron, Bobo192, Boccobrock, Bomac, Bonaparte, BradBeattie, Brichcja, BrokenSegue, Bry9000, CJLL Wright, COMPATT, CX, Campjacob, Cerber, Charlesverdon, Chasingsol, Chris cardillo, Chung33, Ckatz, Compfun, Coppertwig, Countincr, Crissidancer88, D. Recorder, DSRH, DVD R W, DabMachine, Daniellemonkey64, Darth Panda, Denys, DerHexer, DocWatson42, Dominus, Dpv, EdC, Efe, Eliyak, Ely ely82, Epbr123, Equilibrioception, Erebus555, Erick.Antezana, Evercat, Evilphoenix, Fanra, Fenice, Filemon, FlyHigh, Fresheneesz, Frongle, Frymaster, Fvw, GPHemsley, Galoubet, Giftlite, Glenn, Globalanarchist, Graham Berrisford, Grubber, Grunt, Helix84, Hetar, HisSpaceResearch, Husond, Hveziris, IZAK, Iancarter, Informatwr, Iwillgoslow3126, J.delanoy, Jackfork, Jcs2006, Jgd9, Jiang, Jojhutton, Jossi, Joymmart, Jpbowen, Jrgetsin, Jtneill, Justin1234345, Jwdietrich2, Kazakmike, Kazooou, Kevin Baas, Kevin Breitenstein, Khatru2, Klenod, KnowledgeOfSelf, La goutte de pluie, LeaveSleaves, Lexor, Lightmouse, LtNOWIS, Luna Santin, MGTom, Mange01, Mani1, Manuelkuhs, Marc Venot, Marek69, Maxí, Mdd, Mentifisto, Michael Hardy, MichaelBillington, Mike Cline, Miriel, Mkoval, Mkoyle, Nahtmmm, NatusRoma, Newscript, Niemeyerstein en, Nihil novi, Normxxx, Octahedron80, Oddeivind, OliviaGuest, Optikos, Ordermaven, Orgwiki, Orphic, Ott, Patho, Patrick, Pde, Phanerozoic, Phil Sandifer, Philwelch, Photouploadrr, Qst, Quester67, Ram4eva, Ratjed, Rdsmith4, Reddi, Res2216firestar, RexNL, RichardF, Rjm at sleepers, RoyBoy, S Roper, ST47, Samuel Curtis, ScierGuy, Scott Gall, Sevilledade, Sholto Maud, Simoes, Simply south, Smjg, Snek01, Snowmanradio, Someguy1221, Soulwork, SpookyMulder, StaticVision, Steven J. Anderson, Stevertigo, StradivariusTV, Strait, Sunray, Surfer2233, Thanatos666, The Anome, Thseamon, Tiddly Tom, TobyJ, Tobyw87, TowerDragon, Tporter2010, Treisijs, Trevor MacInnis, Vapier, Vlaton, Wiccan Quagga, Wmahan, Wolfkeeper, Woohookitty, Xdcdx, Xtifr, Xyzzyplugh, Yidisheryid, Zeneka, Zntrip, Zzuuzz, Александър, Милан Јелисавчић, 277 anonymous edits

Dynamics *Source*: http://en.wikipedia.org/w/index.php?oldid=294284169 *Contributors*: (jarbarf), 16@r, 212.59.194.xxx, ABusa, AVand, Ap, Apapte, Army1987, BanyanTree, Bbtommy, Boleyn2, Brendy15, Brews ohare, Bright.eyes.love, Bryan Derksen, Burburchacha, Camembert, Canaen, Conversion script, Danceswithzerglings, Dynamic, Dzordzm, Geniac, George100, Heron, Imusade, Inwind, It Is Me Here, J.delanoy, Jaibe, Javawizard, Jtvisona, Leafyplant, MAG1, Mac, MagneticFlux, Merphant, Murdockh, NBS, NaibStilgar, Nsaa, Off!, Piet Delport, Rinzai uk2, Robofish, Sadi Carnot, SebastianHelm, Stephenb, THEN WHO WAS PHONE?, TaffRees1203, Tarquin, Tcncv, Template namespace initialisation script, The Cat and the Owl, TheAllSeeingEye, ThorinMuglindir, Tristanlbailey, Ugur Basak, Usgnus, Wolfkeeper, 57 anonymous edits

Dynamical system *Source*: http://en.wikipedia.org/w/index.php?oldid=293524521 *Contributors*: 0, 195.186.254.xxx, Aaronp808, Adam majewski, Aleksandar Guzijan, Altenmann, AntOnTrack, Ap, Athkalani, AxelBoldt, Bluemoose, Brazzouk, CX, Caesium, Charles Matthews, Chetvorno, Chopchopwhitey, Complex01, Complexica, Cumi, Cutler, Daniele.tampieri, Dino, Dmharvey, Dysprosia, EPM, El C, Epbr123, Epolk, Everyking, Evilphoenix, Filemon, Filur, Finn-Zoltan, Fredrik, Gandalf61, Giftlite, Headbomb, Hesam7, Highlightened, Hve, Hydroli, Jabernal, Jay Gatsby, Jeff3000, Jeffrey Smith, JerrySteal, Jitse Niesen, JocK, Jugander, K-UNIT, KYPark, Karol Langner, Kayvan45622, Kenneth M Burke, Kotepho, Kzzl, Lakinekaki, Lightmouse, Linas, ManiacK, Marj Tiefert, MathMartin, Mathmanta, Mct mht, Mdd, Meersan, Michael Hardy, Milly.mortimer, Msh210, Neelix, Nn123645, Noeckel, Oleg Alexandrov, Orange Death, OrgasGirl, Patrickdepinguin, PetaRZ, Pgan002, Phys, PlatypeanArchcow, RedWolf, Reddi, Reinderien, Revolver, Rhythmiccycle, Rich Farmbrough, Rintrah, Sadi Carnot, Salgueiro, Salix alba, Sam Korn, Samuelbf85, SilverSurfer314, Snoyes, Solace098, Sverdrup, Template namespace initialisation script, The Anome, The wub, Tobias Hoevekamp, Tomisti, Tommyjs, Tosha, Volfy, Voretus, Waitati, WhiteC, WillowW, XJamRastafire, XaosBits, Zsniew, 119 anonymous edits

Complex system *Source*: http://en.wikipedia.org/w/index.php?oldid=293710129 *Contributors*: 130.94.122.xxx, Acadac, Adam M. Gadomski, AdamRetchless, Adastra, Amalas, Antoni Barau, Apothecia, Ballista, Betacommand, Betamod, Bryan Derksen, Calltech, Cazort, Centrx, Childhoodsend, Choesarian, Chopchopwhitey, Christopher Agnew, Cmbarton54, Danthemankhan, DarwinPeacock, Dggreen, Diza, Duracell, ELApro, Emmazunz84, Erkan Yilmaz, Examtester, FT2, Fmrafka, G.A.S, GarOgar, Grosscha, Hu, IPSOS, Ifaomo, Iterator12n, J.A.McCoy, JFromm, Jmeppley, Jon Awbrey, Jpbowen, Jrgetsin, Jrouquie, Jshadias, Jwdietrich2, KYPark, Karada, Karol Langner, Kevinbrowning, Kimiko, La goutte de pluie, Letranova, Lexor, Lucasbfr, MagnaMopus, Magntuve, Malvaro, MathShaman, McVities, Mdd, Megane, Mfmoore, Miketam, MisterCharlie, Mkoval, Mr3641, Msuzen, N2e, Nbrown@unicistinstitute.org, Nigholith, Noeckel, Northfox, Oatmealcookiemon, Ott, Overix, Paresnah, Passw0rd, Peterdjones, Pjacobi, RedWolf, RodC, Ronz, Rsanz, Salix alba, Sectryan, Sengkang, Simoes, Sina2, Sir Nicholas de Mimsy-Porpington, Spook`, Steamturn, Steve Masterson, Supermanchander, The wub, Thsgrn, Took, Truthflux, Tschild, Versageek, Vonkje, Xtifr, Xyzzyplugh, Yhkhoo, Yonidebest, Zzuuzz, 148 anonymous edits

Complexity *Source*: http://en.wikipedia.org/w/index.php?oldid=294117635 *Contributors*: 1diot, APH, Ahoerstemeier, Ak6128, Allen Moore, Amalas, Andreas Kaufmann, Ashaktur, Bcastel3, Bookandcoffee, COMPATT, CRGreathouse, CX, Cacycle, Cambyses, Celendin, Conversion script, Corti, Damiens.rf, David Haslam, Dessimoz, Dggreen, Digx, Douglas R. White, Dpv, DrPTThomas, Duracell, EBRJoseph, ESkog, Elvismcgrady, Emersoni, Erudecorp, Examtester, Flammifer, Flipjargendy, Fram, G2kdoe, Giftlite, GraemeL, GregorB, Hirzel, Ingenuity Arts, Innohead, Ioannes Pragensis, Isaac, Iterator12n, Jccarteron, Jfurr1981, JiE, John courtneidge, Jon Awbrey, Joy, Jpbowen, Jstanley01, Kiand, Kilmer-san, Ksyrie, Kurykh, Lambiam, Letranova, Lexor, Liberatus, Linda Vandergriff, MER-C, MLCommons, MMX, Malvaro, Marco Krohn, Matthew Stannard, Matthew Yeager, Maurice Carbonaro, Mav, MaxHund, Mdd, Monkeez, Mr3641, Mrholybrain, Multipundit, Nbrown@unicistinstitute.org, Nick Green, Nicolesc, Ninadelis, Ninjavitus, Oleg Alexandrov, Ordermaven, Ott, P3d0, Patrick, Pjt111, Pleasantville, Pring, Rade Kutil, Rainman321, Ripper234, Rjwilmsi, Robinh, RodC, Rvsole, Ryguasu, Sacramentis, Samtheboy, Sbonin, Shaane, Simguy, SimonP, Skier Dude, Smithfarm, Srinivasasha, StN, Stanbeek, Suisui, Tasc, TimVickers, Tom harrison, TowerDragon, Trovatore, Veinor, Wricardoh, Zepard, Zinp, 107 anonymous edits

Complex systems *Source*: http://en.wikipedia.org/w/index.php?oldid=294489773 *Contributors*: Ashaktur, Ask123, Bcastel3, Cmbarton54, Complexica, DarwinPeacock, Dggreen, DrPTThomas, DrSethBullock, Erkan Yilmaz, Fuller.s.a., Gaius Cornelius, Giftlite, Jamessungjin.kim, Jcbohorquez, Joehawayek, Margaret9mary, Masoninman, Mdd, Mr3641, N2e, Nbaig, Nigholith, Normonics, Philr 99, Qwfp, Resonance or resonant?, Ronz, Rosemartin, Sevenoaks, SimonP, Smee44, Spaecious, Will R Turner, ^demon, 46 anonymous edits

Systems biology *Source*: http://en.wikipedia.org/w/index.php?oldid=294364649 *Contributors*: APH, Aciel, Alan Liefting, AlirezaShaneh, Amandadawnbesemer, Amirsnik, Andreas td, Arthena, Asadrahman, Aua, Bad Cat, Bamess, BatteryIncluded, Bci2, Bio-ITWorld, Biochaos, Biophysik, Blueleezard, Boku wa kage, Broadbeer, CRGreathouse, CX, Can't sleep, clown will eat me, Captain-tucker, Ceolas, Charlenelieu, CharonZ, Ckatz, Claronow, ColinGillespie, Cquan, Crodriguel, D6, DanielNuyu, Delta Xi, Dmb000006, Drgarden, Droyarzun, Duelist135, Edaddison, Edward, Electric sheep, Erick.Antezana, Erkan Yilmaz, Eveillar, FLeader, Fences and windows, Fenice, Fletcher04, Foggy29, Fredrik, Garychurchill, Gcm, Gdrahnier, Ggonnell, Giftlite, Gwolfe, Heisner, HexiToFor, IPSOS, JRSocInterface, JaGa, Jdegreef, Jethero, Jondel, Jongbhak, Jpbowen, JulioVeraGon, Jwdietrich2, Kane5187, Karthik.raman, Kcordina, KirbyRandolf, Kku, Klenod, Klipkow, Lauranrg, Lenov, Letranova, Lexor, Lilia Alberghina, Linkman21, Lkathmann, Massbiotech, Mdd, Michael Fourman, Michael Hardy, Miguel Andrade, MikeHucka, Mkotl, Mkuiper, Mmxx, Mobashirgenome, Molelect, N2e, NBeale, NIH Media, Natelewis, Nbaliga, Neilbeach, Nemenman, Netsnipe, Nick Green, O RLY?, Ombudsman, Opertinicy, Patho, PaulGarner, Pkahlem, Pvosta, Quarl, Rajah, Reggiebird, Rich Farmbrough, Rjwilmsi, Robnpov, Rvencio, Rwcitek, Satish.vammi, SeeGee, Senu, Seuss01, Sholto Maud, Srlasky, Steinsky, Stewartadcock, Synthetic Biologist, Tagishsimon, Template namespace initialisation script, Thorwald, Triamus, Unauthorised Immunophysicist, Vangos, Versus22, Vonkje, WLU, Waltpohl, Wavelength, Whosasking, Xeaa, Zargulon, Zlite, Zoicon5, عیسی, 350 anonymous edits

Neural network *Source*: http://en.wikipedia.org/w/index.php?oldid=293708296 *Contributors*: A tzoumerkioti, Adodge, Ahoerstemeier, Ahyeek, Alan Au, Alanbly, Aleksey Vayner, AnAj, Andrew C. Wilson, Andrew4010, Anetode, Antandrus, Anthony Appleyard, Antonio Lopez, Appljax123, Arcadian, ArielGold, Arru, Arun143123, Avenue, Az1568, B9 hummingbird hovering, BSTRhino, Batch1, Benbread, BeverlyCrusher, Bhadani, BlaiseFEgan, Bonnarj, Bpavel88, Brian0918, BrokenSegue, Bungopolis, Butros, CX, Caltas, Cameltrader, Canley, Cantus, Cap'n Refsmmat, Cburnett, CharlesGillingham, Chase me ladies, I'm the Cavalry, ChrisKennedy, Clarissaelsinger, Coladie, CommodiCast, Cs007, D.Strand, Daev, Dan Guan, Dan100, Dank, Darkride, David R. Ingham, Daytona2, Delirium, Denn, Denoir, Deodar, Dicklyon, Diegotorquemada, Digfarenough, Dlohcierekim's sock, Doczilla, Doyley, Driper, Dstorkey, Duncharris, Dusti, Dysprosia, Ecalot, Efbfweborg, Elsendero, Epsilon60198, EyrianAtWork, Fangz, Fbarl, Fenice, Figma, Fippy Darkpaw, Foobar, G.A.S, Gauss, Ghepeu, Giftlite, Gilhamto, Gioto, Goyston, Guntram, Gurch, Herschel, HighKing, Holme215, Hongooi, Hu, Hu12, Ines it, Itistoday, Jacj, James-davis, Jhedengren, Jitse Niesen, John Broughton, John Quiggin, JohnJBarton, Jon Awbrey, JonEAhlquist, JonHarder, Jpbowen, Jwdietrich2, Jwmillerusa, Kaimiddleton, Karbinski, Khoikhoi, Kl4m, Klaus, Kuru, La goutte de pluie, LaggedOnUser, Levin, LilHelpa, Lordvolton, Luna Santin, MER-C, MSchmahl, MaBe, Maian, Male1979, Mamling, Mareino, Mathaddins, Matthew, Mattisse, Mbell, Mcasl, Mcsee, Mdd, Mekem, Michael Hardy, MikeWren, Mosquitopsu, Mr.winston, Murx, Mwanner, Mysid, Nguyengiap84, Nicholasnice, Nikai, Nikoladie, Ninly, Nunh-huh,

Oleg Alexandrov, Olethros, Omegatron, Onasraou, Opelio, Originalname37, Pakcw, Pankaj Sharma Computers, Passino, Pavel Vozenilek, Pepperpot80, Pescadero, Peter.kabai, Pietro16, Pjetter, Prolog, PseudoSudo, Psyphen, RJASE1, RWillwerth, Raven1977, Rayc, Reddi, Reedbeta, Reedy, Requestion, RexNL, Riana, Rich Farmbrough, Richard001, Ritchy, Rjwilmsi, Rmyeid, Romanm, Ronz, Royalguard11, Rtcpenguin, Rwp, SDas, SallyForth123, SamuelRiv, Satishsnandihalli, Schwnj, Seans Potato Business, Shadmanzafar, Shawndkc, Slakr, Sonyjose, SparkOfCreation, Spazzm, Spudtater, Steve Crossin, Stevenj, Stoni, Strangerer, Strife911, Supten, Tarotcards, Tbutzon, That Guy, From That Show!, The sunder king, TheJosh, Thomasmeeks, Titoxd, Tribaal, Trilobitealive, Twri, Vendettax, Violetriga, Waggers, Wbrameld, Wduch, Wernerseyfried, Will316, Wingchi, Wknight8111, Wmahan, Wraithdart, Zeiden, Zybler, Ömer Cengiz Çelebi, 439 anonymous edits

Gene regulatory network *Source*: http://en.wikipedia.org/w/index.php?oldid=291641856 *Contributors*: Adiel lo, BigHaz, BryanD, Carl T, Charles Matthews, DanielYamins, Delaszk, FlorianMarkowetz, Frederickmercury, Gaius Cornelius, Jknabe, Jpbowen, Jugander, Kku, Lexor, Massbiotech, Mdd, Memestream, Michael Hardy, Mortalsyn, Narayanese, Natelewis, Neurocompute, PFHLai, Peak, Ronz, Rverduzco, Slambeck, Stewartadcock, Tekhnofiend, WAS 4.250, WriterHound, Wynand.winterbach, 84 anonymous edits

Genomics *Source*: http://en.wikipedia.org/w/index.php?oldid=294542432 *Contributors*: *drew, 5dPZ, AdamRetchless, Adenosine, Alex naish, Andreadb, Anthere, Apers0n, Aphextwin5678, AxelBoldt, Barrylb, Bill.albing, Braidwood, Branttudor, Brion VIBBER, Bryan Derksen, Calvinthe1337, Ceyockey, Combio, CommodiCast, DabMachine, Dave Nelson, David D., Dekisugi, Dicklyon, Dmb000006, DoctorDNA, Dolfin, Drgarden, El C, Eubulides, Eugene, Fred Bradstadt, Gary King, Genometer, GeoMor, Ghostoroy, Giftlite, Gilliam, Habj, Hadal, Hbent, Heron, Jenks, Jethero, Jfdwolff, Joconnol, Joerg Kurt Wegner, Johntex, Johnuniq, Jongbhak, Larssono, Lexor, Lightmouse, Lost-theory, Mariusz Biegacki, Marj Tiefert, Mav, Mike Lin, Natarajanganesan, Nitwitpicker, Oleginger, Para, Peak, Pgan002, Pharmtao, Pion, Pvosta, Quizkajer, RandomP, Recury, Rein0299, Rich Farmbrough, Ronz, Rppgen, Sairen42, Scewing, Shanes, SimonP, Sjjupadhyay, Spitfire ch, Springmn, Starshadow, Stonedhamlet, Syp, Template namespace initialisation script, TheObtuseAngleOfDoom, Thkim75, Thorwald, Tiddly Tom, Toddst1, Touchstone42, Unyoyega, VashiDonsk, W09110900, Wavelength, Wayne530, Williamb, Wmahan, Wuzzybaba, Xanthoptica, ZayZayEM, ZimZalaBim, 129 anonymous edits

Genetic algorithm *Source*: http://en.wikipedia.org/w/index.php?oldid=295135438 *Contributors*: "alyosha", .:Ajvol:., 2fargon, A. S. Aulakh, A.Nath, AAAAA, Aabs, AdamRaizen, Ahoerstemeier, Ahyeek, Alex Kosorukoff, Algorithms, Aliekens, AlterMind, Andreas Kaufmann, Antandrus, AnthonyQBachler, Antonielly, Antzervos, Arbor, Arkuat, Armchair info guy, Arthur Rubin, Artur adib, Asbestos, AussieScribe, Avinesh (usurped), BAxelrod, Baguio, Beetstra, Bidabadi, Bjtaylor01, Bobby D. Bryant, Bockbockchicken, Bovineone, Bradka, Brat32, Breeder8128, Brick Thrower, Brinkost, BryanD, Bumbulski, CShistory, Carl Turner, Centrx, CharlesGillingham, Chipchap, Chocolateboy, Chopchopwhitey, CloudNine, Cngoulimis, Cnilep, CoderGnome, Conway71, CosineKitty, Cpcjr, Crispin Cooper, Curps, DabMachine, David Martland, DavidCBryant, DerrickCheng, Diroth, DixonD, Diza, Djhache, Duncharris, Dúnadan, Edin1, Edrucker, Edward, Eleschinski2000, Esotericengineer, Euhapt1, Evercat, Ewlyahoocom, Felsenst, Freiberg, Frongle, Furrykef, Gaius Cornelius, George100, Giftlite, Goobergunch, Gpel461, GraemeL, Gragus, GregorB, Grein, Grendelkhan, Guang2500, Hellisp, Hike395, Hippietrail, Hu, J.delanoy, Jasper53, Jcmiras, Jeff3000, Jitse Niesen, Josilber, Jwdietrich2, Jyril, KaHa242, Kane5187, Kdakin, Keki Burjorjee, Kindyroot, Kjells, Kku, Klausikm, KrakatoaKatie, Kuzaar, Kwertii, Kyokpae, LMSchmitt, Larham, Lawrenceb, Lee J Haywood, Leonard^Bloom, Lexor, LieAfterLie, Ludvig von Hamburger, MER-C, Madcoverboy, Magnus Manske, Malafaya, Male1979, Marco Krohn, Mark Krueger, Marksale, Massimo Macconi, MattOates, Mdd, Michael Hardy, MikeMayer, Mikeblas, Mikołaj Koziarkiewicz, Mohan1986, Mohdavary, Mpo, Negrulio, Nikai, No1sundevil, Nosophorus, Novablogger, Oleg Alexandrov, Oli Filth, Omicronpersei8, Oneiros, Open2universe, Otolemur crassicaudatus, Paskornc, Pecorajr, Pegship, Pelotas, PeterStJohn, Phyzome, Plasticup, Postrach, Poweron, Projectstann, Pruetboonma, Purplesword, Qed, QmunkE, Radagast83, Raduberinde, Ratfox, Redfoxtx, RevRagnarok, Rfl, Riccardopoli, Rjwilmsi, Roberta F., Robma, Ronz, Ruud Koot, SDas, SSZ, ST47, SamuelScarano, Sankar netsoft, Scarpy, Shyamal, Silver hr, Simeon, Simpsons contributor, SlackerMom, Smack, Soegoe, Spoon!, SteelSoul, Stefano KALB, Steinsky, Stephenb, Stewartadcock, Stochastics, Sundaryourfriend, Swarmcode, Tailboom22, Tameeria, Tapan bagchi, Tarantulae, Tarret, Taw, Techna1, TempestCA, Temporary-login, Terryn3, Texture, The Epopt, Thomas weise, Thric3, TimVickers, Timwi, Toncek, Toshke, Tribaal, Tulkolahten, Twelvethirteen, Unixcrab, Unyounyo, Useight, Versus, Vietbio, Vignaux, Vincom2, VladB, Waveguy, William Avery, Wjousts, Xiaojeng, Xn4, Yinon, Yuanwang200409, Zawersh, Zwgeem, 428 anonymous edits

Metabolic network *Source*: http://en.wikipedia.org/w/index.php?oldid=231870375 *Contributors*: Blastwizard, Ceyockey, Oleginger, PDH, TheParanoidOne, TimVickers, Zephyris, 2 anonymous edits

Metabolic network modelling *Source*: http://en.wikipedia.org/w/index.php?oldid=285580489 *Contributors*: AKC, Barticus88, Choster, Dylan Lake, Glane23, Karthik.raman, Leptictidium, Mdd, Noalaignorancia, PDH, Parakkum, Piotrus, Ragesoss, Sarebi, TimVickers, Uncle G, 20 anonymous edits

Protein-protein interaction *Source*: http://en.wikipedia.org/w/index.php?oldid=293638290 *Contributors*: 56869kltaylor, 7bd1, A wandering 1, Alboyle, Apfelsine, Ashcroft, Bci2, Biophys, Clicketyclack, Cpichardo, D-rew, DarkSaber2k, Delldot, Djstates, Dsome, FreeKill, Giftlite, GracelinTina, Hendrik Fuß, Hotheartdog, Jeandré du Toit, Jkbioinfo, Jkwaran, Jn3v16, Jongbhak, Keesiewonder, Kkmurray, Kuheli, Kyawtun, Lafw, Lemchesvej, Lenticel, Longhair, Meb025, Michael Hardy, MichaelMcGuffin, Miguel Andrade, NickelShoe, Ninjagecko, Nnh, Rajah, Reb42, Riana, Ronz, Seans Potato Business, Snowolf, TheParanoidOne, Thorwald, Uthbrian, Victor D, Wenzelr, Whosasking, Wintrag, 67 anonymous edits

Proteomics *Source*: http://en.wikipedia.org/w/index.php?oldid=290412197 *Contributors*: 2over0, Aiko, Akriasas, Alan Liefting, AlistairMcMillan, Apfelsine, ArazZeynili, Babbage, Bdekker, Bezapt, Bill.albing, Borgx, Boy in the bands, Bryan Derksen, Calimo, CathCarey, Chaos, Chris the speller, Cjb88, Clicketyclack, Cpiggee, Dancter, Dave Nelson, Dfornika, Dhart, Dicklyon, Djstates, Dmb000006, Download, El C, Flowanda, Gacggt, Gaius Cornelius, GraemeLeggett, Graham87, Hadal, Iamunknown, IlyaHaykinson, Iridescent, Itub, Jambell, Janbrogger, Jason.nunes, JeLuF, Jfdwolff, Johannesvillavelius, JonHarder, Jóna Þórunn, Kbelhajj, Kevyn, Kjaergaard, Kkmurray, Kku, Kosigrim, Kukini, Lexor, LiDaobing, Lights, Lupin, Lysdexia, MStreble, Maartenvdv, Mani1, Manyanswer, Mathisuresh, Mav, Mjensen@nas.edu, N2e, Nick Y., Nina Gerlach, Nwbeeson, Oddwick, Oleginger, Ottava Rima, PDH, Paul Drye, Pcarvalho, Perissinotti, Pgan002, Pganas, Plumbago, Proteomicon, ProveIt, Pscott22, Pvosta, Quintote, RDBrown, Raymond Hui, Remi0o, Rich Farmbrough, Roadnottaken, Sater, Schutz, Senski, Shizhao, Smcarlson, Someguy1221, Sorfane, Sp3000, Springatlast, Srlasky, StevieNic, Systemfolder, Template namespace initialisation script, TestPilot, Tim@, Tregonsee, Trevor MacInnis, Triwbe, Tstrobaugh, Versus22, Voyagerfan5761, Whosasking, Wisdom89, Xeaa, Zashaw, ZimZalaBim, आशीष भटनागर, 211 anonymous edits

Interactomics *Source*: http://en.wikipedia.org/w/index.php?oldid=293358442 *Contributors*: Bci2, Bdevrees, Erick.Antezana, Erodium, Jong, Jongbhak, Karthik.raman, Lexor, Llull, Niteowlneils, PDH, Pekaje, Rajah, Tucsontt, 8 anonymous edits

Mathematical biology *Source*: http://en.wikipedia.org/w/index.php?oldid=294590245 *Contributors*: Adoniscik, Agilemolecule, Agricola44, Alan Liefting, Anclation, Andreas td, Aua, Audriusa, Bci2, Bduke, Berland, BillWSmithJr, Ceyockey, Charvest, Chopchopwhitey, Commander Nemet, Constructive editor, Cquan, Den fjättrade ankan, Durova, Dysprosia, Eduardoporcher, Fredrik, Gandalfxviv, Geronimo20, Guettarda, Henriok, Honeydew, Imoen, Jag123, Jaibe, Jennavecia, JonHarder, Jonsafari, Jpbowen, Jwdietrich2, Karl-Henner, Kripkenstein, Leptictidium, Lexor, Lquilter, M stone, MATThematical, Malcolm Farmer, Mathmoclaire, Maurreen, Melcombe, Michael Hardy, Oldekop, Oli Filth, Open4D, Owlmonkey, Percy Snoodle, PeterStJohn, PhDP, Plw, Porcher, Rich Farmbrough, Sintaku, Sir hubert, Squidonius, Ssavelan, StN, Stemonitis, Tompw, Triwbe, Vina, Wavelength, 69 anonymous edits

Theoretical biology *Source*: http://en.wikipedia.org/w/index.php?oldid=295122668 *Contributors*: Alamowit, Alan Liefting, Bci2, Boninho, Boresal, Cacycle, Cazort, Den fjättrade ankan, Dj Capricorn, Doctorage, Dr Oldekop, Eliezg, G716, Guettarda, Hans Adler, Hydro2, Jeff3000, Jrf, Jwdietrich2, Kakugo, Lexor, Lquilter, Marashie, Marsian, Maurreen, Mdd, Occhanikov, Oldekop, PhDP, Pietrosperoni, Porcher, Ruud Koot, Tassedethe, Tompw, Txomin, WebDrake, WilliamBonfieldFRS, 24 anonymous edits

Bifurcation theory *Source*: http://en.wikipedia.org/w/index.php?oldid=288854389 *Contributors*: Amitkashwar, Anandhan, Arthur Rubin, Athkalani, Cat2020, Deeptrivia, Dharma6662000, Dmr2, Giftlite, Guckurev, Harriv, Headbomb, Hesacon, JanSuchy, Jheald, Jugander, K-UNIT, Kenneth M Burke, Linas, Mathmoclaire, Mmernex, Paul Matthews, Pizza1512, Pt, Rhythmiccycle, Saziel, Squidonius, Stefankroon, Sviemeister, Voretus, Zanaq, Zsniew, 33 anonymous edits

Catastrophe theory *Source*: http://en.wikipedia.org/w/index.php?oldid=274510127 *Contributors*: Aetheling, Alberto da Calvairate, Alexandrov, Bryan Derksen, Chabacano, D4g0thur, DavidWBrooks, Fenice, Giftlite, Hgilbert, Hotfeba, Iamfscked, Itai, Jheald, Justin Foote, Jwdietrich2, Kbk, Keilana, Linas, Magmi, Maphyche, Matusz, Mdd, Michael Hardy, Mkch, Mr giant sloth, Oleg Alexandrov, Porcher, Profangelo, Roadrunner, Salix alba, Seidenstud, Silverfish, Ultraviolet scissor flame, Vinograd19, Wasell, Zirconscot, 31 anonymous edits

Chaos *Source*: http://en.wikipedia.org/w/index.php?oldid=293195071 *Contributors*: 999, A2Kafir, Aces lead, Adinls, Alansohn, Alexius08, AlphaEta, AmosWolfe, Ana Pérez Vega, Andre Engels, Andre'd. lopez, AndrewDressel, Anetode, AngelOfSadness, Angela, Angr, Antandrus, Anthony Appleyard, Ashami, Atlant, AtmanDave, BD2412, Badgernet, BearcatFootball, Bearly541, Ben-Zin, Bfigura's puppy, Bigwyrm, Blue loonie, Bobo192, BorgQueen, Caltas, Can't sleep, clown will eat me, Carabinieri, Carlossuarez46, Castanea dentata, Catgut, Chameleon, ChaosPirate, Chaoschad220, Chaosfeary, Chaosgio, Chaosplanet, ChazYork, Chrislk02, Col. Trinquier, Conversion script, CorbinSimpson, Ctachme, Curps, CyberSkull, Cyktsui, Cyp, D J L, DIEGO RICARDO PEREIRA, DJ Clayworth, DMacks, Dakart, Danhuby, Daniel Lawrence, David R. Ingham, Dawn Bard, Dayewalker, De728631, Deor, Diberri, Didactohedron, DirkvdM, Dpv, Dylan Lake, Dysepsion, Emijrp, Eschnett, FETuriousness, Fiyinw, Flammifer, FlavusBelisarius, Flewis, Former user 6, Frecklefoot, Frehley, Frosty Blizzard, Frymaster, Gfalco, Giftlite, GraemeMcRae, Granpuff, Grapevine.79, Grenavitar, Gtrmp, Hanuman Das, Hapsiainen, Hede2000, Hibana, HumanFrog, Hydragon, I need a better name, IPSOS, Icelight, Improbcat, Intothewoods29, Ioscius, JForget, Janey Dowerstruffel, Jazhara7, Jbasson, Jeff3000, Jmlipton, Joelster, Johan Elisson, John courtneidge, Jordansc, Josh Grosse, Jugander, Justin.case.meiners, K-UNIT, KDerbyshire, Keesiewonder, Khaosinfire, Kimon, Kku, Kpjas, KumoriKage, Lambiam, Leoadec, LestatdeLioncourt, Lexor, LizardWizard, Localhandson, Lokal Profil, LordRM, Luxdormiens, MER-C, Macmelvino, Maestlin, Marcprumpeler, Matthew Husdon, MaulYoda, Mav, Mdd, Meachly, Megatron must be stopped, Mhking, Michael Hardy, Mipadi, Mjgullans, Ms2ger, Mushroom, NHRHS2010, Nai, Nakon, NawlinWiki, Neo-Jay, NickelShoe, NielsenGW, Nishkid64, Oleg Alexandrov, Ossmann, Pak21, Parleyvous, Paul Matthews, Paul Murray, Pearle, Pendimethalin, Peng, Pharillon, Piano non troppo, Pjacklam, Pratheepps, Premeditated Chaos, Prsephone1674, Puckly, Qmwne235, RPH, Racooon, Radagast3, Radagast83, RadicalBender, RatedRestricted, RayAYang, Razorflame, ReallyNiceGuy, RedWolf, Renamed user 4, RepublicanJacobite, Revolver Ocelot, Ripper234, Robin Hood, Rundquist, SJP, Sabzafrica, Sam Hocevar, Samaritan, Samw, Sangajin, SasIsBlack, Satanael, Schneelocke, Selket, Semorrison, Shariku Onikage, Shiftchange, Sidneep, Sik0fewl, Sionus, Sjakkalle, Smack, SparrowsWing, Ste4k, Steel, Stephen e nelson, Steveprutz, Strait, Supermorff, Synergy, Tchalvak, The Ronin, The Singing Badger, The Traveller, TheDoober, TheRingess, TheSpectator, Thehotelambush, Thingg, Tkandell, Tomisti, Torzsmokus, Trifon Triantafillidis, Tubedogg, Tucci528, Tunnels of Set, Twirk88, Until It Sleeps, Urhixidur, Useight, V.V zzzzz, Van helsing, Versus22, Vicarious, Vinmax, Vishvax, Vuo, Wetman, Whispering, Willking1979, Wyatt915, Xpspectre, Zacharyvince, Zack Holly Venturi, Zanimum, Zikar, ₹, 404 anonymous edits

Chaotic system *Source*: http://en.wikipedia.org/w/index.php?oldid=16370981 *Contributors*: -

Chaos theory *Source*: http://en.wikipedia.org/w/index.php?oldid=294741812 *Contributors*: 19.168, 19.7, 3p1416, 4dhayman, 7h3 3L173, Abrfreek777, Academic Challenger, Adam Bishop, Africangenesis, Ahoerstemeier, Alansohn, Alereon, Alex Bakharev, Alex brollo, Alpharius, AngelOfSadness, Antandrus, Apsimpson02, Aranherunar, Aremith, Arjun01, Arthena, Arthur Rubin, Arvindn, Asyndeton, Atomic Duck!, Axel-Rega, AxelBoldt, Ayonbd2000, Banno, Barnaby dawson, Ben pcc, Benjamindees, Bernhard Bauer, Bevo, Bhadani, Blazen nite, Brazucs, Brewhaha@edmc.net, Brian0918, Brighterorange, Brother Francis, BryanD, Bryt, Buchanan-Hermit, CS2020, CSWarren, CX, Cactus.man, Cal 1234, Can't sleep, clown will eat me, Candybars, Captain head, Carlylecastle, Cat2020, CathNek, Chaos, Charles Matthews, Charukesi, Chinju, Chopchopwhitey, Cmarnold, Coffee2theorems, CommodoreMan, Complexica, Conversion script, Convictionist, Coppertwig, Corvus cornix, Crownjewel82, Cumi, Curps, Cwkmail, Cyrloc, DGaw, DSRH, DV8 2XL, DancingMan, DancingPenguin, Dave Bass, Daveyork, David R. Ingham, Dbroadwell, Debresser, Decumanus, Deeptrivia, Defunkt, DerHexer, Devanshi.shah, Dfred, Dgurubaran, Dicklyon, Diderot7, Diego pmc, Dino, Discospinster, Djr32, Dlyons493, Doc Tropics, DougsTech, Drini, Dumelow, Dysprosia, Edward, Enochlau, Enuja, Epbr123, Erkan Yilmaz, Escape Orbit, Evercat, Everyking, Evil Monkey, Evilphoenix, Experiment123, Fairandbalanced, Fenice, Fennec, Filip Larsen, Flammifer, Flewis, Flower Priest, Flux.books, Fluzwup, Fratrep, Freakonaleashnj, Funky Fantom, GabrielPere, Gail, Gandalf61, Gandhi gaurav, Gareth Owen, Garravogue, Gene Ward Smith, Gfalco, Giftlite, Goethean, Gommert, Gorgeorgeus, GraemeL, Greg321, GregorB, Grein, Guardian of Light, Gumpu, Gwernol, GwydionM, Headbomb, HeckXX, Helixblue, Hellfire81, HenkvD, Herakles01, Heron, Hervegirod, Hillman, Hu, Husond, Hut 8.5, IPSOS, IPrussian, IRP, Igny, Iluvcapra, Imaninjapirate, InnocuousPseudonym, Inquisitus, Invisifan, Iridescent, Isomorphic, Izehar, J.delanoy, JAK2112, JJIG, JLBernstein, JaGa, JabberWok, Jacksonroberts25, Jacob501, Jahiegel, Jamesevi, JavierMC, Jayvdb, Jcsellak, Jdthood, Jeff G., Jill.marleigh, Jitse Niesen, Jj137, Jmeppley, Jmlipton, Joakim Munkhammar, JocK, JoergenB, JorgeGG, Jorunn, Jose Ramos, Joshua Andersen, Joyous!, Jugander, Justinodem, K1Bond007, Karada, Karol Langner, Katalaveno, Kelisi, Kevin Baas, KickAir8P~, Kieff, Kilmer-san, Kim Bruning, Kirrages, KoshVorlon, Kotinopoulos, Kristen Eriksen, Ksmadden, Kungfuadam, Kuru, Kızılsungur, La Pianista, Lakinekaki, Lambiam, Lapaz, Lectonar, Letranova, Lexor, Linas, Loren36, Luigi-ish, Lulu of the Lotus-Eaters, Lumidek, MER-C, MKoltnow, MONGO, Madrazz, Magmi, Mako098765, Maksim-e, ManaUser, Marquez, MathMartin, Mav, Maxlittle2007, Mbc362, Mct mht, Mdd, MetsFan76, Michael Hardy, Mikaey, Mike2vil, MikeBaharmast, Milly.mortimer, Mimigary, Mini-Geek, Mjkelley79, Mosgiel, Msh210, Mukadderat, Munford, Nabarry, Nakon, Neurophyre, NewEnglandYankee, Nicholasink, Nivix, Niyse, Nori Llane, Nutter13, Obradovic Goran, Ojigiri, OleMaster, Oliver202, Oneiros, Onesius, Oshron, Oxymoron83, P3d0, PSOfan2000, PaePae, Pascal.Tesson, Paul August, Paul Murray, Perrygogas, PeterSymonds, Phgao, Phil Boswell, Plastikspork, Pleasantville, Pmanderson, Pmg, Poeloq, Prashmail, Prokaryote1234, PurpleRain, Quarl, QuasiAbstract, Radagast3, RainbowOfLight, Rakeshfern, Raven4x4x, Rdsmith4, Readopedia, Red Thunder, Reject, RexNL, Rich Farmbrough, Richardseel, Roadrunner, Romarin, Romon, Rosaak, Rror, Ruud Koot, Rwh, SUPERSONICOOHHHOHOH, Sagaciousuk, Salih, Sam Hocevar, SamuraiGabe, Scartol, SchfiftyThree, Schneelocke, Scid, Scientio, Scriberius, Semper discens, Sethmasters, Shappy, Sheildofthunder, Shiftchange, Shoeofdeath, Simetrical, SimmonsNorwood, SimonP, Sligocki, Smack, Smithbrenon, Smjg, Snoyes, Sodium, Soir, SomeUsr, Somecreepyoldguy, Spectrogram, Spidermanizdabest, Splash, StaticGull, Stellea, Stephenb, Stephenchou0722, Sudharsansn, Sunayana, Suso, Sverdrup, Tagalong99, Tasoskessaris, Tassedethe, Tez, The Equilibrium, The hoodie, TheDestitutionOfOrganizedReligion, TheObtuseAngleOfDoom, TheRingess, Thegreenj, This, that and the other, Thismarty, Thomas G Graf, Timrem, TitusCarus, To Fight a Vandal, Tobias Hoevekamp, Tojuro, Tomasao, Tommi Ronkainen, Torcini, Tosun, Totnesmartin, Traroth, Tunnels of Set, UXs, Un Piton, Universal Hero, Unyoyega, V.V zzzzz, VanFullOfMidgets, Versus22, Vicarious, Vicky sekar, Voorlandt, Vsmith, WAS 4.250, Waitati, Warhammer 8, Wbblaze4, West Brom 4ever, WhiteC, Widenet, Wikiskimmer, Wile E. Heresiarch, William M. Connolley, Winonanick, Wolfrock, XFreakonaLeashX, XaosBits, Yamamoto Ichiro, Yanksox, Yodler, Yoeb137, Ytbau, Zakholdsworth, Zardoze, Zfr, Zojj, Zsniew, Zubras, 1011 anonymous edits

Rössler attractor *Source*: http://en.wikipedia.org/w/index.php?oldid=285829667 *Contributors*: AndrewKepert, Can't sleep, clown will eat me, Carrp, Cquimper, El C, Eteq, Fnielsen, GregorB, Hqb, Joakim Munkhammar, Kwamikagami, Logicalrealist, Oleg Alexandrov, Selket, SnowRaptor, Tó campos, 27 anonymous edits

Standard map *Source*: http://en.wikipedia.org/w/index.php?oldid=239166272 *Contributors*: 3p1416, El C, Pmg, Robertd, Shepelyansky, 6 anonymous edits

Synchronization of chaos *Source*: http://en.wikipedia.org/w/index.php?oldid=283959552 *Contributors*: 3p1416, IRP, Jitse Niesen, PhilKnight, QueenCake, Versus22, 21 anonymous edits

Molecular dynamics *Source*: http://en.wikipedia.org/w/index.php?oldid=295309203 *Contributors*: Agilemolecule, Alex.g, Amire80, Ammatsun, Anthracene, Anxdo, ApJilly, Astavats, Ayucat, Bbullot, Bduke, BenFrantzDale, Bubba hotep, Chris the speller, Cortonin, Cwassman, DMacks, DRider, Dacb, DeadEyeArrow, Demus Wiesbaden, Dicklyon, Dietmar.paschek, DragonflySixtyseven, Drswenson, Ebuchol, Ehdr, Gentgeen, Giftlite, Huckit, Itamblyn, Itub, JWSchmidt, Jerome Charles Potts, Jorgenumata, Jugander, Kaihsu, Karol Langner, Katherine Folsom, Kennylam, Kevyn, Kjaergaard, Knordlun, Laurent1979, Lexor, LiDaobing, Linas, Lomenoldur, Ludx, Maduixa, Marchuta, Marx Gomes, Mateusz Galuszka, Mattopia, Md Arshad Iqbal, Mihoopes, Mr Marie Weaver, Msuzen, Nicolasbock, Oiramrasec, Opabinia regalis, Ossi, P99am, Paul.raymond.brenner, Pedrito, Pelister, PhCOOH, Pksach, PrometheusX303, Raviwiki4, Rob Hooft, Roo1812, Sandycx, Shura58, Smoe, Smremde, Stewartadcock, Sudiarta, TStein, Themfromspace, Thorwald, Utcursch, Van helsing, Whanrott, Wikimcmd, Wittgenstein77, Wiz9999, Xavier andrade, Yrtgm, 200 anonymous edits

Monte Carlo method *Source*: http://en.wikipedia.org/w/index.php?oldid=294459459 *Contributors*: *drew, ABCD, Aardvark92, Adfred123, Aferistas, Agilemolecule, Akanksh, Alanbly, Albmont, AlexBIOSS, AlexandreCam, AlfredR, Alliance09, Altenmann, Andrea Parri, Andreas Kaufmann, Angelbo, Aniu, Apanag, Aspuru, Atlant, Avalcarce, Aznrocket, BAxelrod, BConleyEEPE, Banano03, Banus, Bduke, BenFrantzDale, BenTrotsky, Bender235, Bensaccount, BillGosset, Bkell, Blotwell, Bmaddy, Bobo192, Boffob, Boredzo, Broquaint, Btyner, CRGreathouse, Caiaffa, Charles Matthews, ChicagoActuary, Cibergili, Cm the p, Colonies Chris, Coneslayer, Cretog8, Criter, Cybercobra, Cython1, DMG413, Damistmu, Ddcampayo, Ddxc, Digemedi, Ds53, Duck ears, Duncharris, Dylanwhs, ERosa, EldKatt, Elpincha, Elwikipedista, Eudaemonic3, Ezrakilty, Fastfission, Fintor, Flammifer, Frozen fish, Furrykef, G716, Giftlite, Gilliam, Goudzovski, GraemeL, GrayCalhoun, Greenyoda, Grestrepo, Gtrmp, Gökhan, Hanksname, Hawaiian717, Hokanomono, Hu12, Hubbardaie, ILikeThings, IanOsgood, Inrad, Itub, Jackal irl, Janpedia, JavaManAz, Jeffq, Jitse Niesen, Joey0084, JohnOwens, Jorgenumata, Jsarratt, Jugander, Jérôme, K.lee, KSmrq, KaHa242, Karol Langner, Kenmckinley, Kimys, Knordlun, Kroese, Kummi, Kuru, Lambyte, LeoTrottier, Levin, Lexor, LoveMonkey, Malatesta, Male1979, ManchotPi, Marcofalcioni, Martinp, Masatran, Mathcount, MaxHD, Maxentrope, Maylene, Melcombe, Michael Hardy, Mikael V, Misha Stepanov, Mnath, Moink, Mtford, Nagasaka, Nanshu, Narayanese, Nelson50, Nosophorus, Nsaa, Nuno Tavares, Nvartaniucsd, Ohnoitsjamie, Oli Filth, Oneboy, Orderud, OrgasGirl, P99am, Paul August, PaulxSA, Pbroks13, Pcb21, Pete.Hurd, PeterBoun, Pgreenfinch, Philopp,

Pibwl, Pinguin.tk, PlantTrees, Pne, Popsracer, Poupoune5, Qadro, Quantumelfmage, Quentar, Qxz, RWillwerth, Ramin Nakisa, Redgolpe, Renesis, Richie Rocks, Rinconsoleao, Rjmccall, Ronnotel, Rs2, SKelly1313, Sam Korn, Samratvishaljain, Sergio.ballestrero, Shreevatsa, Snoyes, Somewherepurple, Spellmaster, Splash6, SpuriousQ, Stefanez, Stefanomione, StewartMH, Stimpy, Storm Rider, Superninja, Tarantola, Taxman, Tayste, Tesi1700, Theron110, Thirteenity, Tiger Khan, Tim Starling, Tom harrison, Tooksteps, Trebor, Twooars, Urdutext, Vipuser, VoseSoftware, Wile E. Heresiarch, Yoderj, Zarniwoot, Zoicon5, Zr40, Zuidervled, 307 anonymous edits

Quantum chemistry *Source*: http://en.wikipedia.org/w/index.php?oldid=292520271 *Contributors*: 144.189.40.xxx, 208.40.185.xxx, 4lex, Acroterion, Alansohn, Ayla, BTDenyer, Bci2, Bduke, Bob, BrianY, Bubbha, CDN99, Capecodeph, ChemGardener, CloudNine, Cmdrjameson, CommonsDelinker, Conversion script, Cool3, Cypa, EdJohnston, Edsanville, EmilyT, Euryalus, Gentgeen, Gershom, Giftlite, Glenn, GregorB, Haljolad, HappyCamper, Holdran, Hugo-cs, Ian Pitchford, Itub, James 007, Jantop, JerrySteal, Kaliumfredrik, Karol Langner, Keenan Pepper, Keilana, Koinut, Krash, La goutte de pluie, Lampuchi, Ligulem, Lijuni, Looxix, M stone, Martin Hedegaard, Meisfunny, Milo, Nickptar, Noisy, Nzzl, Okedem, Perelaar, Ratol, Rifleman 82, SHL-at-Sv, SQL, Sadi Carnot, Salsb, Shalom Yechiel, Shanel, Sidhekin, Smoe, Sunev, Tasudrty, Terhorstj, Timwi, UninvitedCompany, Vb, Vgy7ujm, Vig vimarsh, Voigfdsa, Vsmith, W.F.Galway, Wiki alf, Yurik, Zarniwoot, Zeimusu, Александър, مل ع ب وب ح م, 132 anonymous edits

Quantum Monte Carlo *Source*: http://en.wikipedia.org/w/index.php?oldid=295037474 *Contributors*: Acipsen, Amyoungil, Bci2, Conscious, Henry Delforn, Isilanes, Jheriko, Karol Langner, Lucaskw, Mdt26a, Melcombe, Michael Hardy, NNemec, Pablomme, Paulcardan, Rbonvall, Rich Farmbrough, Rjwilmsi, Sagaciousuk, Supersion77, TestPilot, Trigger hippie77, UkPaolo, Veinor, Vgy7ujm, Vyznev Xnebara, WilliamDParker, Wirawan0, 38 anonymous edits

DNA Dynamics *Source*: http://en.wikipedia.org/w/index.php?oldid=295133476 *Contributors*: Auntof6, Bci2, Chris the speller, CommonsDelinker, Ironholds, Potatoswatter

Molecular models of DNA *Source*: http://en.wikipedia.org/w/index.php?oldid=293388965 *Contributors*: Bci2, Chris the speller, CommonsDelinker, Oscarthecat

Protein folding *Source*: http://en.wikipedia.org/w/index.php?oldid=293585634 *Contributors*: 168..., 5beta5, Adriferr, Agilemolecule, Akane700, Andraaide, Arcadian, Banus, Barticus88, Bci2, Bendzh, Bfinn, Bioinfo177, Biophys, Biophysik, Blainster, Blooooo, Brianga, Bryan Derksen, Cacycle, Calvero JP, Cathalgarvey, Cburnett, ChicXulub, Clicketyclack, Computor, Cyberman, Czhangrice, D. Recorder, DannyWilde, Davepntr, DennisDaniels, Dhatz, Donarreiskoffer, Dwmyers, ESkog, Eequor, Erencexor, Erwinser, Fawzin, Fuzheado, Gcrossan, Gowantervo, Grimlock, Harley peters, Herd of Swine, Hooplehead, Intangir, Ixfd64, JJ TKOB, Jacobsman, Jammedshut, JeramieHicks, Katherine Folsom, Kevyn, Kierano, Kjaergaard, Konstantin, Kostmo, Leptictidium, Lexor, Lfh, LiDaobing, Lir, Lostart, Lucaaah, M stone, Macintosh10000, Madeleine Price Ball, Magnus Manske, Malcolm Farmer, Mark Renier, Michael Hardy, Miguel Andrade, Minghong, MockAE, Movado73, Movalley, Myscrnnm, Netesq, Opabinia regalis, Otvaltak, P99am, Piotrus, Ptr123, RDBrown, Rebroad, Richyfp, Rjwilmsi, RoyBoy, RunninRiot, Samrat Mukhopadhyay, Sharkman, Shmedia, SirHaddock, Smelialichu, Snowmanradio, Splette, SpuriousQ, Stepa, Sunny17152, Taw, Terrace4, TestPilot, The Anome, The Great Cook, Tijmz, TimVickers, Tomixdf, Tommstein, Toytoy, Trenchcoatjedi, Trusilver, V8rik, Waerloeg, WillowW, WorldDownInFire, Xnuala, Xy7, Yves-henri, Zoicon5, רואיל, 146 anonymous edits

Image Sources, Licenses and Contributors

Image:System boundary.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:System_boundary.svg *License*: unknown *Contributors*: User:Stannered

Image:Lorenz attractor yb.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Lorenz_attractor_yb.svg *License*: Creative Commons Attribution-Sharealike 2.5 *Contributors*: User:Dschwen, User:Wikimol

Image:LinearFields.png *Source*: http://en.wikipedia.org/w/index.php?title=File:LinearFields.png *License*: Creative Commons Attribution 2.5 *Contributors*: XaosBits

Image:complexity-map-overview.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Complexity-map-overview.png *License*: unknown *Contributors*: Bcastel3

Image:sociology-complexity-legend.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Sociology-complexity-legend.png *License*: unknown *Contributors*: Bcastel3

Image:Braitenberg.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Braitenberg.jpg *License*: GNU Free Documentation License *Contributors*: Rxke

Image:Complex-adaptive-system.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Complex-adaptive-system.jpg *License*: GNU Free Documentation License *Contributors*: User:Acadac

Image:Genomics GTL Pictorial Program.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Genomics_GTL_Pictorial_Program.jpg *License*: Public Domain *Contributors*: Mdd

Image:Signal transduction v1.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Signal_transduction_v1.png *License*: GNU Free Documentation License *Contributors*: Original uploader was Roadnottaken at en.wikipedia

Image:Neural network example.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Neural_network_example.png *License*: Public Domain *Contributors*: Murx, Wiso, 1 anonymous edits

Image:Gene Regulatory Network.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Gene_Regulatory_Network.jpg *License*: Public Domain *Contributors*: Mdd

Image:Gene Regulatory Network 2.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Gene_Regulatory_Network_2.jpg *License*: Public Domain *Contributors*: Mdd

Image:A thaliana metabolic network.png *Source*: http://en.wikipedia.org/w/index.php?title=File:A_thaliana_metabolic_network.png *License*: GNU Free Documentation License *Contributors*: Original uploader was TimVickers at en.wikipedia

Image:Metabolic Network Model for Escherichia coli.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Metabolic_Network_Model_for_Escherichia_coli.jpg *License*: Public Domain *Contributors*: Mdd

Image:protein pattern analyzer.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Protein_pattern_analyzer.jpg *License*: Public Domain *Contributors*: Athaenara, Emesee, 1 anonymous edits

Image:Cell cycle bifurcation diagram.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Cell_cycle_bifurcation_diagram.jpg *License*: unknown *Contributors*: User:Squidonius

Image:Saddlenode.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:Saddlenode.gif *License*: unknown *Contributors*: Deeptrivia, 1 anonymous edits

Image:Chaosorderchaos.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Chaosorderchaos.png *License*: unknown *Contributors*: Darapti, Deeptrivia

Image:fold_bifurcation.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Fold_bifurcation.svg *License*: Public Domain *Contributors*: Darapti, Jheald

Image:cusp_catastrophe.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Cusp_catastrophe.svg *License*: Public Domain *Contributors*: Jheald

Image:cusp_shape.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Cusp_shape.svg *License*: Public Domain *Contributors*: Jheald

Image:pitchfork_bifurcation_left.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Pitchfork_bifurcation_left.svg *License*: Public Domain *Contributors*: Darapti, Jheald

Image:GustaveMoreau03.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:GustaveMoreau03.jpg *License*: unknown *Contributors*: AndreasPraefcke, Bibi Saint-Pol, Mattes, MobyDick, Olivier2, Wst, Xenophon, 3 anonymous edits

Image:LogisticMap BifurcationDiagram.png *Source*: http://en.wikipedia.org/w/index.php?title=File:LogisticMap_BifurcationDiagram.png *License*: Public Domain *Contributors*: Adam majewski, CiaPan, Darapti, Nagy, Oleg Alexandrov, PAR, 2 anonymous edits

Image:Fractal fern1.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Fractal_fern1.png *License*: Public Domain *Contributors*: Adam Mihályi

Image:Airplane vortex edit.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Airplane_vortex_edit.jpg *License*: Public Domain *Contributors*: User:Fir0002

Image:Club Lambda.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Club_Lambda.png *License*: unknown *Contributors*: User:Brewhaha@edmc.net

Image:Damped driven chaotic pendulum - double period behavior.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Damped_driven_chaotic_pendulum_-_double_period_behavior.png *License*: Public Domain *Contributors*: Brian0918

Image:Roessler attractor.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Roessler_attractor.png *License*: Creative Commons Attribution-Sharealike 2.5 *Contributors*: User:Wofl

Image:RosslerStereo.png *Source*: http://en.wikipedia.org/w/index.php?title=File:RosslerStereo.png *License*: GNU Free Documentation License *Contributors*: Benc17388, Eteq, 1 anonymous edits

Image:RosslerstdXY.png *Source*: http://en.wikipedia.org/w/index.php?title=File:RosslerstdXY.png *License*: GNU Free Documentation License *Contributors*: ChrisRuvolo, Kieff, Logicalrealist

Image:Eigenvectors.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Eigenvectors.png *License*: GNU Free Documentation License *Contributors*: User:Iain

Image:Rosslerstd3D.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Rosslerstd3D.png *License*: GNU Free Documentation License *Contributors*: Logicalrealist

Image:Poincare.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Poincare.png *License*: GNU Free Documentation License *Contributors*: Logicalrealist

Image:Tentmap.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Tentmap.png *License*: GNU Free Documentation License *Contributors*: Logicalrealist

Image:Bifurcation DiagramB.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Bifurcation_DiagramB.png *License*: unknown *Contributors*: SnowRaptor

Image:Bifurcation.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Bifurcation.png *License*: GNU Free Documentation License *Contributors*: Logicalrealist

Image:VaryingC.png *Source*: http://en.wikipedia.org/w/index.php?title=File:VaryingC.png *License*: GNU Free Documentation License *Contributors*: ChrisRuvolo, Kieff, Logicalrealist

Image:standard-map.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Standard-map.png *License*: unknown *Contributors*: Darapti, Pmg

Image:Mdalgorithm.PNG *Source*: http://en.wikipedia.org/w/index.php?title=File:Mdalgorithm.PNG *License*: Public Domain *Contributors*: User:Knordlun

Image:Monte carlo method.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Monte_carlo_method.svg *License*: Public Domain *Contributors*: --pbroks13talk? Original uploader was Pbroks13 at en.wikipedia

File:DNA-fragment-3D-vdW.png *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA-fragment-3D-vdW.png *License*: Public Domain *Contributors*: Benjah-bmm27

File:ADN animation.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:ADN_animation.gif *License*: Public Domain *Contributors*: Aushulz, Bawolff, Brian0918, Kersti Nebelsiek, Magadan, Mattes, Origamiemensch, Stevenfruitsmaak, 3 anonymous edits

File:Simple harmonic oscillator.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:Simple_harmonic_oscillator.gif *License*: Public Domain *Contributors*: User:Oleg Alexandrov

File:A-DNA orbit animated small.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:A-DNA_orbit_animated_small.gif *License*: GNU Free Documentation License *Contributors*: User:Bstlee, User:Zephyris

File:3D model hydrogen bonds in water.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:3D_model_hydrogen_bonds_in_water.jpg *License*: GNU Free Documentation License *Contributors*: User:snek01

File:DNA labels.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_labels.jpg *License*: GNU Free Documentation License *Contributors*: User:Raul654

File:A-B-Z-DNA Side View.png *Source*: http://en.wikipedia.org/w/index.php?title=File:A-B-Z-DNA_Side_View.png *License*: Public Domain *Contributors*: Original uploader was Thorwald at en.wikipedia

File:ABDNAxrgpj.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:ABDNAxrgpj.jpg *License*: GNU Free Documentation License *Contributors*: I.C. Baianu et al.

File:BU Bio5.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:BU_Bio5.jpg *License*: Creative Commons Attribution-Sharealike 2.0 *Contributors*: Original uploader was Elapied at fr.wikipedia

File:Parallel telomere quadruple.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Parallel_telomere_quadruple.png *License*: unknown *Contributors*: User:Splette

File:Four-way DNA junction.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:Four-way_DNA_junction.gif *License*: Public Domain *Contributors*: Aushulz, Molatwork, Origamiemensch, TimVickers, 1 anonymous edits

File:Sarfus.DNABiochip.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Sarfus.DNABiochip.jpg *License*: unknown *Contributors*: Nanolane

File:DNA nanostructures.png *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_nanostructures.png *License*: unknown *Contributors*: (Images were kindly provided by Thomas H. LaBean and Hao Yan.)

File:Circular DNA Supercoiling.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Circular_DNA_Supercoiling.png *License*: GNU Free Documentation License *Contributors*: Richard Wheeler (Zephyris)

File:Rothemund-DNA-SierpinskiGasket.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Rothemund-DNA-SierpinskiGasket.jpg *License*: Creative Commons Attribution 2.5 *Contributors*: Antony-22

File:RNA pol.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:RNA_pol.jpg *License*: Public Domain *Contributors*: InfoCan

File:Primase 3B39.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Primase_3B39.png *License*: Public Domain *Contributors*: own work

File:MGMT+DNA 1T38.png *Source*: http://en.wikipedia.org/w/index.php?title=File:MGMT+DNA_1T38.png *License*: Public Domain *Contributors*: own work

File:DNA damaged by carcinogenic 2-aminofluorene AF .jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_damaged_by_carcinogenic_2-aminofluorene_AF_.jpg *License*: Public Domain *Contributors*: Brian E. Hingerty, Oak Ridge National Laboratory Suse Broyde, New York University Dinshaw J. Patel, Memorial Sloan Kettering Cancer Center

File:Chr2 orang human.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Chr2_orang_human.jpg *License*: Creative Commons Attribution-Sharealike 2.5 *Contributors*: Verena Schubel, Stefan Müller, Department Biologie der Ludwig-Maximilians-Universität München.

File:Plasmid emNL.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Plasmid_emNL.jpg *License*: GNU Free Documentation License *Contributors*: Denniss, Glenn, Rasbak

File:3D-SIM-3 Prophase 3 color.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:3D-SIM-3_Prophase_3_color.jpg *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Lothar Schermelleh

File:Chromosome2 merge.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Chromosome2_merge.png *License*: Public Domain *Contributors*: Original uploader was Evercat at en.wikipedia

File:Chromosome Conformation Capture Technology.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Chromosome_Conformation_Capture_Technology.jpg *License*: Public Domain *Contributors*: User:Kangyun1985

File:PCR.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:PCR.svg *License*: unknown *Contributors*: User:Madprime

File:Pcr gel.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Pcr_gel.png *License*: GNU Free Documentation License *Contributors*: Habj, Ies, PatríciaR, Retama, Saperaud

File:SFP discovery principle.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:SFP_discovery_principle.jpg *License*: unknown *Contributors*: User:Agbiotec

File:Cdnaarray.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Cdnaarray.jpg *License*: unknown *Contributors*: Mangapoco

File:Expression of Human Wild-Type and P239S Mutant Palladin.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Expression_of_Human_Wild-Type_and_P239S_Mutant_Palladin.png *License*: unknown *Contributors*: see above

File:Plos VHL.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Plos_VHL.jpg *License*: Creative Commons Attribution 2.5 *Contributors*: Akinom, Anniolek, Filip em, Thommiddleton

File:Random genetic drift chart.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Random_genetic_drift_chart.png *License*: unknown *Contributors*: User:Professor marginalia

File:Signal transduction v1.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Signal_transduction_v1.png *License*: GNU Free Documentation License *Contributors*: Original uploader was Roadnottaken at en.wikipedia

File:Genomics GTL Pictorial Program.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Genomics_GTL_Pictorial_Program.jpg *License*: Public Domain *Contributors*: Mdd

File:Maldi informatics figure 6.JPG *Source*: http://en.wikipedia.org/w/index.php?title=File:Maldi_informatics_figure_6.JPG *License*: Public Domain *Contributors*: Rbeavis

File:Stokes shift.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Stokes_shift.png *License*: unknown *Contributors*: User:Mykhal

File:CARS Scheme.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:CARS_Scheme.svg *License*: unknown *Contributors*: Onno Gabriel

File:HyperspectralCube.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:HyperspectralCube.jpg *License*: Public Domain *Contributors*: Dr. Nicholas M. Short, Sr.

File:MultispectralComparedToHyperspectral.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:MultispectralComparedToHyperspectral.jpg *License*: Public Domain *Contributors*: Dr. Nicholas M. Short, Sr.

File:Tirfm.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Tirfm.svg *License*: Public Domain *Contributors*: Dawid Kulik

File:Fluorescence microscop.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Fluorescence_microscop.jpg *License*: unknown *Contributors*: Masur

File:Microscope And Digital Camera.JPG *Source*: http://en.wikipedia.org/w/index.php?title=File:Microscope_And_Digital_Camera.JPG *License*: GNU Free Documentation License *Contributors*: User:Zephyris

File:FluorescentCells.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:FluorescentCells.jpg *License*: Public Domain *Contributors*: DO11.10, Emijrp, NEON ja, Origamiemensch, Splette, Tolanor, 5 anonymous edits

File:Yeast membrane proteins.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Yeast_membrane_proteins.jpg *License*: unknown *Contributors*: User:Masur

File:S cerevisiae septins.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:S_cerevisiae_septins.jpg *License*: Public Domain *Contributors*: Spitfire ch, Philippsen Lab, Biozentrum Basel

File:Dividing Cell Fluorescence.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Dividing_Cell_Fluorescence.jpg *License*: unknown *Contributors*: Will-moore-dundee

File:HeLa Hoechst 33258.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:HeLa_Hoechst_33258.jpg *License*: Public Domain *Contributors*: TenOfAllTrades

File:FISH 13 21.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:FISH_13_21.jpg *License*: Public Domain *Contributors*: Gregor1976

File:Bloodcell sun flares pathology.jpeg *Source*: http://en.wikipedia.org/w/index.php?title=File:Bloodcell_sun_flares_pathology.jpeg *License*: Public Domain *Contributors*: Birindand, Karelj, NEON ja, 1 anonymous edits

File:Carboxysome 3 images.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Carboxysome_3_images.png *License*: Creative Commons Attribution 3.0 *Contributors*: Prof. Todd O. Yeates, UCLA Dept. of Chem. and Biochem.

File:DNA_nanostructures.png *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_nanostructures.png *License*: unknown *Contributors*: (Images were kindly provided by Thomas H. LaBean and Hao Yan.)

File:Holliday junction coloured.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Holliday_junction_coloured.png *License*: GNU Free Documentation License *Contributors*: Original uploader was Zephyris at en.wikipedia

File:Holliday Junction cropped.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Holliday_Junction_cropped.png *License*: GNU Free Documentation License *Contributors*: Original uploader was TimVickers at en.wikipedia

File:Atomic force microscope by Zureks.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Atomic_force_microscope_by_Zureks.jpg *License*: unknown *Contributors*: User:Zureks

File:Atomic force microscope block diagram.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Atomic_force_microscope_block_diagram.png *License*: Public Domain *Contributors*: Original uploader was Askewmind at en.wikipedia

File:AFM view of sodium chloride.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:AFM_view_of_sodium_chloride.gif *License*: Public Domain *Contributors*: Courtesy of prof. Ernst Meyer, university of Basel

File:Single-Molecule-Under-Water-AFM-Tapping-Mode.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Single-Molecule-Under-Water-AFM-Tapping-Mode.jpg *License*: unknown *Contributors*: User:Yurko

Image:Sarfus.DNABiochip.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Sarfus.DNABiochip.jpg *License*: unknown *Contributors*: Nanolane

Image:Spinning DNA.gif *Source*: http://en.wikipedia.org/w/index.php?title=File:Spinning_DNA.gif *License*: Public Domain *Contributors*: USDA

File:Methanol.pdb.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Methanol.pdb.png *License*: Creative Commons Attribution-Sharealike 2.5 *Contributors*: ALoopingIcon, Benjah-bmm27

File:DNA chemical structure.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_chemical_structure.svg *License*: unknown *Contributors*: Madprime, Wickey, 1 anonymous edits

File:DNA replication.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_replication.svg *License*: unknown *Contributors*: Bibi Saint-Pol, Elborgo, En rouge, Helix84, Jnpet, LadyofHats, Leafnode, M.Komorniczak, MetalGearLiquid, Nikola Smolenski, Nl74, Sentausa, WarX, 2 anonymous edits

Image:Methanol.pdb.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Methanol.pdb.png *License*: Creative Commons Attribution-Sharealike 2.5 *Contributors*: ALoopingIcon, Benjah-bmm27

File:Bragg diffraction.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Bragg_diffraction.png *License*: GNU General Public License *Contributors*: user:hadmack

File:DNA in water.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_in_water.jpg *License*: unknown *Contributors*: User:Bbkkk

File:X ray diffraction.png *Source*: http://en.wikipedia.org/w/index.php?title=File:X_ray_diffraction.png *License*: unknown *Contributors*: Thomas Splettstoesser

File:X Ray Diffractometer.JPG *Source*: http://en.wikipedia.org/w/index.php?title=File:X_Ray_Diffractometer.JPG *License*: GNU Free Documentation License *Contributors*: Ff02::3, Pieter Kuiper

File:SLAC detector edit1.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:SLAC_detector_edit1.jpg *License*: unknown *Contributors*: User:Mfield, User:Starwiz

File:ISIS exptal hall.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:ISIS_exptal_hall.jpg *License*: unknown *Contributors*: wurzeller

File:Dna-SNP.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Dna-SNP.svg *License*: unknown *Contributors*: User:Gringer

File:DNA Under electron microscope Image 3576B-PH.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_Under_electron_microscope_Image_3576B-PH.jpg *License*: unknown *Contributors*: Original uploader was SeanMack at en.wikipedia

File:DNA Model Crick-Watson.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_Model_Crick-Watson.jpg *License*: Public Domain *Contributors*: User:Alkivar

File:AT DNA base pair pt.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:AT_DNA_base_pair_pt.svg *License*: Public Domain *Contributors*: User:Lijealso

File:Museo Príncipe Felipe. ADN.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Museo_Príncipe_Felipe._ADN.jpg *License*: Creative Commons Attribution-Sharealike 2.0 *Contributors*: Fernando

File:AGCT DNA mini.png *Source*: http://en.wikipedia.org/w/index.php?title=File:AGCT_DNA_mini.png *License*: unknown *Contributors*: Iquo

File:Rosalindfranklinsjokecard.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Rosalindfranklinsjokecard.jpg *License*: unknown *Contributors*: Bci2, J.delanoy, Martyman, Nitramrekcap, Rjm at sleepers, 3 anonymous edits

Image:Rosalindfranklinsjokecard.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Rosalindfranklinsjokecard.jpg *License*: unknown *Contributors*: Bci2, J.delanoy, Martyman, Nitramrekcap, Rjm at sleepers, 3 anonymous edits

File:DNA Repair.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:DNA_Repair.jpg *License*: Public Domain *Contributors*: Courtesy of Tom Ellenberger, Washington University School of Medicine in St. Louis.

File:Chromatin chromosom.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Chromatin_chromosom.png *License*: Public Domain *Contributors*: User:Magnus Manske

File:Chromosome.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Chromosome.svg *License*: unknown *Contributors*: User:Dietzel65, User:Magnus Manske, User:Tryphon

File:Transkription Translation 01.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Transkription_Translation_01.jpg *License*: Public Domain *Contributors*: User:Kuebi

File:RibosomaleTranskriptionsEinheit.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:RibosomaleTranskriptionsEinheit.jpg *License*: GNU Free Documentation License *Contributors*: User:Merops

File:Mitochondrial DNA and diseases.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Mitochondrial_DNA_and_diseases.png *License*: unknown *Contributors*: User:XXXL1986

File:Co-dominance Rhododendron.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Co-dominance_Rhododendron.jpg *License*: Creative Commons Attribution 2.0 *Contributors*: Ayacop, Cillas, FlickrLickr, FlickreviewR, Horcha, Kanonkas, Kevmin, MPF, Para

File:AFMimageRoughGlass20x20.png *Source*: http://en.wikipedia.org/w/index.php?title=File:AFMimageRoughGlass20x20.png *License*: Public Domain *Contributors*: Chych

File:Confocalprinciple.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:Confocalprinciple.svg *License*: GNU Free Documentation License *Contributors*: Danh

File:3D-SIM-1 NPC Confocal vs 3D-SIM detail.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:3D-SIM-1_NPC_Confocal_vs_3D-SIM_detail.jpg *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Changes in layout by the uploader. Only the creator of the original (Lothar Schermelleh) should be credited.

File:Inverted microscope.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Inverted_microscope.jpg *License*: unknown *Contributors*: Nuno Nogueira (Nmnogueira) Original uploader was Nmnogueira at en.wikipedia

File:FluorescenceFilters 2008-09-28.svg *Source*: http://en.wikipedia.org/w/index.php?title=File:FluorescenceFilters_2008-09-28.svg *License*: unknown *Contributors*: User:Mastermolch

Image:Protein folding.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Protein_folding.png *License*: Public Domain *Contributors*: DrKjaergaard, PatríciaR

Image:Protein folding schematic.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Protein_folding_schematic.png *License*: Public Domain *Contributors*: User:Tomixdf

License

Version 1.2, November 2002

0.PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.
This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.
We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1.APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.
A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.
A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.
The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.
The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.
A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".
Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.
The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.
A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.
The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2.VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.
You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3.COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.
If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.
If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.
It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4.MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

1. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
2. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
3. State on the Title page the name of the publisher of the Modified Version, as the publisher.
4. Preserve all the copyright notices of the Document.
5. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
6. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
7. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

8. Include an unaltered copy of this License.
9. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
10. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
11. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
12. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
13. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
14. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
15. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5.COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6.COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7.AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8.TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9.TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10.FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright (c) YEAR YOUR NAME.
>
> Permission is granted to copy, distribute and/or modify this document
>
> under the terms of the GNU Free Documentation License, Version 1.2
>
> or any later version published by the Free Software Foundation;
>
> with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
>
> A copy of the license is included in the section entitled "GNU
>
> Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with...Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the
>
> Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.