Bulletin of Mathematical Biology (2007) 69: 635–657
DOI 10.1007/s11538-006-9141-5

ORIGINAL ARTICLE

Information Theory in Living Systems, Methods, Applications, and Challenges

Robert A. Gatenby[a,b,*], B. Roy Frieden[c]

[a] *Department of Radiology, University Medical Center, University of Arizona, 1501 North Campbell Avenue, Tucson, Arizona 85724, USA*
[b] *Department of Applied Mathematics, University of Arizona, 1501 North Campbell Avenue, Tucson, Arizona 85724, USA*
[c] *Department of Optical Science, University of Arizona, Tucson, Arizona 85724, USA*

Received: 16 February 2006 / Accepted: 22 May 2006 / Published online: 3 November 2006
© Society for Mathematical Biology 2006

Abstract Living systems are distinguished in nature by their ability to maintain stable, ordered states far from equilibrium. This is despite constant buffeting by thermodynamic forces that, if unopposed, will inevitably increase disorder. Cells maintain a steep transmembrane entropy gradient by continuous application of information that permits cellular components to carry out highly specific tasks that import energy and export entropy. Thus, the study of information storage, flow and utilization is critical for understanding first principles that govern the dynamics of life. Initial biological applications of information theory (IT) used Shannon's methods to measure the information content in strings of monomers such as genes, RNA, and proteins. Recent work has used bioinformatic and dynamical systems to provide remarkable insights into the topology and dynamics of intracellular information networks. Novel applications of Fisher-, Shannon-, and Kullback–Leibler informations are promoting increased understanding of the mechanisms by which genetic information is converted to work and order. Insights into evolution may be gained by analysis of the the fitness contributions from specific segments of genetic information as well as the optimization process in which the fitness are constrained by the substrate cost for its storage and utilization. Recent IT applications have recognized the possible role of nontraditional information storage structures including lipids and ion gradients as well as information transmission by molecular flux across cell membranes. Many fascinating challenges remain, including defining the intercellular information dynamics of multicellular organisms and the role of disordered information storage and flow in disease.

Keywords Information theory · Entropy · Shannon information

*Correspondance Author.
E-mail address: rgatenby@radiology.arizona.edu (R. A. Gatenby).

1. Introduction

Living organisms are both stable highly-ordered structures, and complex dynamical, semiopen systems functioning far from thermodynamic equilibrium (Morowitz, 1955). The task of maintaining this low entropy state is formidable. Disorder tends to continuously increase as large molecules are broken down, ions and small molecules diffuse along concentration gradients, and metabolism of glucose and other substrate produces metabolites and heat. Thus, a stable highly-ordered intracellular state requires a steep transmembrane entropy gradient maintained through appropriate energy- and information-dependent exchange mechanisms. The ability of cells to import energy and export entropy requires, among other things, accurate identification of atomic and molecular structures so that carbon chains can be imported and efficiently metabolized followed by export of the metabolic products which would otherwise increase intracellular entropy. This can be accomplished only by continuous application of information through multiple cellular components such as membrane transporters. Although information storage, transmission, reception, and utilization clearly play central roles in maintaining a living state, the dynamics governing these processes remain incompletely understood.

In a 1970 review (Johnson, 1970), Johnson characterized information theory (IT) as a "general calculus for biology." It is clear that life without matter and energy is impossible. Johnson's manuscript emphasized that *life without information is likewise impossible*. Since the article, remarkable progress has been made toward understanding the informational basis for life. Automated techniques now allow complete characterization of the RNA and protein contents of cellular populations and, even, individual cells. The human genome has been catalogued along with those of increasing numbers of other organisms. Correlations between specific gene mutations and various human diseases are reported almost daily.

In parallel with these important advances, IT has grown beyond the methods and concepts that dominated investigation of information dynamics in living systems in 1970. Some of these approaches use new applications of traditional mathematical models developed by Fisher (1959), Shannon (1948), and Kullback (1959). Others address limitations of the IT methodology by applying new statistical and modeling approaches to information dynamics, including bioinformatics, dynamical systems, game theory, graph theory, and measurement theory. This research has provided insights, through statistical inference and fundamental analysis, into biological processes over many orders of magnitude.

Clearly information storage, transmission, and utilization is unique to and necessary for living systems. A comprehensive survey of information dynamics in biology is not feasible (or probably even possible). Instead, our focus in this communication will be upon new, and sometimes controversial, IT developments that have provided novel insights into the general principles that govern information storage, transmission, reception, and processing. Results of these approaches will be discussed, along with future goals and challenges.

As a matter of nomenclature, we use the acronym IT to stand both for information theories in general and for those informations that are special cases

of a general information measure called Kullback–Leibler information or cross-entropy. The context will define the particular sense in use.

2. Ideal requirements of biological information

While it seems intuitively apparent that information storage and processing is a fundamental characteristic of living systems, the first principles governing information dynamics in biology remain unclear. The problem is exacerbated by uncertainty in the choice of an appropriate and unique measure for the information. Even settling upon a definition for information in biological systems is surprisingly difficult and often contentious. For example, information can be variably expressed in terms of its relationship to a variety of other system traits including (Shannon, 1948; Trincher, 1965; Pierce, 1980; Gilbert, 1966; Szilard, 1929)

- randomness,
- uncertainty,
- predictability, and
- thermodynamic entropy.

In the context of biology, information is probably best viewed as a *quantity* that is deeply imbedded in a dynamical *process*, so that the information content of a message cannot be separated from the underlying system, which includes the

- meaning or instructions encoded at the source,
- ability of a receiver to use that information to perform work or some other function, and
- energy cost of transcription, transmission and utilization of the information.

Thus, biological information is distinguished from physical or thermodynamic entropy because it requires both order *and* meaning, and can be examined only within the context of the complete system. For example, a protein synthesized by *random* attachment of amino acids is an ordered structure that decreases local entropy but carries no information, just as a word consisting of randomly placed letters is meaningless. Similarly, the information encoded in a protein can be used to perform work/function only by the specific ligand to which it can bind. Thus, the information encoded in a growth factor is conveyed only if a growth factor receptor is present, just as a message in English carries no information to a receiver who speaks only Russian. This critical role of context and meaning, neither of which can be readily quantified, in information is a significant part of the problem in trying to develop a comprehensive mathematical formalism to describe information dynamics in living systems.

Although there is currently no information metric that measures all these traits, forms of Shannon and Fisher Information seem to come closest in accomplishing them. These information measures are respectively used in system models that regard a biological organism as a communication channel or a classical estimation channel.

Information theory (IT) is typically defined as the study of information storage, communication and utilization. Its origin is usually traced to Shannon's pioneering

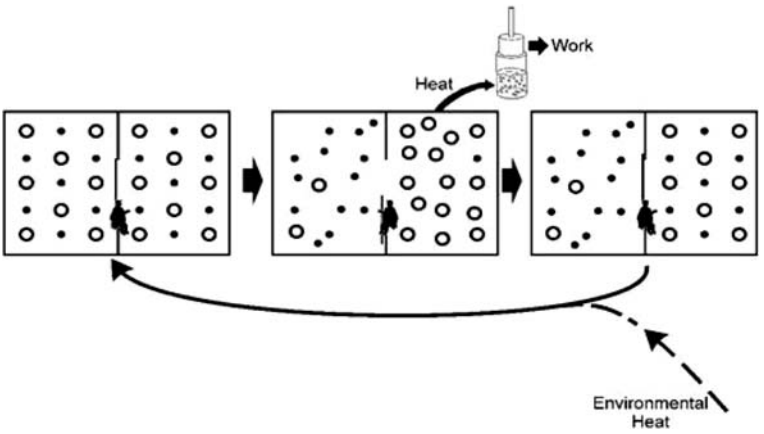

Fig. 1 A modified version of Maxwell's famous Demon *gedanken* experiment. *Open* and *closed circles* represent fast and slow gas molecules, respectively. In the initial state, they pass freely through an open channel so that the particles are distributed equally in both containers which are, therefore, at equal temperature. Here we will assume that the same distribution is present in the environment around the box. Subsequently a demon opens and closes a gate across the channel. It allows only *open circle* to move left to right and *closed circles* right to left producing a new state in which one box contains mostly *open circles* and the other mostly closed ones. The increased temperature in the left box can then be used to do work as shown. Subsequently, the internal state of the boxes is at a lower energy (temperature) than the surroundings. Heat will then flow into the box with a return to the initial state. Note that if the work is exported from the system, it loses energy with each cycle. If the work is performed in the environment in contact with the boxes (i.e., within the system), under ideal conditions, the process would be adiabatic. In reality, of course, energy would be inevitably lost in the process of performing work and metabolism of the demon in making measurements and moving the gate. Note that a hidden cost of the *gedanken* experiment is the energy the demon expended to gain the information necessary to perform its molecule sorting function. This simple experiment illustrates the dynamical and complex interactions of biological information and thermodynamics. The change in molecular distribution due to the demon's efforts is the result of its use of information to distinguish between the two types of molecules and judge the direction of their motion. This utilization of information is manifested thermodynamically by heat flow between the boxes which reduces thermodynamic entropy. The nonrandom distribution of the molecules, by Eq. (8), now confers a nonzero value of H in the containers. This information/energy is then dissipated by performing work returning the containers to their original state.

article (Shannon, 1948) which describes a theory of communication. Interestingly, Shannon generally eschewed the term "information theory," possibly because the term "information" was already in use from the early 1920's onward in describing Fisher's information (Fisher, 1925). In fact the role of information in physical processes, particularly thermodynamics, had been investigated by others (Szilard, 1929) since the era of Maxwell's famous 19th century (Maxwell, 1880) *gedanken* experiment demon (Fig. 1).

2.1. Information metrics

Through the years, many forms of information have been invented and utilized (Ebeling and Frommel, 1998). Aside from those due to *Fisher* and *Shannon*, a

partial list includes informations described by *Renyi*, *Wootters*, *Hellinger*, *Gini–Simpson*, *Tsallis and Kullback–Leibler*. In general, each type of information is tailored to satisfy the needs of a particular area of application. For example, the Gini–Simpson information is sensitive to the degree of diversity that is present in the population mix of a given ecology (Frieden, 2004). Also, certain of these informations are so general as to include some of the others within their scope.

2.1.1. Kullback–Leibler information

Chief among these is the Kullback–Leibler (or K–L) information, denoted as $K_{\mathrm{KL}}(p/r)$ and defined as (Reza, 1961; Prigogine, 1965)

$$K_{\mathrm{KL}}(p/r) \equiv -\sum_i p(y_i) \log_b \frac{p(y_i)}{r(y_i)}. \tag{1}$$

All sums over i are from $i = 1$ to $i = N$. The logarithmic base b is at the discretion of the user. The choice imparts an artificial unit to K_{KL}. For $b = 2$, the unit is the "bit." Although this term was not actually coined by Shannon, it came into popular use through his work. The curve $p(y_i)$ is the subject probability law (usually regarded as unknown and to be found) for events $y_i, i = 1, \ldots, N$, and $r(y_i)$ is a known or input probability law. It is often called a *reference* probability law (Reza, 1961), as discussed below.

Equation (1) is inclusive of Fisher information and both Shannon transinformation and Shannon information (shown below).

The form of Eq. (1) is that of a "distance measure," specifically, the distance between the two curves $p(y_i)$ and $r(y_i)$. As a check, note that if $p(y_i) = r(y_i)$ at all y_i, i.e., the two curves are the *same* curve, any legitimate measure of the distance between them should be zero. In fact the K-L distance, or information, *is* identically zero under these circumstances ($\log_b 1 = 0$). Conversely, the distance in Eq. (1) grows in magnitude as $p(y_i)$ and $r(y_i)$ increasingly differ from one another. As a fine point, the distance is, however, a "directed" one, in that the distance from p to r is not equal to that from r to p, i.e., from Eq. (1) $K_{\mathrm{KL}}(p/r) \neq K_{\mathrm{KL}}(r/p)$.

Another interesting case is when r is "a differential away" from p, i.e.,

$$r(y_i) = p(y_i + \Delta y), \quad \Delta y \to 0. \tag{2}$$

Then if also the subdivision $y_i \equiv i\,\Delta y$, the discrete sum in Eq. (1) becomes a continuous integral over y obeying

$$K_{\mathrm{KL}}(p/r) \to -\lim_{\Delta y \to 0} \frac{\Delta y^2}{2} I. \tag{3}$$

Here, the amplitude of the K-L information becomes essentially the Fisher information, which differs from the more well-known Shannon measure in that it is a measure of an observer's ability to estimate some parameter (Fisher, 1925, 1959; Frieden, 2004). Usually the parameter numerically fixes an observed phenomenon in the form of a set of measurements (i.e., the difference between $p(y_i)$

and $r(y_i)$). Therefore, information about the parameter is provided by the measurements. Fisher information is usually written

$$I = \int \frac{1}{p(\varepsilon)} \left(\frac{\mathrm{d}p(\varepsilon)}{\mathrm{d}\varepsilon} \right)^2 \mathrm{d}\varepsilon. \tag{4}$$

This measures how well the measurements can be used to define the actual value of the phenomenon. The relevant probability function is $p(\varepsilon)$ where ε measures the deviation of a measurement from the ideal or true value of the variable. Large values for ε imply the measurement is substantially different from the ideal value, while small values indicate the measurement is very close to that value. Thus, if a set of measurements is very inaccurate, the deviation ε might, with *equal* probability, be very large or very small. Then $p(\varepsilon) \approx$ const. and Eq. (4) gives $I \approx 0$, very little information. This makes sense for these conditions.

2.1.1.1. Principle of extreme K-L information. The PDF $r(y_i)$ is often called a "reference" law, for the following reason. Suppose that, for a given PDF $r(y_i)$, the $p(y_i)$ are sought that extremize $K_{\mathrm{KL}}(p/r)$ subject to constraints, as in a problem

$$-\sum_i p(y_i) \log_b \frac{p(y_i)}{r(y_i)} + \sum_j \lambda_j \sum_i p(y_i) f_j(y_i) = \text{extrem}. \tag{5}$$

This is called the "principle of extreme K-L information" (Frieden, 2001). Here the $f_j(y_i)$, $j = 1, \ldots, J$ are known constraint kernels. (An example might be $f_j(y_i) = {y_i}^j$, where the constraints are moments of order j.) This is characteristically a Bayesian approach to estimating the $p(y_i)$. That is, the constraints are any the observer thinks "reasonably characterize" the given process. They are not necessarily, for example, the most important or most fundamental in some sense. Obviously, the resulting solution for $p(y_i)$ cannot, then, be regarded as fundamental either. Rather, it is considered to be as "reasonable" as were the constraints.

The solution to Eq. (5) is obtained simply by differentiating $\partial/\partial p(y_k)$ it and equating this to zero. The result is a solution

$$p(y_k) = r(y_k) \mathrm{e}^{-1} b^{\sum_j \lambda_j f_j(y_k)}, \qquad k = 1, \ldots, N. \tag{6}$$

This shows a proportionality

$$p(y_k) \propto r(y_k) \tag{7}$$

between the estimate $p(y_k)$ and the input law $r(y_k)$. That is, the estimate $p(y_k)$ is biased toward $r(y_k)$. Or, $r(y_k)$ acts as a *reference function* for $p(y_k)$. Also, from Eq. (7) $p/r \sim 1$. This is the reason for the division notation $K_{\mathrm{KL}}(p/r)$, and also why this information is often called the "cross entropy" between p and r.

2.1.1.2. A transition to EPI principle. The principle of extreme K-L information includes the EPI (Extreme Physical Information) principle within its scope. Briefly, the working hypothesis of EPI (Frieden, 2004) is that the data in any measurements result from a flow of information—specifically, Fisher information—that proceeds from an information source to a sink in data space. The information source is the phenomenon that underlies the data. Its information level is denoted as J. The data information is called I, and the information flow is symbolized by the transition $J \rightarrow I$. The flow physically consists of material particles that carry the information. Each phenomenon generally utilizes different messenger particles (for example, in an optical microscope the particles are photons). The source information J represents the Fisher information that is intrinsic to the measured phenomenon. Assuming that we are dealing with an isolated, passive system, the received information can at best equal the source level. That is, $I \leq J$. The level J corresponds as well to the bound—as opposed to free—information.

The EPI principle is that $I - J =$ extremum. The extremization is attained through choice of the probabilities $p(y_k)$ of the scenario or their square roots (probability amplitudes). These $p(y_k)$ are the predicted system probabilities defining the given system.

In the continuous limit $y_i \equiv i\Delta y$, $\Delta y \rightarrow 0$, and with the special reference function choice Eq. (2) correspondence Eq. (3) holds, so that extreme K-L principle becomes

$$-\frac{\Delta y^2}{2} I + \sum_j \lambda_j \sum_i p(y_i) f_j(y_i) = \text{extrem.} \tag{8}$$

If the constraints j are chosen such that their sum in Eq. (8) is effectively the Fisher source function $J = J[f_1(y), \ldots, f_J(y)]$ for the problem, then Eq. (8) takes the form of the EPI principle. This means that for any problem that may be addressed by EPI, the problem could also have been approached using the extreme K-L principle, provided that the appropriate reference choice was used and that the user constructed (exercising good intuition) the special set of constraints that is identical with the theoretical source functional J for the problem. We have found the EPI approach to have significant biological applications particularly in cancer as outlined below.

2.1.1.3. Biological interplay of system and reference probabilities. An important but little recognized constraint on the dynamics of information gathering, utilization and storage in biological systems is the corresponding cost (see below). That is, information is not free and the energy cost of measuring and reconstructing the external probability distribution function (where $p(y_i)$ is an internal picture of the external environment described by $r(y_i)$) requires optimization strategies that trade off accuracy for cost. The most accurate measurement of the environment corresponding to the greatest quantity of K-L information requires all of the senses to be maximally "tuned" to detect all possible sights, sounds, smells etc. (i.e., $p_i = r_i$). In fact, most organisms scan their environment with a single highly

tuned sensory system while the others are used much less. We propose these strategies have evolved to give the organism a sufficiently accurate picture of the environment to maximize its survival probability while minimizing the cost in terms of energy output both for the use of the sensory system and the cost of storing, transmitting, and transcribing the greater heritable information content required to build more complex, high resolution systems.

Thus, examining the interplay of p and r through K-L information seems to have great potential in understanding biological systems and their evolution. For example, consider a species of animals as they apply K-L cross entropy and continuously generate internal measures of critical environmental parameters such as threats, food sources, mating opportunities etc. This ability to accurately model the environment allows the individual to avoid dangers, evade predators, and find food. The fitness and function of each individual is, therefore, dependent on the accuracy of the K-L function to maximize its fitness. At the same time each component of the internal picture $p(y_i)$ requires energy so that increasing K-L information also requires the individual find and acquire more resources from the environment. In this context, evolution serves as an optimization process that balances the survival and proliferation benefits of information against the cost of obtaining it. Of course the local ecology is also in continuous flux due to climate changes, evolution or local species, invasion by foreign species and so on. This continuous change in $r(y_i)$ acts as a forcing function on $p(y_i)$ and drives adaptation and evolution of the K-L optimization in each population.

2.1.1.4. Application of K-L principle to developmental biology. Many interesting potential applications for Kullback–Leibler information may be found in *developmental biology*. Here the information content of a single fertilized egg is translated into large, complex multicellular organisms, through the process of growth. Then, for efficient development, the probability distribution function that describes the *mature* organism—call it $p(y_i)$—must be nearly identical to the information encoded in a probability distribution function within the *egg*—call it $r(y_i)$. That is, the bias property Eq. (7) must hold again. Developmental biology can thus be modeled as a K-L process in which the probability function of the fertilized egg acts as a reference function for future development and is replicated stepwise with maximal accuracy. This process should, then, obey the extremization principle, with appropriate constraint kernels $f_j(y_i)$. This relationship ought to define the general principles that govern developmental biology. To our knowledge, this idea has not yet been carried through. We predict that its implementation will lead to new insights into developmental biology, both regarding normal and abnormal development.

As a working alternative, extremization principle can also go over into the *EPI principle*, as we showed above. The EPI principle is mathematically different enough from the more general principle to give different results. Hence, it might well be worth trying. It certainly worked in application to abnormal, i.e., cancerous, growth as discussed below.

We next examine special cases of K-L information that are of the Shannon form.

2.1.2. Shannon information types

Shannon information and transinformation follow as special cases $r(y_i)$ of the K-L information. In the case of a constant, unit reference function $r(y_i) = 1$, Eq. (1) becomes

$$K_{\mathrm{KL}}(p/r) \equiv H(Y) = -\sum_i p(y_i) \log_b p(y_i). \tag{9}$$

This is called the "Shannon information" (sometimes the "Shannon entropy").

In the alternative case of joint, or *two* dimensional probabilities $\mathbf{y} \equiv (y_i, y_j)$, $i, j = 1, \ldots, N$, Eq. (1) becomes

$$K_{\mathrm{KL}}(p/r) \equiv -\sum_{i,j} p(y_i, y_j) \log_b \frac{p(y_i, y_j)}{r(y_i, y_j)}. \tag{10}$$

If the reference function is constructed as $r(y_i, y_j) = p_1(y_i)p_2(y_j)$, where the latter are the marginal laws for $p(y_i, y_j)$, the K-L form in Eq. (10) becomes

$$K_{\mathrm{KL}}(p/r) \equiv -S = -\sum_{i,j} p(y_i, y_j) \log_b \frac{p(y_i, y_j)}{p_1(y_i)p_2(y_j)}. \tag{11}$$

This is exactly the negative of the Shannon trans- (or mutual) information S. Information S represents the information in a signal y_i about *another* signal y_j (thus the "trans"); or vice-versa.

The Shannon information in Eq. (8) defines H as "information, choice, and uncertainty" in a signal Y which possess possible configurations y_i with respective probabilities $p(y_i)$.

In summary, a variety of information metrics are available to apply to biological systems. It is conceptually useful to start with the K-L information because the more commonly used biological metrics, including Shannon information or Fisher information, follow from it depending on the particular choice of the reference function $r(y)$.

2.1.3. Information as an expenditure of energy

Shannon also defined H as an "entropy," because of its formal similarity to H in Boltzmann's theorem. This has led to the concept of information as "negative" entropy (sometime called "negentropy" although neither term is technically correct since entropy is a negative number). Clearly there is a qualitative relationship between information and thermodynamic entropy in that the information content of a system increases only if entropy (randomness) decreases. Similarly inflow of information increases system order (Fig. 1). A quantitative relationship, although subject to disagreement, between information and energy has been proposed (Pierce, 1980; Schneider, 1991a,b):

$$H_{\mathrm{bit}} \geq kT \ln 2. \tag{12}$$

Here H_{bit} is the amount of energy required to acquire 1 bit of information, k is the Boltzman constant and T is the absolute temperature.

Equation (12) allows quantification of the information-energy optimization process that governs information dynamics as discussed above. Thus, any biological process that obtains, stores, or uses information must expend *at least* the equivalent energy in Eq. (12) for that information. This places limits on living systems as they balance their need for information with their ability to pay the energy cost of its storage and transmission. Optimization of this process by evolution will favor organisms that maintain enough information to encode a complex and robust system, *and no more*.

2.1.4. Some problems with biological applications Shannon information

Perhaps the first comprehensive analysis of information dynamics in living systems was performed by Schrodinger in his book "What is Life" (Schrodinger, 1944). At the time, the need for information storage and transfer across generations was apparent but the mechanisms largely unknown. Schrodinger derived general principles for the biology of information storage by noting that cells are continuously buffeted by thermodynamic perturbations. For this reason, atomic level intracellular structures are not sufficiently stable to encode heritable information. Rather, he predicted that transgenerational information must be encoded in molecules with an aperiodic structure. Subsequent discoveries of the DNA structure, consisting of sequences of the four nucleotides A, T, C, and G, of course confirmed his prediction. As a sequence of symbols, polynucleotides and polypeptides are ideally suited to analysis by Shannon's methods (Hariri et al., 1990; Strait and Dewey, 1996; Zeeberg, 2002; Weiss et al., 2000). A variety of authors have used information theory to quantify the information content of genes, examine the conservation of information during evolution (Brooks et al., 1984; Wallace and Wallace, 1998; Schneider, 2003; Dehnert et al., 2005), and the origin and evolution of biological complexity (Adami et al., 2000; Segre et al., 2000). Application of Shannon methods to ecology have met with some success, particularly in describing the interactions among multiple species (Ulanowicz, 2001; Fath et al., 2003). While quite interesting, this type of analysis was well-described in the previous review of Johnson and will not be examined in detail here.

Rather, the focus will be on new methods, since it seems clear that, in the 35 years since Johnson's original article, IT using traditional Shannon methods has *not* become, as predicted, the "general calculus" of biology. Although the lack of widespread application is probably the result of multiple limitations, several issues stand out:

First, Shannon methods that quantify the content of information in a biological structure say nothing about its *meaning*, *cost*, or *function*.

Second, such traditional IT models typically did not address the precise mechanisms by which information is used to perform work and maintain cellular stability.

Third, IT tended to view information as a quantity that is simply exchanged between individuals. Rather, it has become increasingly clear that biological information flows along complex *pathways*, with positive and negative feedback loops and substantial temporal and spatial plasticity.

The limitations in traditional IT have been addressed by a number of investigators who have applied Shannon methods in novel ways or developed new mathematical approaches to information dynamics.

3. Intracellular information dynamics

3.1. Bioinformatics and network analysis

Some of the most important advances in theoretical understanding of intracellular information dynamics have stemmed from applications of methods from dynamical systems and bioinformatics. The former grew out of graph theory and is used to characterize the structure and interconnections of complex systems such as electricity grids and the internet. Early biological applications examined metabolic pathways (Jeong et al., 2000; Albert and Barabasi, 2002). Other studies have focused on information pathways—the *interactome*, see Fig. 2—in which a node represents an information carrier (typically protein or RNA) and the flow of information defined by the interactions of these carriers (Wagner and Fell, 2000; Yeger-Lotem et al., 2004).

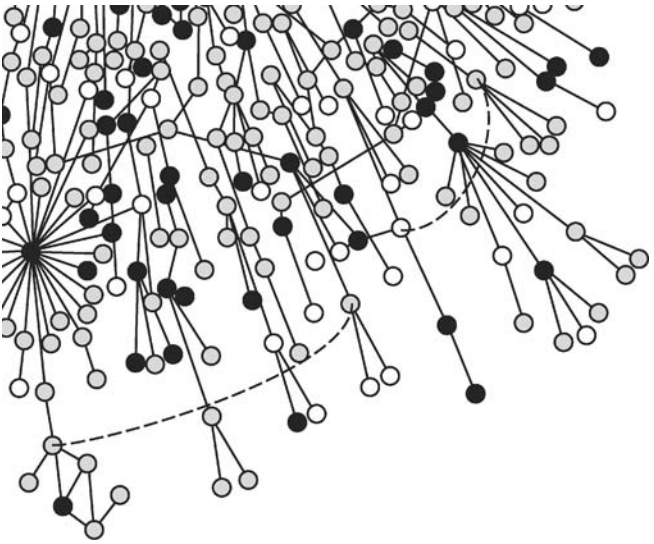

Fig. 2 The organizational principles of a scale free network of information flow in cells. Each node is a class of proteins and the connections are based on direction interactions of the protein molecules or flow of substrate from one protein to another (note this could be metabolites from a enzyme catalyzed reaction or ions from a protein pump within the membraned). The number of connections for each node is described by a power law distribution so that most nodes have few connections but a few nodes have many connections and are designated as hubs. The information flow dynamics of hubs is bimodal with some (party hubs) demonstrating temporally synchronous connections while others (date hubs) exhibit greater heterogeneity and typically connect with nodes that are more physically distant (dashed lines).

This work has been made possible by the rapid advances in genomics and proteomics that measure the intracellular expression of the information in each gene based on the number or RNA or protein copies present. A necessary parallel development included new methods in bioinformatics (Yook et al., 2004), which encompasses diverse statistical methods used to extract information from large data sets produced by the new molecular biology technologies.

Information networks contain junctions, or nodes, which typically represent a distinct protein species or connections based on the interaction of one node with or protein with another either by direct contact or indirectly as the products of one interaction become the substrate for another enzyme. It appears that intracellular information networks are, at least to a first approximation, organized in a scale-free topology, so that nodal connections can be described by a distribution function

$$P(k) \sim k^{-\gamma}, \tag{13}$$

where $P(k)$ is the probability that a node will have k connections, and γ is the scaling exponent. This is a simple power law. Thus, a few nodes have many connections while most have only one or a small number. Highly connected nodes (arbitrarily defined as $\geq$5 connections (Grunenfelder and Winzeler, 2002; Li, 2004) are called *hubs*. This topology has the advantage of allowing high connectivity and small diameter (defined as the average length of the shortest path between any two nodes). Interestingly, power law behavior such as Eq. (13) follows from EPI (Gatenby and Frieden, 2002; Frieden, 2004). This means that power law behavior follows not only from an efficient allocation of energy within a system but also from an efficient flow of information. In effect, the scale free dynamics of intracellular networks follows from principles of information theory.

3.1.1. Information flow through hubs and links

The dynamics of information flow in these biological networks is beginning to be understood. Han et al. (2004) used microarray data before and after various cellular perturbations to define the synchronization of activity among nodes. They found consistent evidence for two distinctly different hub dynamics. Some of the hubs were highly connected with others, so that the groups reacted synchronously—a pattern they described as "party hubs." In other hubs, the connected nodes demonstrated greater temporal heterogeneity in their responses and were termed "date hubs." This bimodal behavior was highly dependent of spatial distribution of the proteins—date hubs tended to involve information flow over greater distances within the cell while party hubs connected proteins that were "near neighbors." Interestingly, they also demonstrated that removal of date hubs resulted in much greater disruption of the network than party hubs. These results demonstrate that information flow within the interactome is complex and probably nonlinear. In fact, maps of cell signalling typically demonstrate multiple parallel, but interacting, pathways. It seems clear that further examination of the principles governing transmission of biological information is necessary and will likely be extremely interesting. For example, it appears that "on–off" phenomena in biological systems (such as neuron firing or cellular entry into the cell cycle) involves

multiple continuous positive and negative information pathways. Thus, a neuron typically receives constant pulses from other fibers which produce local depolarizations that sum over time and space. Only when the depolarization exceeds some threshold value will a propagating wave result (the "on–off" switch). Similarly, environmental signals regarding cellular proliferation in multicellular tissue are carried through multiple interacting pathways with positive and negative feedback loops that may either promote or inhibit entry into the cell cycle. As with neurons, the on–off decision of the cell to enter the cell cycle (like the neuron, this "decision" is irretrievable since a cell, once in the cycle, must either undergo mitosis or die) seems to require that the summation of these positive and negative signals exceed some threshold. These observations suggest that critical cellular messages may, as in the internet, be broken into components and sent along different paths. It is likely this strategy insures rapid and maximally accurate transmission in a noisy system but far more study will be necessary to clarify these complex dynamics.

It is important to recognize that the interactome represents a map of a *potential* information network—the actual network will vary between individuals and in the same individual over time. For example, Luscombe et al. (2004) demonstrated that only a few hubs are completely stable over time, while most exhibit transient variations in connections in response to a variety of intracellular and environmental changes. In other words, information flows through highly plastic networks that are continuously rewired in response to stimuli that carry information from outside the network. These results suggest future work will also need to focus on the dynamics within nodes, so that the flow and connectivity become explicit functions of other system parameters, such as the concentration of the node protein that is present. For example, a node and its connections will be lost when the protein is not expressed, or if its concentration falls below some threshold. Similarly, the number of functioning links to a node or the speed of flow through the links may be dependent on protein concentration.

Many additional challenges remain for this exciting, and highly productive, avenue of investigation. Significant variations among the data sets on which the models are based will need to be resolved and other analytic methods such as the network motif approach should be examined.

3.1.2. Dynamics of system failures

Although the general focus of this work is on normal cellular function, much may be learned from understanding the topology and dynamics that lead to failures in the intracellular information system. Topological analysis of scale-free networks demonstrates that their heterogeneity confers resistance to failure caused by random loss of nodes. For example, it has been demonstrated that network function is maintained even when up to 8% of the nodes are randomly removed (Albert et al., 2000; von Mering et al., 2002). On the other hand, scale-free networks are vulnerable to selective attack. For example, removal of 8% of the *most-connected* hubs resulted in a 500% decline in network function as measured by average path length (Callaway et al., 2000; Jeong et al., 2001).

The dynamical aspects of system failure are suggested by Zhao et al. (2004), who proposed that in each time step of normal system function the relevant quantity (in this case information) is exchanged between every pair of nodes and travels along the shortest possible pathway. A node i will ordinarily have some number of shortest pathways running through it—defined as its load L. The capacity C of the node is the maximum load it can carry, with

$$C_i = (1 + \alpha) L_i \tag{14}$$

where α is the tolerance parameter. They found that the loss of a node results in a redistribution of flow through other nodes. When this exceeds the node capacity, cascading failure will result in system disintegration. Consequently the vulnerability of the system to attack is dependent on the value of α.

Another way of stating this is that the robustness of intracellular information networks comes at a cost—the expenditure of resources necessary to increase the number of connections and/or the nodal capacity (Zhao et al., 2004). Thus, both topological and dynamical analysis of system failure illustrates the central biological cost-benefit calculation apparent in Eq. (13).

Living systems must balance their need for robust information dynamics with cost. Therefore, eventually the energy costs of forming and maintaining large information networks will need to be included in the analyses. These constraints place boundary conditions on biological information networks such that each cell can be expected to evolve system topology and dynamics that approach the minimum necessary to achieve sufficient robustness. A level of robustness is needed that enables the cell to resist the range of typical environmental disturbances. Chief among these are perturbations in physical parameters such as temperature and biological disturbances due to parameters such as infectious agents.

3.2. Information and cellular fitness

An alternative approach to addressing the value and cost of information assigns to the information content of each gene (determined by Eq. (9)) a fitness value defined by its contribution to the survival minus its cost (Gatenby and Frieden, 2002, 2005a). In the context of the dynamical system models, date hubs, for example, will have a higher value than party hubs, which in turn will typically have a higher value than individual nodes. This potential variation in the fitness value of each component of the information network may play a role in its observed plasticity. For example, the value of the nodes and connections of the four proteins that constitute glycine decarboxylase under normal conditions have a low fitness value resulting in normal conditions, and are therefore not detectable. However, this component of the system assumes great value when a cell is compelled to use glycine as a carbon source. Since the benefit now exceeds the cost, this hub becomes highly active.

In addition the "fitness" of cells in multicellular organisms is highly contextual. When it is defined as survival and function of the multicellular organism, the value of cellular differentiated genes is high because they maintain the stability of the

entire system. However, if fitness is defined simply by individual cellular proliferation, then gene segments that encode differentiated cell function have a *negative* value (i.e., their contribution to proliferation is 0, so subtraction of the energy cost of the information yields a negative fitness value). This binary fitness effect may be observed during carcinogenesis, as the transformed cells act increasingly as individuals rather than as components of a multicellular society. Thus, random mutations that disable differentiated functions will be favored during somatic evolution—a process that manifests itself as dedifferentiation, and is characteristically observed in invasive cancers.

3.3. Cellular information utilization

The preceding methods have provided many insights into the general organization and dynamics of information flow within cells, but do not address the nitty gritty question of how this information is actually *used* by the cell to perform work. Schneider and colleagues (1991a,b, 1997) have examined biological utilization of information through a novel approach in which macromolecules (typically polypeptides or polynucleotides) or macromolecular complexes function as *isothermal machines* operating in a cycle. Each such machine contains information originally encoded in the genome, then transmitted through the sequence of amino acids and, finally, manifest in a tertiary structure which permits highly specific allosteric binding. The cycles (Fig. 3) begin when the machine is primed to a higher energy state so that its binding sites are available and, therefore, ready to do work. During the cycle, the machine expands its information content by "choosing" a specific ligand over all other possibilities—a process manifested by the energy of binding. In the *operation* phase of the cycle, the machine dissipates this energy by performing work that benefits the cell. This approach explicitly integrates information with cellular thermodynamic processes, and allows the value and meaning of the encoded information to be defined by the specific cellular function performed.

Other studies suggest that this general approach can be broadly applied to mechanical and enzymatic protein function (Lahoz-Beltra, 1997). In fact a theoretical model of cells as a collection of molecular machines has been proposed (Alberts, 1998) and it seems likely that a combination of molecular machines and network analysis will yield substantial insights into the coupling of cellular information dynamics with cellular function.

3.4. Potential Controversies—is all cellular information stored in the genome and transmitted by proteins?

Implicit in vitually all of the work on cellular information is an assumption that the genome and its RNA and protein products represent a complete representation of information dynamics. However, the results of Luscombe et al. (2004) demonstrate that the interactome, in fact, adapts to information flow from *external* sources which may be intra- or extracellular.

Gatenby and Frieden (2005b) have proposed the genome in fact represents only one component of a broad ensemble of cellular information carriers. They point out that Eq. (9) demonstrates information may be encoded in any cellular structure

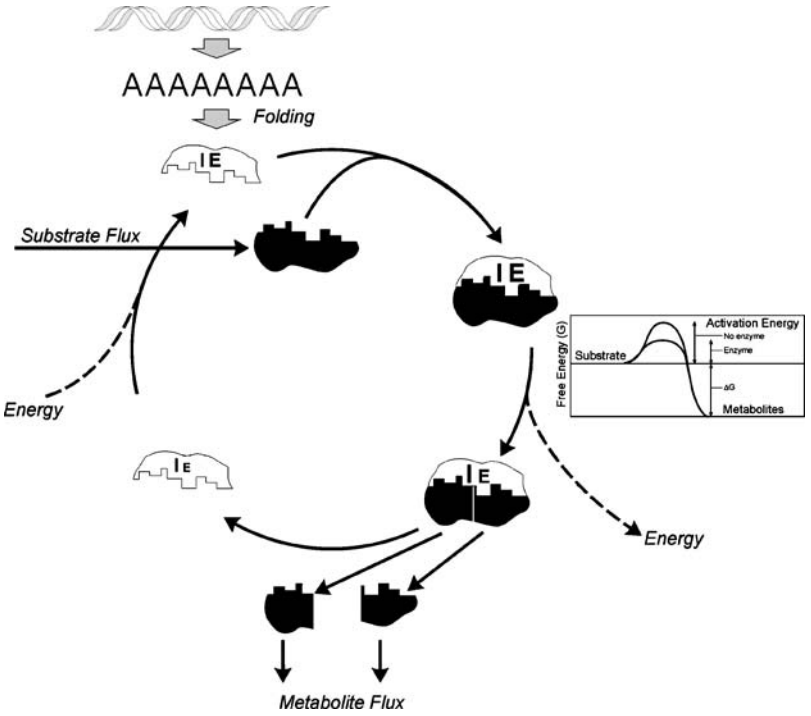

Fig. 3 This sequence illustrates the molecular machine model that translates information into useful work developed by Schneider. Information stored in DNA flows to proteins through the amino acid sequence which, after protein folding, results in a specific shape. On the top left the protein is primed with a baseline level of information and energy (designated by the size of the **I** and **E**). The protein chooses to bind with a specific substrate (and reject all other possibilities). This creates a molecular complex in which the information content and energy (through binding) is maximized. This is dissipated by reducing the activation energy of the reaction allowing the substrate to be split into two metabolites with release of energy (ΔG). The release of the metabolites reduces the energy and informational content of the enzyme, which is then primed by the addition of new energy from the environment (e.g., through thermodynamic fluctuations or ATP). The net cycle allows information to be used to generate work. Note that the reaction produces molecular flux in which substrate is consumed and metabolites are generated (see Fig. 4).

whose structure or distribution is described by a nonrandom probability function. The linear distribution of nucleotides A, T, C and G along a DNA "string" represents a clear and well-recognized example of such an information encoding structure. However, there are many other highly nonrandom cell structures including lipids, cell membranes, and transmembrane gradient of ions. Some components of the membrane such as the ABO antigen are directly encoded by the DNA content through expression of specific enzymes and are thus directly linked to genetic information. However, other components of the membrane such as lipids demonstrate significant variations in molecular structure and highly nonrandom concentration distributions among different cells (Alberts et al., 1994). There is no apparent direct link between the information encoded (by Eq. (9)) in these ordered structures and the genome suggesting they may be sources of cellular information

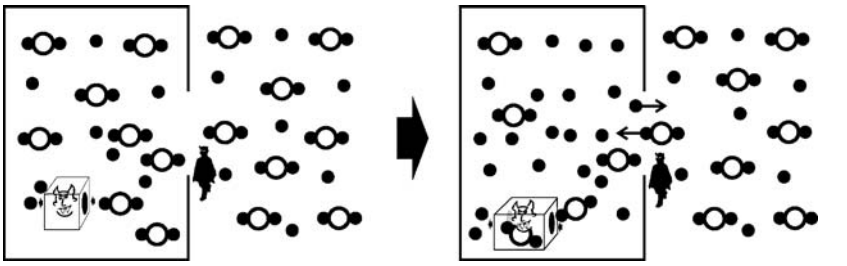

Fig. 4 An alternative demon *gedanken* experiment in which a container is initially at equilibrium with an environment so that two different molecules are equally distributed (shown in the left box). Within the container is a box housing an ogre who at some time t_0 begins to "eat" the central portion of the larger molecule and releases the two peripheral metabolites back into the environment (similar to the enzyme in Fig. 3). As a result of the ogre's appetite, the concentration of particles within the box changes producing a gradient with the environment and molecular flux through the door. A demon stationed at the door can, by measuring the concentration gradients or the flux, gain information regarding the consumption rate of the ogre. In other words, concentration gradients and fluxes of small molecules represent potential sources of information storage and transmission separate from the protein interactome in Fig. 2.

independent of DNA, RNA, and proteins. In fact, the potential information in membrane lipids of a typical mammalian cell has been calculated to be on the order of 5×10^{10} bits (significantly larger than the information content of the genome). Similarly, transmembrane ion gradients represent substantial displacements from randomness and, therefore, the possibility of substantial encoded information. For example, the potential information content of the transmembrane ion gradients of Na^+f02b can be calculated (assuming intracellular concentration of about 20 mM and extracellular concentration of 140 mM) by Eq. (9) as approximately 1.2×10^8 bits per cell.

What role might these information carriers play? Clearly the genome acts as the long term storage mechanism for heritable information. However, information flow from the environment into the cell or within cellular components over short time periods could very well utilize nongenomic, nonprotein information carriers. For example, information encoded in molecular gradients could relate to the constantly changing metabolic status of the cell (Figs. 3 and 4), or to the extracellular concentrations of substrate or signals. Finally, long-term information storage and transgenerational transmission is costly, but can be augmented if heritable information is expanded by encoding proteins that can, in turn, generate additional information. Thus, proteins that are transmembrane ion pumps can act as Maxwell demons generating additional information (through the trans membrane ion gradients as outline above) in each new cell.

Independently, Keener and colleagues (Dockery and Keener, 2001; Keener, 2005) recently demonstrated that information stored in gradients can be transmitted by molecular and ionic flux (Fig. 4). The general concept is evident in the equation for diffusion of particles across membranes:

$$J_{\mathrm{M}} = \frac{D_{\mathrm{M}} A_{\mathrm{M}}}{L_{\mathrm{M}}} (C_{\mathrm{O}} - C_{\mathrm{I}}) \tag{15}$$

where J_M is the flux across a membrane, D_M is the diffusion coefficient, A_M is the surface area of the membrane, L_M is the length of diffusion across the membrane and C_O and C_I are the outside and inside concentrations. Keener points out that the flux may actually encode information about the parameters of Eq. (15). For example, if the system is subject to perturbations in the extracellular concentration of some nutrient C, a measurement of flux will also allow measurement of the extracellular concentration C_O. He has recently demonstrated that bacteria can measure flux through positive and negative feedback loops, and thereby use that information to obtain extracellular and intracellular information, specifically the size of a local population and the length of flagella, respectively.

Obviously this work remains speculative, it does have the potential to significantly expand knowledge of the topology and dynamics of intra- and extracellular information networks.

4. Extracellular information dynamics

The manner by which information dynamics govern development and function of multicellular living structures remains largely unexplored and an exciting future challenge for IT.

Clearly, the stable synchronization of large numbers of cells requires complex, but accurate information flow. The need for increased information transmission and reception in multicellular organisms is evident in the genome of the cyanobacterium *Nostoc punctiforme*, which forms a simple filamentous multicellular structure and contains 7,400 genes. Another unicellular cyanobacterium *Synechocystis* contains 3,200 genes (Kaiser, 2001).

Insight into at least some of the principles of intercellular information dynamics can be obtained in rudimentary bacterial communication during quorum sensing (Taga and Bassler, 2003; Zhang and Dong, 2004). Typically information is carried by small molecules (autoinducers) produced and excreted into the environment at a regular rate by each bacterium. The molecule freely diffuses through the bacterial wall and, typically, binds to a regulator protein in the cytoplasm. The regulator/autoinducer complex then binds to a specific DNA promoter segment upstream of the protein that catalyzes formation of the autoinducer. This produces a positive feedback loop (Surette et al., 1999). Once a critical concentration of the autoinducer is exceeded, the signal is dramatically amplified by the autocatalytic pathway. In turn, this produces a population phase change in which all of the individuals act in unison. Here the content of the message is simply: "I am here." The resulting environmental information is the summation of these messages. These are conveyed to each individual through a transmembrane molecular flux that is coupled to an intracellular positive feedback loop.

Similar but more complex information dynamics is observed in the aggregation of the amoebae Dictyostelium, and has been extensively modeled (Franca-Koh and Devreotes, 2004; Chamaraux, 2005). This is a creative reaction to an imminent state of starvation. First, the individual amoebae must accurately detect information that starvation is imminent. Second, individuals in the population must be instructed to aggregate into a fruiting body through a chemotactic relay system.

The system is based on detection and production of cAMP, which tells individuals to undergo a phase transition from proliferation to differentiation, and to then migrate in a specific direction.

The information dynamics in these rudimentary multicellular networks seems to be based on pathways that integrate small diffusing molecular messages with the intracellular interactome. In other words, nodes from one individual link to nodes within others via a diffusing messenger. It will be interesting to see if these interactions form a scale-free network that echoes the organizational principles within cells.

Formation of a multicellular society from a single germ cell requires sufficient information content in the initial cell to form a complex multicellular organism. In the case of humans, the latter consists of approximately 10^{14} cells. Understanding this synthesis is undoubtedly the greatest challenge for information theory. Indications are that organizational principles of developmental biology include multiple diffusing information carriers and specific receptors. These build on communication strategies observed in rudimentary multicell societies. However, it appears that more complex multicellular organisms have also developed additional mechanisms of information exchange between cells in contact. This exchange allows a flow of information across cell membranes via direct interactions of membrane proteins. These interactions also occur through intercellular channels that are formed by gap junctions, and probably involve a bidirectional flux of ions and small molecules. Since these links are dependent on cell contact they may well exhibit a Poisson rather than a power law distribution. This suggests that complex multicellular organisms may use a combination of contact-mediated links between nodes of next-nearest neighbors, and longer distance, nonrandom links through diffusing information carriers and specific receptor molecules. This combination of random and scale free networks may produce information dynamics substantially different from that of the cellular interactome.

4.1. Information and disease

Ultimately, the goal of IT is to use insights into intra- and extracellular information dynamics in order to understand disease and suggest therapeutic strategies (Morris, 2001). It seems clear that biological information networks are designed to be adaptive and robust and, therefore, resistant to failures due to environmental perturbations or direct attack. Disease must represent some failure of the system due to mutations or attacks for which it was unprepared.

The role of information degradation in disease appears to exhibit both simple linear and highly complex nonlinear dynamics. For example, sickle cell disease is a straightforward result of biallelic mutations that alter a specific node in the interactome (the hemoglobin protein). Similarly, some infectious agents (HIV for example) selectively attack specific nodes of specific cell types, producing a complex but relatively straightforward series of events in the cells and organism. However, given the complexity of the informational networks it is not surprising that many diseases appear to result from a disruption of multiple nodes and interactions, such that consistent correlation with a single gene mutation or external stimulus will not be possible.

The multistep transition from normal tissue through a variety of intermediates to invasive cancer represents an interesting model for these dynamics. Carcinogenesis requires accumulating genetic mutations and chromosomal defects—a process often termed "somatic evolution" (Garcia et al., 2000). In fact, Loeb and others hypothesize an increased mutation rate as a necessary condition for cancer formation (Loeb, 2001).

Clearly, degradation of the genome due to mutations will result in disruption of critical information pathways. It therefore appears that somatic evolution of the malignant phenotype requires a loss of cellular information. However, while random genetic mutations appear necessary for carcinogenesis, two observations indicate that the information dynamics are more complex than this simple picture:

1. A constant mutation rate, in the absence of modifying constraints, will result in an "error catastrophe" (Eigen and Schuster, 1977; Kendal, 1990; Sole and Deisboeck, 2004) in which genomic information decays to the point that it insufficient to maintain life, even disordered, nonfunctioning carcinogenic life.
2. The cancer phenotype exhibits apparent *nonrandomness*, in that differentiated cell functions are progressively (semideterministically) lost, while genes necessary for proliferation remain functional (and often up-regulated) even in the most advanced cancers.

Using IT, it has been demonstrated that information degradation from the mutator phenotype is highly constrained by Darwinian selection during somatic evolution of the malignant phenotype (Gatenby and Frieden, 2002, 2005a). These selection dynamics will preserve the genomic information content necessary for growth, while the information content of gene segments that suppress growth such tumor suppressor genes will be rapidly degraded. As described above, genes that encode the differentiation function will have negative fitness values, and are subject to progressive degradation. These dynamics allow tumor cells to preserve information that is necessary for survival and proliferation, while progressively losing information that encodes cellular function in a multicellular society (and the energy demands for maintaining and transcribing that information). That is their dual strategy for winning out.

The summation of these effects is that the global information content of cancer cells will asymptotically approach the minimum necessary to maintain proliferation. Interestingly, the assumption of a minimal information state leads to a unique prediction for the growth law governing an in situ tumor mass. Using EPI, it is found that under "free field" growth conditions, i.e., unconstrained by host response, substrate limitation or any curative program, the growth law has the power law form t^{γ}, with t the age of the tumor and the exponent $\gamma = 1.62$. This prediction is in remarkably good agreement with six clinical studies which show that small human breast cancers do empirically exhibit power law growth, and, with an exponent of 1.74 ± 0.22 (Gatenby and Frieden, 2002).

Clearly, IT has potential for defining the informational basis of disease and this application remains a fascinating future challenge.

5. Conclusions

Living systems, uniquely in nature, exist at the dynamical interface of information and thermodynamics. Thus, examination of information storage, processing and communication is integrally linked to the study of biology. Past decades have seen remarkable technological developments in this direction, yielding substantial insight into the informational bases of life. Meanwhile, information theory—the study of information storage, communication and processing—has similarly made great strides in providing conceptual frameworks to define the underlying dynamics. These provide novel insights into the underlying structure of information networks and the mechanisms by which information is converted to work and to function. These theoretical developments have been accomplished using both traditional Shannon methods and new mathematical techniques such as network analysis, Fisher measurement theory, EPI and game theory. It is likely that, in coming years, these tools will be applied to many other important biological problems, including information flow among cellular and subcellular components of multicellular organisms and the disruption of biological information in disease. The results will almost certainly be fascinating and the prospects are exciting.

References

Adami, C., Ofria, C., Collier, T.C., 2000. Evolution of biological complexity. Proc. Natl. Acad. Sci. 97, 4463–4468.

Albert, R., Barabasi, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97.

Albert, R., Jeong, H., Barabasi, A.-L., 2000. Error and attack tolerance of complex networks. Nature 406, 378–382.

Alberts, B., 1998. The cell as a collection of protein machines:preparing the next generation of molecular biologists. Cell 92, 291–294.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D., 1994. Molecular Biology of the Cell. Garland Publishing Inc., New York.

Brooks, D.R., Leblond, P.H., Cumming, D.D., 1984. Information and entropy in a simple evolution model. J. Theor. Biol. 109, 77–93.

Callaway, D.S., Hopcraft, J.E., Kleinberg, J.M., Newman, M.E., Strogatz, S.H., 2000. Network robustness and fragility: Percolation on random graphs. Phys. Rev. Lett. 85, 5468–5471.

Chamaraux, F., Fache, S., Bruckert, F., Fourcade, B., 2005. Kinetics of cell spreading. Phys. Rev. Lett. 94, 158102–158110.

Dehnert, M., Helm, W.E., Hutt, M.-T., 2005. Information theory reveals large-scale synchronization of statistical correlations in eukaryote genomes. Gene 345, 81–90.

Dockery, J.D., Keener, J.P., 2001. A mathematical model for quorum sensing in Pseudomonas aeruginosa. Bull. Math. Biol. 63, 95–116.

Ebeling, W., Frommel, C., 1998. Entropy and predictability of information carriers. Biosystem 46, 47–55.

Eigen, M., Schuster, P., 1977. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Naturwissenschaften 64, 541–565.

Fath, B.D., Cabezas, H., Pawlowski, W., 2003. Regime changes in ecological systems: an information theory approach. J. Theor. Biol. 222, 517–530.

Fisher, R.A., 1925. Theory of statistical estimation. Proc. Cambridge Phil. Soc. 22, 700–725.

Fisher, R.A., 1959. Statistical methods and scientific inference, 2nd edition. Oliver and Boyd, London, pp. 1–112

Franca-Koh, J., Devreotes, P.N., 2004. Moving forward: Mechanisms of chemoattractant gradient sensing. Physiology 19, 300–308.

Frieden, B.R., 2001. Probability, Statistical Optics and Data Testing, 3rd edition. Springer-Verlag, Berlin.
Frieden, B.R., 2004. Science from Fisher Information, 2nd edition. Cambridge University Press, Cambridge, U.K.
Garcia, S.B., Novelli, M., Wright, N.A., 2000. The clonal origin and clonal evolution of epithelial tumors. Int. J. Exp. Path. 81, 89–116.
Gatenby, R.A., Frieden, B.R., 2002. Application of information theory and extreme physical information to carcinogenesis. Cancer Res. 62, 3675–3684.
Gatenby, R.A., Frieden, B.R., 2005a. Information dynamics in carcinogenesis and tumor growth. Mutat. Res. 568(2), 259–227.
Gatenby, R.A., Frieden, B.R., 2005b. The role of non-genomic information in maintaining thermodynamic stability in living systems. Math. Biosci. Eng. 2(1), 43–51.
Gilbert, E.N., 1966. Information theory after 18 years. Science 152, 320–326.
Grunenfelder, B., Winzeler, E.A., 2002. Treasures and traps in genome-wide data sets: Case examples from yeast. Nat. Rev. Genet. 3, 653–661.
Han, J-D., et al., 2004. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 430, 88–93.
Hariri, A., Weber, B., Olmsted, J., 1990. On the validity of Shannon-information calculations for molecular biological sequences. J. Theor. Biol. 147, 235–254.
Jeong, H., et al., 2001. Lethality and centrality in protein networks. Nature 411, 41–45.
Jeong, H., et al., 2000. The large-scale organization of metabolic networks. Nature 407, 651–654.
Johnson, H.A., 1970. Information theory in biology after 18 years. Science 168, 1545–1550.
Kaiser, D., 2001. Building a multicellular organism. Annu. Rev. Genet. 35, 103–123.
Keener, J.P., 2005. A model for length control of flagellar hooks of Salmonella typhimurium. J. Theor. Biol. 234, 263–275.
Kendal, W.S., 1990. The use of information theory to analyze genomic changes in neoplasia. Math. Biosc. 100, 143–159.
Kullback, S., 1959 Information Theory and Statistics. Wiley, New York.
Lahoz-Beltra, R., 1997. Molecular automata assembly: Principles and simulation of bacterial membrane construction. Biosys. 44, 209–229.
Li, S., Armstrong, C.M., Bertin, N., et al., 2004. A map of the interactome network of the metazoan *C. elegans*. Nature 303, 540–543.
Loeb, L.A., 2001. A mutator phenotype in cancer. Cancer Res. 61, 3230–3239.
Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A., Gerstein, M., 2004. Genomic analysis of regulatory network dynamics reveals large topological changes. Nature 431, 308–312.
Maxwell, J.C., 1880. Theory of Heat, 6th edition. D. Appleton Co., New York.
Morowitz, H.J., 1955. Some order-disorder considerations in living systems. Bull. Math. Biophys. 17, 81–86.
Morris, J.A., 2001. Information theory: A guide to the investigation of disease. J. Biosci. 26, 15–23.
Pierce, J.R., 1980. Information theory and physics, in introduction to information theory. symbols, signals, and noise. 2nd edition. Dover Publications, New York, pp. 184–207.
Prigogine, I., 1965. Steady states and entropy production. Physica 31, 719–724.
Reza, F.M., 1961. An Introduction to Information Theory. McGraw-Hill, New York.
Schneider, T.D., 2003. Evolution of biological information. Nucleic Acids Res. 28, 2794–2785.
Schneider, T.D., 1997. Information content of individual genetic sequences. J. Theor. Biol. 189, 427–441.
Schneider, T.D., 1991a. Theory of molecular machines. I. Channel capacity of molecular machines. J. Theor. Biol. 148, 83–123.
Schneider, T.D., 1991b. Theory of molecular machines II. Energy dissipation from molecular machines. J. Theor. Biol. 148, 125–137.
Schrodinger, E., 1944. What is Life? Cambridge University Press, Cambridge, U.K.
Segre, D., Ben-Eli, D., Lancet, D., 2000. Compositional genomes: Prebiotic information transfer in mutually catalytic non-covalent assemblies. Proc. Natl Acad. Sci. 97, 4112–4117.
Shannon, C.E., 1948. A mathematical theory of communication. Bell System Tech. J. 27, 379–623.
Sole, R.V., Deisboeck, T.S., 2004. An error catastrophe in cancer? J. Theor. Biol. 228, 47–54.
Strait, B.J., Dewey, T.G., 1996. The Shannon information entropy of protein sequences. J. Biophys. 71, 148–155.

Surette, M.G., Miller, M.B., Bassler, B.L., 1999. Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: A new family of genes responsible for autoinducer production. Proc. Natl. Acad Sci. 96, 1639–1644.
Szilard, L., 1929. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. Z. Physik. 53, 840–848.
Taga, M.E., Bassler, B.L., 2003. Chemical communication among bacteria. Proc. Natl. Acad. Sci. 100, 14549–14554.
Trincher, K.S., 1965. Biology and Information: Elements of Biological Thermodynamics. Consultants Bureau, New York.
Ulanowicz, R.E., 2001. Information theory in ecology. Comput. Chem. 25, 393–399.
von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., Bork, P., 2002. Comparative assessment of large-scale data sets of protein-protein interactions. Nature 417, 399–403.
Wagner, A., Fell, D., 2000. Technical Report No. 00-07-041. Santa Fe Inst.
Wallace, R., Wallace, R.G., 1998. Information theory, scaling laws and the thermodynamics of evolution. J. Theor. Biol. 192, 545–559.
Weiss, O., Jimenez-Montano, M.A., Herzel, H., 2000. Information content of protein sequence. J. Theor. Biol. 206, 379–386.
Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S., Milo, R., Pinter, R.Y., Alon, U., Margalit, H., 2004. Network motifs in integrated cellular networks of transcription and protein-protein interactions. Proc. Nat. Acad. Sci. 101, 5934–5939.
Yook, S-H., Oltvai, Z.N., Barabasi, A.-L., 2004. Functional and topological characterization of protein interaction networks. Proteomics 4, 928–942.
Zhang, L.-H., Dong, Y.-H., 2004. Quorum sensing and signal interference:diverse implications. Mol. Micro. 53, 1563–1571.
Zeeberg, B., 2002. Shannon information theoretic computation of synonymous codon usage biases in coding regions of human and mouse genomes. Genome. Res. 1944–1955.
Zhao, L., Park, K., Lai, Y.-C., 2004. Attack vulnerability of scale-free networks due to cascading breakdown. Phys. Rev. E. 70, 1–4.