

Article

Entropy, Information and Complexity or Which Aims the Arrow of Time?

George E. Mikhailovsky ^{1,*} and Alexander P. Levich ²

¹ CALIBRE, 6354 Walker Lane, Alexandria, VA 22310, USA

² Department of Biophysics, Biology Faculty, Moscow Lomonosov University, Leninskie gory, 1-12, 119991 Moscow, Russia; E-Mail: APL@chronos.msu.ru

* Author to whom correspondence should be addressed; E-Mail: gmikhai@yahoo.com; Tel.: +703-801-6453; Fax: +703-797-8502.

Academic Editor: Kevin H. Knuth

Received: 16 May 2015 / Accepted: 29 June 2015 / Published: 10 July 2015

Abstract: In this article, we analyze the interrelationships among such notions as entropy, information, complexity, order and chaos and show using the theory of categories how to generalize the second law of thermodynamics as a law of increasing generalized entropy or a general law of complication. This law could be applied to any system with morphisms, including all of our universe and its subsystems. We discuss how such a general law and other laws of nature drive the evolution of the universe, including physicochemical and biological evolutions. In addition, we determine eliminating selection in physicochemical evolution as an extremely simplified prototype of natural selection. Laws of nature do not allow complexity and entropy to reach maximal values by generating structures. One could consider them as a kind of “breeder” of such selection.

Keywords: complexity; category theory; entropy; information; evolution

1. Introduction

All of us always feel the irreversibility of time in our everyday lives as a principal difference between yesterday and tomorrow. Such irreversibility was also assumed in science and philosophy, where it is a result of the irreversibility of cause-effect chains, where the cause usually precedes the effect. The only

exception is goal setting, when a goal that is always located in the future plays the role of the cause. In this case, effect is actually an effort to achieve the goal, and it naturally precedes the goal. Nonetheless, even in such a situation, cause-effect relationships suppose an essential difference between the past and the future and, respectively, the irreversibility of time.

However, all of these reasons had either common or philosophical sense, but not a scientific, physical character. Newtonian mechanics considered time as a reversible parameter, and all of its laws are completely symmetrical in time. If one considers moving planets in a solar system or balls on an ideal (without friction) pool table from past to future or from future to past, there would not be any differences from a mechanical point of view. The situation had changed only in the middle of the nineteenth century. In 1851, Rudolf Clausius, based on the ideas of Sadi Carnot, formulated the laws of thermodynamics, the second of which claimed the increasing of entropy in isolated systems. Thus, entropy became the first physical quantity that varies in time asymmetrically.

2. Entropy, Information, Complexity and Time

2.1. Entropy, the Boltzmann H-Theorem and His Formula

After 21 years, Ludwig Boltzmann [1] proved his H-theorem, which shows how the irreversibility of entropy increase can be derived from the microscopic reversibility of processes obeying Newtonian mechanics. As a result, Ludwig Boltzmann deduced the formula given by Max Planck to its well-known form:

$$S = K_b \ln W \quad (1)$$

where S is entropy, K_b the Boltzmann constant equal to 1.38×10^{23} J/K and W the number of microstates related to a given macrostate.

Boltzmann's formula suggested radically new statistical interpretation of entropy and demonstrated its connection to the number of ways in which atoms and molecules could be arranged without a change of the system's thermodynamic state. Never before had one equation been linked to values related to different levels of the hierarchy and having a parameter irreversible in time as a result.

2.2. Information, Shannon Formula and Its Similarity with the Boltzmann One

The true value of the Boltzmann formula that goes far beyond thermodynamics was only realized nearly a hundred years later, in the middle of the twentieth century. In 1948, Claude Shannon, while working on information theory, suggested a formula for informational entropy (H):

$$H = -K \sum p_i \log p_i \quad (2)$$

where K is an arbitrary positive constant and p_i the probability of possible events. Boltzmann's constant K_b from Equation (1) is determined in joules by Kelvin degrees. Both of them, however, are units of energy, and if we make the Boltzmann constant a nondimensional quantity, it becomes equal to just one. Therefore, if we define the events by microstates, consider them equally probable and choose the nondimensional Boltzmann constant, the Shannon Equation (2) becomes the Boltzmann Equation (1). Thus, the Shannon formula can be considered as a generalization of the Boltzmann one to a case with

different probabilities of microstates or, using Shannon's terminology, different probabilities of letters in a message.

Very soon, however, it became clear that Equation (2) quantitatively describes any information. Here, we should note that Claude Shannon did not contrast information and informational entropy. He wrote [2] (p. 50): "Quantities of the form $H = -\sum p_i \log p_i$ (the constant K merely amounts to a choice of a unit of measure) play a central role in information theory as measures of information, choice and uncertainty. The form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics, where p_i is the probability of a system being in cell i of its phase space." At the same time, he distinguished maximum entropy as information of an absolutely unordered state or message created by a purely discrete Markov process, relative entropy as a ratio of H from Equation (2) to maximum entropy and redundancy that equals one minus relative entropy. Maximum entropy is calculated by the same Equation (2) when all of the p_i probabilities are equal one to another. However, when they are not equal and redundancy respectively is more than zero, this formula could be used as a measure of information in quite different fields of knowledge. Claude Shannon himself used it in linguistics [2]. Robert H. MacArthur [3] and Kermit Hutcheson [4] applied Shannon entropy (H), which they prefer to name information (I), to assess species diversity in biology. Norbert Wiener [5] began using the Shannon formula in cybernetics that he created, and Umberto Eco even in poetics [6]. At the same time, there was a confusion related to the fact that, despite the obvious similarities and even homology of the Shannon and Boltzmann formulas, entropy was traditionally seen as something negative, while the information from the very beginning was considered as something obviously positive. As a result, there were attempts to relate the information with some negative entropy (negentropy) [7], considered at the same time as the free energy. This is definitely complicated in all cases because the free energy, as we discuss below, usually destroys or wipes out the information. It should be highlighted here that we cover exclusively quantitative aspects of the information followed from Equation (2) and do not consider its semantic aspects. For example, the message "There is rain in the Sahara today" contains far more semantic information than "There is rain in London today". At the same time, the first has 3.41 (less) while the second 3.66 bits of informational entropy calculated by Equation (2). It is the context of our knowledge of geography and global climate that creates a huge difference in semantic information. Respectively, highly ordered systems described as "AAAAAAAAAAAAAAAAA..." are extremely rare in our world, and therefore, semantic information in such a message would be extremely high. However, the quantity of information in this message measured by Equation (2) equals zero bits.

2.3. Interpretation Entropy as Information

Perhaps entropy inherited its negative connotation from its relation to disorder revealed by Ludwig Boltzmann. Disorder, which is an antonym of the German "Die Ordnung", could not be perceived positively in the nineteenth century. With the advent of quantum mechanics in the twentieth century, order began to be associated with the degeneracy of the wave function, simplification of the quantum system and the loss of information by it. Now, the concept of order not only ceases to be perceived as certainly positive, but even scientific rigor of the term faced serious doubts [8,9]. In our opinion, however, it is more important that the approaches of thermodynamics and information theory are of quite different

scales: the first one is absolutely not interested in the trajectories of individual molecules, while the second one (unlike the approach of hardware technologists) completely ignores the heat generated by the information transmission and processing.

Let us explain this idea with an actual example. A crowd of thousands of American spectators at an international hockey match chants during the game “U-S-A! U-S-A!” We have an extremely ordered, extremely degenerated state with minimal entropy and information. Then, as soon as the period of the hockey game is over, everybody is starting to talk to each other during a break, and a clear slogan is replaced by a muffled roar, in which the “macroscopic” observer finds no meaning. However, for the “microscopic” observer who walks between seats around the arena, each separate conversation makes a lot of sense. If one writes down all of them, it would be a long series of volumes instead of three syllables endlessly repeated during the previous 20 minutes. As a result, chaos replaced order, the system degraded and its entropy sharply increased for the “macroobserver”, but for the “microobserver” (and for the system itself in its entirety), information fantastically increased, and the system passed from an extremely degraded, simple, ordered and poor information state into a much more chaotic, complex, rich and informative one. In other words, everything depends on the point of view. If the observer is at the micro-level and is interested first of all in information about microstates, he will not name it “chaos” or “noise”. If, on the other hand, the observer is located at the macro-level, he as “commander-in-chief” is interested in the total disposition of troops and their fighting spirit and not in relations between the two individual soldiers, their fears and ideals.

Table 1. Examples of high-ordered and low-ordered states.

High-ordered Sates	Low-ordered States
Entropy is low	Entropy is high
Information is low	Information is high
Order	Chaos/Disorder
Simplicity	Complexity
Uniformity	Diversity
Subordination	Freedom
Rigidity	Ductility
Structuring	Randomization
Lattice	Tangle
Predictive	Counterintuitive
Manageability (direct)	Unmanageability/indirect control
Immaturity (for ecosystems)	Climax (for ecosystems)
Agrocenoses	Wild biocenoses
Dictatorship (for society)	Democracy (for society)
Planned economy	Free market
Mono-context	Poly-context
Poor synonym language	Rich synonym language
Artificiality	Naturalness
Degradation	Diversification
Code	Language
Army	Civil society

In addition, estimates of order as something unconditionally positive essentially changed since the nineteenth century. Take a look at Table 1, which contains examples of states with low entropy and information on the left and analogous states with high values of these parameters on the right. It is hard to find such a fanatical fan of “Die Ordnung”, who in all cases prefers the states of the left column.

We would like to summarize this section with a citation from Murray Gell-Mann [10]: “Entropy and information are very closely related. In fact, entropy can be regarded as a measure of ignorance. When it is known only that a system is in a given macrostate, the entropy of the macrostate measures the degree of ignorance the microstate is in by counting the number of bits of additional information needed to specify it, with all of the microstates treated as equally probable.” We would like only to specify that this ignored additional information is not simply needed to describe macrostates with all of their microstates. It is immanent in each of these microstates that we ignore, describing the macrostate.

2.4. Algorithmic Complexity and Its Similarity to Entropy

However, in addition to information and entropy, there is another value closely related to the concept of order and chaos. This is the complexity, often understood very widely [11]. For a long time, it has existed as a qualitative concept, but in the second half of the twentieth century, thanks to the work of Ray Solomonoff [12,13], Andrey Kolmogorov [14] and Gregory Chaitin [15], complexity was strictly defined as a mathematical quantity. All three authors independently came to the definition of algorithmic or descriptive complexity $K(x)$ of the final object x as the length of the shortest computer program that prints a full, but not excessive (*i.e.*, minimal), binary description of x and then halts.

In other words, if D is a set of all possible descriptions x d_x , L the equipotent set of lengths l_x of all of the descriptions d_x and l_{pr} the binary length of the program mentioned above, algorithmic complexity $K(x)$, usually named Kolmogorov complexity, equals to:

$$K(x) = l_{pr} + \text{Min}(l_x) \quad (3)$$

or, if x is not binary, but some other description using n symbols, then:

$$K(x) = l_{pr} + \text{Min}((1/n)\sum p_i \log_2(p_i)) \quad (4)$$

To calculate $K(x)$ for some natural system, you have to describe it, e.g., create its mathematical model, then realize the model as a computer program and, finally, write another program that prints the first program out and then halts. Of course, the length of this second program will be only the first approximation to algorithmic complexity of a given system, because any real model always is not a full description of the system and cannot completely avoid some redundancies. However, creating better and better models, we will approximate step by step the exact value of the complexity. Since the minimum length of the binary description equals, according to the theory of information, the number of pieces of non-redundant information describing the object x in bits, it assumes the relationship between the thus-defined complexity and information or entropy. Indeed, Andrey Kolmogorov himself noted this similarity of algorithmic complexity and information in his paper of 1983 [16]. Later, Thomas M. Cover *et al.* [17] pointed out that the Shannon information and the algorithmic complexity “are pretty much the same” and called this a “beautiful fact”. Then, Seth Lloyd and Heinz Pagels [18] considered

the thermodynamic depth (equal to the entropy up to a constant) as a measure of complexity. Recently, this topic was discussed in a very lively manner [19–25].

Adonai Sant’Anna [23] directly determines the algorithmic complexity of a thermal system as its entropy. Andreia Teixeira *et al.* [24] considers Kolmogorov complexity and Shannon entropy as conceptually different measures, but convincingly proves that for any recursive probability distribution, the expected value of complexity is equal to the entropy up to a constant. Peter Grunwald and Paul Vitanyi [20] point out that $K(x)$ is not a recursive function, and Kolmogorov complexity in general is not calculable in the sense that there is no a computer program that could calculate K in the Equation (3) for any description x , although in simpler and special cases, it may be calculated, while for more complex cases, it could be semi-calculated (*i.e.*, approximately estimated). However, the same holds true for Shannon entropy in the sense that a program for revealing irredundant information from any arbitrary messages does not also exist in general. As a result, the expected Kolmogorov complexity, though not identical, is very close to entropy. James Ladyman *et al.* [21], analyzing the different definitions of complexity, came to the conclusion that the algorithmic definition is the best. They argue that such complexity increases monotonically with increasing disorder, Shannon information entropy and Boltzmann entropy, considering this, however, one of its few shortcomings. Seth Lloyd [18] determines complexity as thermodynamic depth to understanding it, on the one hand, as entropy up to a constant and, on the other hand, as the amount of memory needed to get description x using the shortest program. Finally, David H. Wolpert [25] concludes that the second law of thermodynamics basically determines the increase of complexity over time. In other words, he argues that the complexity and entropy at least have the same sign of the first derivative with respect to time, and the sign is a plus.

Summarizing this overview, we can conclude that, although Equation (4) for complexity is not equivalent completely to Equations (1) and (2), it can be regarded as their generalization in a broader sense. Nonetheless, in relatively simple cases, algorithmic complexity can be considered as approximately equal to Shannon information up to a coefficient and free term l_{pr} . We will demonstrate this by examples in Section 3.4. In more general and complicated cases, such complexity at least grows monotonously with increasing information and entropy.

3. Non-Statistical Approach to the Definition of Generalized Entropy, Using the Theory of Categories and Functors

3.1. Category Theory as a Language for Describing a Wide Variety of Systems

All of the formulas mentioned above were obtained for solving this or that specific problem: the Boltzmann formula gave the molecular-kinetics foundation of thermodynamics, the Shannon formula suggested a quantitative measure of information and Kolmogorov complexity gave algorithmic interpretation of information theory. We in this paper will try to get the same result on the basis of the general consideration using the language of category theory applicable to almost any system.

This language was created in the middle of the twentieth century for describing mathematical structures. The first essential feature of category theory is the consideration not of a separate set with some structure, but of inclusion into the field of view where all of the sets have the same structure (in

other words, sets together with a given axiomatic, for example, with relationships, laws of composition, topologies, and so on). The collection of all such sets is a class of category objects.

Another significant feature of the category theory is explicit consideration not only of the objects themselves, but also of all possible transformations admissible by their structure (for the category of structured sets, this is correspondences from one set to the others not violating their structure). For instance, for sets with a predetermined relation of order, monotonous conformities are admissible, while for sets with a given structure of elements’ “proximity” (topology), the continuous transformations that transform “close” elements into essentially the same “close” ones play the role of admissible transformations, *etc.* The set of admissible transformations (or “morphisms” in the category theory language) is given in the categories for each pair of objects. Therefore, categories with the same objects, but differing morphisms, are, strictly speaking, different and describe different systems. For example, a system with transformations as arbitrary conformities differs from a system where the same set of objects transforms only one-to-one. Processes taking place in the first system are richer than in the latter: the first one allows transitions between states of a variable number of elements, while the second one requires the same number of elements in different states.

Thus, we define an abstract representation of the system as a category that combines a class of objects and a class of morphisms. Objects of the category explicate the system’s states and morphisms: admissible transitions from one state to another (*i.e.*, transformations of states).

3.2. Formulation of the Generalized Entropy for Categories Applicable to Any System Described by These Mathematical Structures

Let us take a system described by category S and the system states X and A , identical to objects X and A from S . Invariant $I_S^X(A)$ is a number of morphisms from X to A in the category S preserving the structure of objects. In the language of systems theory, invariant I is a number of transformations of the state X into the state A , preserving the structure of the system. We interpret the structure of the system as its “macrostate”. Transformations of the state X into the state A will be interpreted as ways of obtaining the state A from state X , or as “microstates”. Then, the invariant of a state is the number of microstates preserving the macrostate of the system, which is consistent with the Boltzmann definition of entropy in Equation (1). More strictly: we determine generalized entropy of the state A of system S (relating to the state X of the same system) as a value:

$$H^X(A) = \ln\left(\frac{I_{\bar{Q}}^X(A)}{I_S^X(A)}\right) \tag{5}$$

where $I_{\bar{Q}}^X(A)$ is the number of morphisms from set X into set A in the category of structured sets Q , and $I_S^X(A)$ is the number of morphisms from set X into set A in the category of structureless sets \bar{Q} with the same cardinality as in category Q , but with an “erased structure”.

As an example, consider the category of sets with partitions, where morphisms are mappings:

$$H_S^X(A) = -\ln\left(\frac{I_S^X(A)}{I_{\bar{S}}^X(A)}\right) = -\ln\left(\frac{\prod_{i=1}^w n_{iA}^{n_{iX}}}{n_A^{n_X}}\right) = -n_X \left(\sum_{i=1}^w \frac{n_{iX}}{n_X} \ln \frac{n_{iA}}{n_A}\right) \tag{6}$$

where n_{iX} and n_{iA} are the numbers of elements in classes of partitions of sets X and A , $n_X = \sum_{i=1}^w n_{iX}$, $n_A = \sum_{i=1}^w n_{iA}$, and w the number of classes in partitions. This example demonstrates that, in particular cases, generalized entropy has the usual “Boltzmann” or, if you like, “Shannon” look.

Invariants of many (if not all) mathematical structures may be expressed by invariants of partitions associated with structures. Invariants of sets with partitions are multiplicative, relating invariants of each class of partitions [26,27]. Therefore, logarithms of the invariants have a typical “entropy-like” look of sums by classes of partitions.

Value $\frac{I_Q^X(A)}{I_Q^X(A)}$ is a ratio of the number of transformations preserving the structure by the total number of transformations that can be interpreted as the probability of the formation of the state with a given structure. In addition, according to Equation (5), value $p_{iA} = \frac{n_{iA}}{\sum_{i=1}^w n_{iA}}$ can be interpreted as the probability of realization class i in respective partition and value $\{p_{1A}, p_{2A}, \dots, p_{wA}\}$ as the probability distribution that determines state A . If we understand p_{iA} as the probabilities of equally-probable Boltzmann microstates, Equation (5) can be interpreted up to a constant as Boltzmann Equation (1). If, on the other hand, implementations p_{iA} are not equally probable and correspond to probabilities of symbol i in message A , Equation (5) takes the form of Shannon information (Equation (2)). Finally, if the message corresponding to $\{p_{1A}, p_{2A}, \dots, p_{wA}\}$ is a minimal description of system x , generalized entropy can be considered as a value at least monotonously related, if not proportional, up to a free term (that is, the binary length of a computer program that prints out this description) to algorithmic complexity (Equation (4)).

However, statistical entropy, information and algorithmic complexity are only a few possible interpretations of Equation (5). It is important to emphasize that the formula for the generalized entropy is introduced with no statistical or probabilistic assumptions and is valid for any large or small amounts of elements of the system.

3.3. Competition for Resources and Metabolic Time

There is another set of numeric characteristics of the system which: (1) always accompanies the variability of the system; (2) increases monotonically relating to generalized entropy and, hence, complexity; and thus, (3) may play a role in change parameterization. These are resources that the system consumes. The amount of “consumed” (plus “lost”) resources determines “reading” of the so-called “metabolic clock” of the system [28]. Construction of this metabolic clock implies the ability to count the number of elements replaced in the system. However, the procedure of comparing sets by the number of elements correctly designed in set theory is only for structureless sets, while all of theoretical science uses sets with a structure for any formal descriptions. In particular, systems with multiple resource flows cannot be represented as a hierarchy of structureless sets. Therefore, a non-trivial application of the metabolic approach requires the ability to compare one structured set to another. This ability comes from a functorial method comparison of structures that offers system invariants as generalization of the concept “number of elements” for structureless sets. Thereby, functorial parameterization is a necessary generalization of the metabolic clock. Note that the system that consumes several resources exists in several metabolic times.

Theorems of variational modeling allow us to establish the relationship between entropy and metabolic system times. The entropy of the system is an “averager” of metabolic times, and entropy increases monotonically with the flow of each of metabolic time, *i.e.*, entropy and metabolic times of a system are linked uniquely, monotonously and can be calculated one through the other. This relationship is given by:

$$H(\mathbf{L}) = \sum_{k=1}^m \lambda^k(\mathbf{L}) L^k \tag{7}$$

Here, H is structural entropy, $\mathbf{L} \equiv \{L^1, L^2, \dots, L^m\}$ the set of metabolic times (resources) of system and $\lambda^k(\mathbf{L})$ Lagrange multipliers of the variational problem on the conditional maximum of structural entropy, restricted by flows of metabolic times L^k . For the structure of sets with partitions where morphisms are preserving the partition mapping (or their dual compliances), the variational problem has the form:

$$\begin{cases} H(n_1, n_2, \dots, n_w) = -n \sum_{i=1}^w \frac{n_i}{n} \log \frac{n_i}{n} \longrightarrow \max, \\ \sum_{i=1}^w q_i^k n_i \leq L^k, k = \overline{1, m}, \\ \sum_{i=1}^w n_i = n, \\ n_i \geq 0, i = \overline{1, w}. \end{cases} \tag{8}$$

Here, n_i is the number of elements in partition class i , w the number of partition classes and q_i^k the average “substance content” of metabolic flow k in elements of class i .

It was proven [29] that $\frac{\partial H}{\partial L^k} \geq 0$, *i.e.*, structural entropy monotonously increases (or at least does not decrease) in the metabolic time of the system or entropy “production” does not decrease along a system’s trajectory in its state space (the theorem is analogous to the Boltzmann H-theorem for physical time).

Such a relationship between generalized entropy and resources L^k can be considered as a heuristic explanation of the origin of the logarithm in dependence of entropy on the number of transformations: with logarithms the relationship between entropy and metabolic times becoming a power, not exponential, which in turn simplifies the formulas, which involve both parameterizations of time. Therefore, if the system metabolic time is, generally speaking, a multi-component magnitude and level-specific (relating to hierarchical levels of the system), then entropy time “averaging” metabolic times of the levels parameterizes system dynamics and returns the notion of the time to its usual universality.

3.4. Extremal Principle as a Law of System Variability, Its Entropy and Information Interpretation and the Irreversibility of Time Generated by It

The class of objects that explicates a system of categories can be presented as a system’s state space. The main component of any dynamical theory is a law of variation of the studied system, in other words, a set of rules that allows one to select from all possible states the ones realized in reality and to indicate the “trajectory” of the system in its state space. In mechanics and field theory, such a law has the form “equations of motion”, which are postulates of the theory. An alternative to the postulation of the equations of motion in theoretical physics, biology, economy and other sciences is the postulation of extremal principles that generate variability laws of the systems studied.

What needs to be extreme in a system? The category-functorial description gives “natural” answer to this question, because category theory has a systematical method to compare system states [27,28]. The possibility to compare the states by the strength of their structure allows one to offer an extremal principle for systems’ variation: from a given state, the system goes into a state having the strongest structure.

However, “strength of structure” regulates the state of a system only partially: some of them can appear incomparable. This complicates the application of the proposed extremal principle, since it allows a situation in which the extremal principle does not admit tracing the trajectory of the system in its state space. However, a functorial method of structure comparison [27,28] allows us to overcome these complications with comparison of states by the strength of the structure. This method establishes a linear, rather than a partial, order on a set of equally structured sets or system states. The functorial method of structure comparison not only claims that any of the systems are comparable, but also strictly introduces a numerical function of the state, which varies from state to state monotonically to the strength of their structure. According to the method, this function is the number of transformations admissible by structure of the system. However, a more usual formulation of the extremal principle can be obtained if we consider the monotonic function of the specific amount of admissible transformations that we defined as the generalized entropy of the state; namely given that the state of the system goes into a state for which the generalized entropy is maximal within the limits set by available resources. In the simulation of dynamic systems with the help, for example, of the Jaynes formalism, one uses the formula for information entropy as a postulate of the objective function. A generalized category-theoretic entropy allows not guessing and not postulating the objective functions, but strictly calculating them from the number of morphisms (transformations) allowed by the system structure. This ability to display the extremized value significantly expands the possibilities of successful systems modeling. A category-theoretic approach to modeling has been very useful in the ecology of communities [28,29].

Let us illustrate this with an example. Consider a very simple system consisting of a discrete space of 8×8 (like a chess board without dividing the fields on the black and white) and eight identical objects distributed arbitrary on these 64 elements of the space (cells). These objects can move freely from cell to cell, realizing two degrees of freedom each. The number of degrees of freedom of the system is twice as much as the number of objects due to the two-dimensionality of our space.

We will consider the particular distribution of eight objects on 64 elements of our space (cells) as a system state that is equivalent in this case to a “microstate”. Thus, the number of possible states equals the number of combinations of eight objects from 64 ones:

$$W_8 = 64!/(64 - 8)!/8! = 4,426,165,368$$

Consider now more specific states when seven objects have arbitrary positions, while the position of the eighth one is completely determined by the positions of one, a few or all of the others. In this case, the number of degrees of freedom will reduce from 16 (eight by two) to 14 (seven by two), and the number of admissible states will decrease up to the number of combinations by seven objects, seven from 64 ones:

$$W_7 = 64!/(64 - 7)!/7! = 621,216,192$$

Let us name a set of these states a “macrostate”. Notice, however, that the number of combinations by k elements from n , calculated by the formula $\frac{n!}{k!(n-k)!}$, is the cumulative number of “microstates”

for “macrostates” with 16, 14, 12, and so on, degrees of freedom (Table 2). Therefore, to reveal the number of “microstates” related exclusively to a given “macrostate”, we have to subtract W_7 from W_8 , W_6 from W_7 , etc. Values of W corrected in such way are presented in the third column of Table 2. Even from a quick glance at this column, it is quite clear how much the numbers of “microstates” differ for this or that “macrostate”. For instance, the system in the “macrostate” with eight degrees of freedom is 176.28-times more likely to go into a state of higher (10) than lower (6) degrees of freedom and the number of “microstates”. Similar ratios of chances for other “macrostates” are listed in the fourth column of Table 2.

These figures make quite clear that our simple model system being left to itself will inevitably move into a “macrostate” with more degrees of freedom and a larger number of admissible states, i.e., “microstates”. In parallel with this, statistical and generalized entropies will grow, too. Both of these entropies equal one another in our case due to the equal probabilities of “microstates” that leads to the reduction of Equation (6) to Boltzmann entropy $\ln W$. The information under these conditions will be calculated by the same formula, but with a binary logarithm: $\log_2 W$. Then, complexity, as a minimal complete description of the system, can be calculated as $N \cdot \log_2(64) + P$, where $\log_2(64)$ is the description of an object position in bits, N the number of objects and P the length of a program that prints the description. Let us set $P = 10$ bits for definiteness.

Table 2. Degrees of freedom, thermodynamic probability, entropy, information, length of minimal description and complexity for a discrete 8×8 space with 8 objects.

Degrees of Freedom	W (cumulative)	W	$\frac{W_{+1}}{W_{-1}}$	Entropy $S = \ln W$	Information $I = \log_2 W$	Length of Description (D)	Complexity (D + Program)
2	64	64		4.16	6.00	6	16
4	2016	1952	619.50	7.58	10.93	12	22
6	41,664	39,648	304.16	10.59	15.27	18	28
8	635,376	593,712	176.28	13.29	19.18	24	34
10	7,624,512	6,989,136	113.44	15.76	22.74	30	40
12	74,974,368	67,349,856	78.16	18.03	26.01	36	46
14	621,216,192	546,241,824	56.50	20.12	29.02	42	52
16	4,426,165,368	3,804,949,176		22.06	31.83	48	58

Values of entropy, information, description and complexity calculated in such way are presented in Columns 5–8 of Table 2. As we can see in Figure 1, entropy and algorithmic complexity are connected almost linearly, and although the length of the program that prints the description was chosen quite arbitrarily, it does not influence the revealed dependency. Moreover, all four values monotonously and almost proportionally grow with increasing degrees of freedom (in Figure 2). Similar data and their dynamics for a space of 16×16 with 256 cells and 16 objects are shown in Table 3 and Figures 3 and 4. As you can see, this example demonstrates all of the identified trends and chances to move into a state with greater entropy and complexity; there is nearly an order of magnitude more than in the case of 8×8 . As for the chances of the two such consecutive transitions, they are of the order of a million. Two obvious conclusions immediately follow from these considerations:

- It is far more probable to find a system in a complex state than in a simple one.

- If a system came to a simple state, the probability that the next state will be simpler is immeasurably less than the probability that the next state will be more complicated.

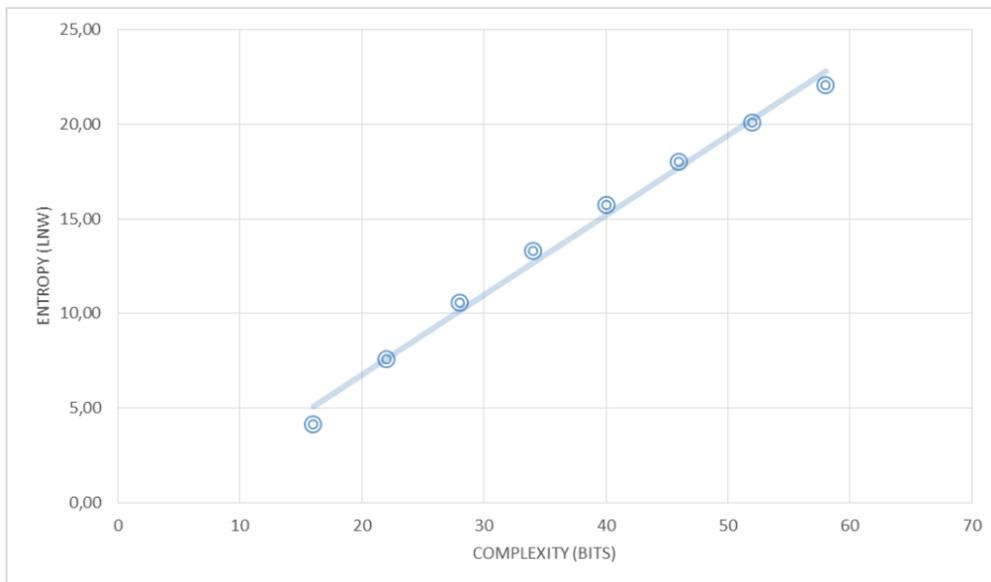


Figure 1. Relation between entropy and complexity for a discrete 8×8 space.

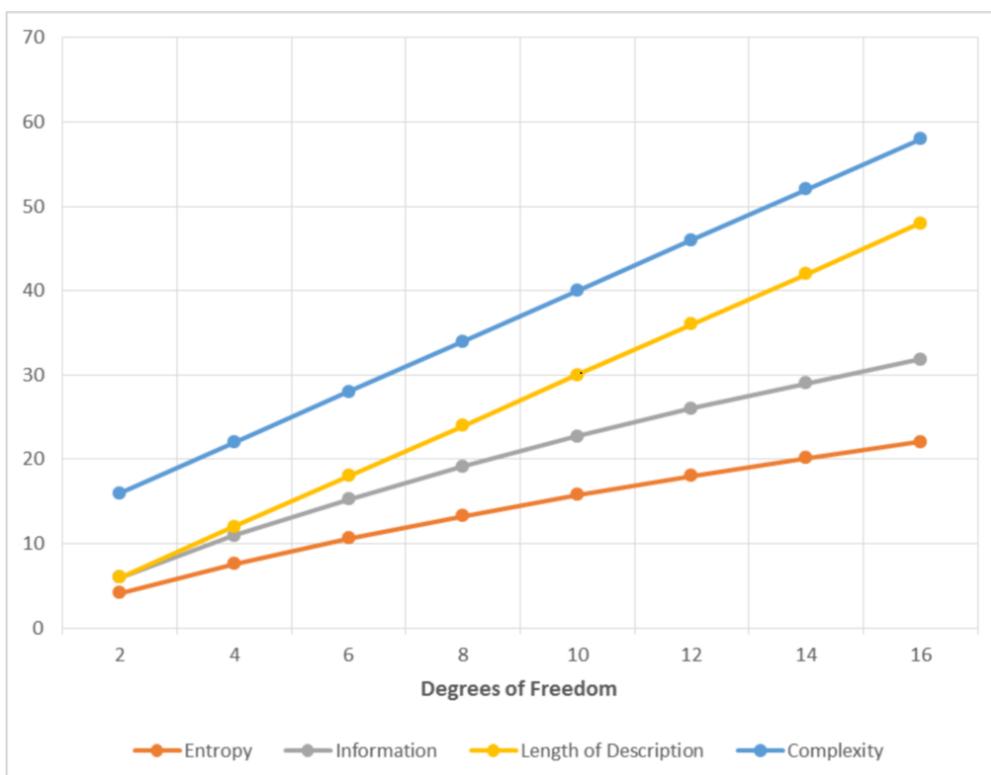


Figure 2. Dependence of entropy, information, length of description and complexity on the degrees of freedom for a discrete 8×8 space.

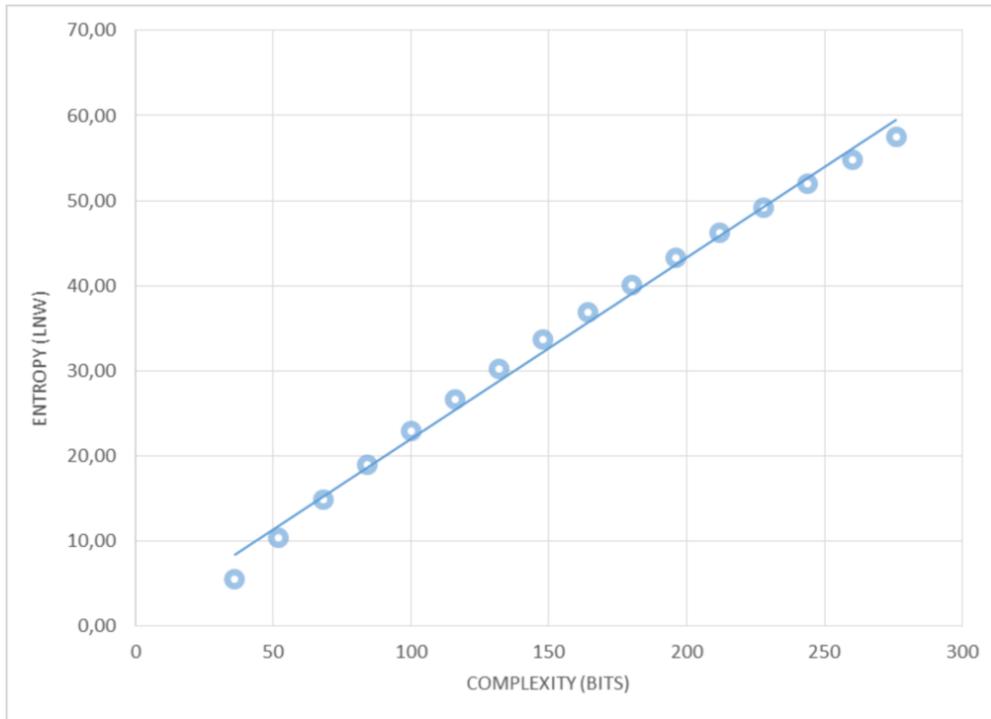


Figure 3. Relation between entropy and complexity for a discrete 16×16 space.

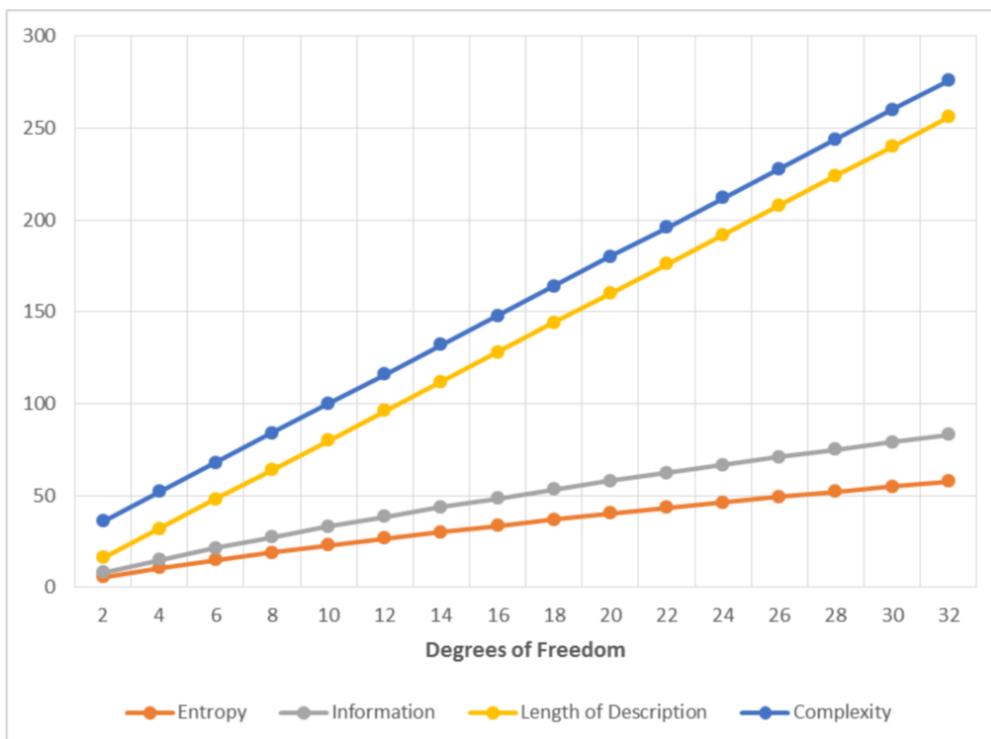


Figure 4. Dependence of entropy, information, length of description and complexity on the degrees of freedom for a discrete 16×16 space.

Table 3. Degrees of freedom, thermodynamic probability, entropy, information, length of minimal description and complexity for a discrete 16×16 space with 16 objects.

Degrees of Freedom	W (cumulative)	W	$\frac{W+1}{W-1}$	Entropy $S = \ln W$	Information $I = \log_2 W$	Length of Description (D)	Complexity (D + Program)
2	256	256		5.55	8.00	16	36
4	32,640	32,384	10,667.50	10.39	14.98	32	52
6	2,763,520	2,730,880	5,312.16	14.82	21.38	48	68
8	174,792,640	172,029,120	3,161.90	18.96	27.36	64	84
10	8,809,549,056	8,634,756,416	2,091.06	22.88	33.01	80	100
12	368,532,802,176	359,723,253,120	1,481.61	26.61	38.39	96	116
14	13,161,885,792,000	12,793,352,989,824	1,102.24	30.18	43.54	112	132
16	409,663,695,276,000	396,501,809,484,000	850.35	33.61	48.49	128	148
18	11,288,510,714,272,000	10,878,847,018,996,000	674.75	36.93	53.27	144	164
20	278,826,214,642,518,000	267,537,703,928,246,000	547.55	40.13	57.89	160	180
22	6,235,568,072,914,500,000	5,956,741,858,271,980,000	452.55	43.23	62.37	176	196
24	127,309,514,822,004,000,000	121,073,946,749,089,000,000	379.77	46.24	66.71	192	212
26	2,389,501,662,813,000,000,000	2,262,192,147,991,000,000,000	322.82	49.17	70.94	208	228
28	41,474,921,718,825,700,000,000	39,085,420,056,012,700,000,000	277.45	52.02	75.05	224	244
30	669,128,737,063,722,000,000,000	627,653,815,344,896,000,000,000	240.75	54.80	79.05	240	260
32	10,078,751,602,022,300,000,000,000	9,409,622,864,958,580,000,000,000		57.50	82.96	256	276

This defines a practically irreversible increase of entropy, information and complexity, leading in turn to the irreversibility of time. For space 16×16 , we could speak about practical irreversibility only, when reversibility is possible, although very improbable, but for real molecular systems where the number of cells is commensurate with the Avogadro’s number (6.02×10^{23}), irreversibility becomes practically absolute. After all, the chances of reversibility become so incredibly small that millions and millions of times of the universe’s existence is not enough to carry out at least one spontaneous transition accompanied by a decrease in the number of degrees of freedom and complexity, which, among other things, supports the idea of Stuart Kauffman [30] about the nonergodicity of the universe.

This absolute irreversibility leads to the absoluteness of the entropy extremal principle, which, as shown above, can be interpreted as an information or as complexity one. This extremal principle implies a monotonic increase of state entropy along the trajectory of the system variation (sequence of its states). Thus, the entropy values parametrize the system changes. In other words, the system’s entropy time does appear [31]. The interval of entropy time (*i.e.*, the increment of entropy) is the logarithm of the value that shows how many times the number of non-equivalent transformations admissible by the structure of the system have changed.

The design of structural and entropy time includes a number of randomly selected formal steps. For instance, a comparison of structured sets that describe the system states is made using injective transformations (by analogy to the comparison of unstructured sets by cardinality). As a representation from the category of structured sets into the category of unstructured ones that generates entropy time, a basic single functor was selected in the functorial method for comparing the structures. These functors appeared to be monotonous to streamline structures, and their values were reducible to entropy-like functions. Injective transformations ordering the structure are unambiguous nesting. In other words, the

evolution of systems, according to the extremal principle, flows from sub-objects to objects. We will discuss this in the next section in more detail. Here, we just note that, in the real world, where the system is limited by the resources, a formalism corresponding to the extremal principle is a variation problem on the conditional, rather than global, extremum of the objective function [26,28]. This type of evolution could be named conservative or causal: the achieved states are not lost (the sub-object “is saved” in the object like some mutations of Archean prokaryotes are saved in our genomes), and the new states occur not in a vacuum, but from their “weaker” (in the sense of ordering by the strength of structure) predecessors.

Therefore, the irreversible flow of entropy time determines the “arrow of time” [32] as a monotonic increase of entropy, information, complexity and freedom as the number of its realized degrees up to the extremum (maximum) defined by resources in the broadest sense and especially by the size of the system. On the other hand, available system resources that define a sequence of states could be considered as resource time that, together with entropy time, explicates the system’s variability as its internal system time.

4. Law Defining the Arrow of Time and Its Direction

4.1. General Law of Complication

The extremal principle in thermodynamics is embodied in the form of its second law. Here, entropy plays the role of the extremal function, and its extremum is maximum. In the previous section, we formulated and proved a far more general extremal principle applicable to any dynamic system (*i.e.*, described by categories with morphisms), including isolated, closed, opened, material, informational, semantic, *etc.*, ones (rare exceptions are static systems without morphisms, hence without dynamics described exceptionally by sets, for example a perfect crystal in a vacuum, a memory chip with a database backup copy or any system at a temperature of absolute zero). The extremum of this general principle is maximum, too, while the extremal function can be regarded as either generalized entropy, or generalized information, or algorithmic complexity.

Therefore, before the formulation of the law related to our general extremal principle, it is necessary to determine the extremal function itself. It seems the most natural to choose for such a role the generalized entropy as it was defined above. Especially taking into account the advice that John von Neumann gave Claude Shannon: “No one knows what entropy really is, so in a debate you will always have the advantage” (cited by [33]). However, jokes aside, it is hard to agree that the term “entropy” could be considered very successful. Its author Rudolf Clausius wrote (cited by [34]): “I propose, accordingly, to call S the entropy of a body, after the Greek word ‘transformation’”. However, as Arieh Ben-Naim [9] fairly remarked, “Entropy, as it is now recognized, does not mean ‘transformation’, or ‘change’, or ‘turn’. It does mean information”. He strongly recommends to opt out of this clearly poor term and replace it everywhere with the term “information”.

However, “information” allows as much if not more different interpretations than “entropy”. The meanings of “information” in communication theory, cybernetics, computer science, semantics, cryptology, *etc.*, though overlapping, are far from identical. Unfortunately, there is not yet a common

definition of information that includes all of these interpretations. Therefore, the use of information without such a common definition as an extremal function, in such a general extremal principle, would inevitably lead to confusion in its interpretation within each particular application.

In contrast to the first two concepts, algorithmic complexity is defined quite strictly and unambiguously. There is still a problem with the fact that the complexity, in general, is not a computable function in the sense that there is no program that calculates the value $K(x)$ for any binary sequence x [35]. However, it is not an essential challenge for the extremal function, given that the complexity at least increases monotonically with increasing entropy and information. At the same time, although algorithmic complexity is determined unambiguously, complexity in a wide sense is interpreted in various ways. In particular, some interpretations consider states of maximum entropy and randomization (chaos) as simple, and the maximum complexity in this approach lies somewhere between order and chaos [36,37]. This is, of course, completely contrary to the general extremal principle. Therefore, it must be emphasized that by extremal function, we mean just algorithmic complexity.

In summary, our generalized extremal principle is the following: the algorithmic complexity of the dynamical system, either being conservative or dissipative, described by categories with morphisms, monotonically and irreversibly increases, tending to a maximum determined by external conditions. Accordingly, the new law, which is a natural generalization of the second law of thermodynamics for any dynamic system described by categories, can be called the general law of complication. We formulate it as follows:

Any natural process in a dynamic system leads to an irreversible and inevitable increase in its algorithmic complexity, together with an increase in its generalized entropy and information.

4.2. Place of General Law of Complication among the Other Laws of Nature

We have already noted that the second law of thermodynamics stands out among the rest of the physical laws. It was the first and almost the only law that is asymmetric with respect to time. Because the general law of complication is a generalization of the second law of thermodynamics, it is natural to expect that it, too, is opposed in this respect to the other laws, clearly distinguishing the past, where the complexity and the generalized entropy/information are less than they are now, and the future, where these values will inevitably be higher. Lately, however, thanks to Seth Lloyd's doctorate thesis [38], forgotten for 25 years, there were serious suspicions that quantum uncertainty leading to quantum entanglement is also responsible for the arrow of time [39]. If this is confirmed, it will be yet another regularity of this kind.

The other obvious difference between the new law, as well as the second law of thermodynamics and all of the other laws of nature is the statistical character of the first two. Typically, laws of nature have an absolute character, predicting effects of this or that cause exactly rather than approximately. You could say that in quantum mechanics, no one can determine exactly and simultaneously the position and velocity of an electron or any other quantum object due to Heisenberg's uncertainty principle [40], and this is true. However, the probability density of the wave function, which is included in the Schrodinger equation, is defined quite precisely and unambiguously. On the contrary, the general law of complication, as we could see from the example analyzed in Section 3.4, determines only the general

direction of system evolution, while each particular transition from one state to another can occur with as increasing or decreasing complexity. The chances for complication, however, are more than for simplification. Moreover, the ratio of chances to transit towards complexity increases with the size of the system, *i.e.*, with number of elements (objects) in it. The statistical nature of the general law of complication, as well as of the second law of thermodynamics leads to inapplicability (in a strict sense) of the well-known Karl Popper falsifiability principle to any of these laws. If we, for instance, found a temperature gradient in an isolated gas volume increasing for a while, it could be just a rare event, rather than the second law of thermodynamics falsification. However, a long enough series of wide-scale experiments could really falsify these and any other statistical laws. Particularly, the general law of complication could be falsified by, for example, evidence that many of galaxies at a distance of billions of light-years (and respectively billions of years younger) have a more complicated structure than our Milky Way or the Andromeda nebula. In addition, Vladimir Ulyanov (Lenin), implementing communist principles, started a wide-scale experiment that should prove that the complicated economics of the free market could be effectively replaced with far more simple and straightforward economics of direct government planning. After 70+ years of revolutions, civil and international wars, repressions and dictatorships in different regions of the world, the experiment demonstrated the complete inconsistency of this idea. Therefore, the fate of communism did not falsify the general law of complication, although, of course, it did not prove it either.

However, the peculiar qualities of the suggested new law are not limited to these two; but to move to the third and the main one, we need to digress to the area of systems' hierarchy, the relationship between the levels of such hierarchy and the forces acting between these levels.

4.3. Hierarchical Forces and the Third Unique Feature of the General Law of Complication

The vast majority of forces considered by physics and other scientific disciplines could be determined as horizontal or lateral ones in a hierarchical sense. They act inside a particular level of hierarchy: for instance, quantum mechanics at the micro-level, Newton's laws at the macro-level and relativity theory at the mega-level. The only obvious exception is thermodynamic forces when the movement of molecules at the micro-level (or at the meso-level if we consider the quantum mechanical one as the micro-level) determines the values of such thermodynamic parameters as temperature, entropy, enthalpy, heat capacity, *etc.*, at the macro-level of the hierarchy. One could name these forces bottom-up hierarchical forces.

We could see the other examples of such bottom-up forces in biology, where the proliferation of cells at the lower level led to the emergence of the biosphere of our planet as a global object. In human society, the activity of very small social groups and even individuals (lower level) can serve as a trigger of revolutionary processes that are radically changing the lives of peoples and states. Although all of these forces are not beyond the interests of modern science, the theory to describe them as comparable with thermodynamics does not exist yet.

In addition to lateral and bottom-up forces, ones that act in the opposite direction exist, too (about 30 years ago, we named them omnicausal forces, emphasizing that their cause is at the higher level of the hierarchy, the level of the whole [41,42]; now, however, such causality and related forces are usually

called top-down [43,44], so we will stick with the common terminology). Each of us feels their existence in his or her own life, and their impact is not always pleasant and positive.

They can be either permanent or sporadic. Examples of the former ones can be gravity, daily and seasonal cycles, inhibition of unlimited cell division in the body of multicellular organisms, the laws and the government in human society, *etc.* Sporadic top-down forces act rarely and look like disasters. Typically, their effect is unpredictable, but sometimes, science reveals laws that govern these forces at the upper level of the hierarchy. We can predict such macro-events as solar and lunar eclipses for thousands of years ahead (because of the extreme simplicity of the Sun-Earth-Moon system), a collision of asteroids and comets with Earth for decades and tsunamis for hours. However, we cannot yet confidently predict earthquakes, volcanic eruptions, climate changes, *etc.*

Considering from this point of view our general law of complication, it is clear that it, as well as the second law of thermodynamics that it generalizes, deals with not lateral, but hierarchical forces and, first of all, bottom-up ones. We have already seen, in Section 3, how the increase in the number of degrees of freedom implemented at the lower level of the hierarchy leads to the increase in complexity and, respectively, generalized entropy and information at the top level. Connection with top-down forces is less clear, because it is implemented by including three levels. If the two-level system is an element of an even higher mega-level, then the more complex this super-system, and, respectively, the more complex the role and behavior of our system as its element, the more often and unpredictable its macrostate changes. Hence, the top-down forces caused by them will more often and unpredictably impact the elements of the system at the micro-level, both constantly (like free energy flux) and sporadically (like catastrophes, which we will discuss in Section 5).

Thus, the third difference between the general law of complications and the other laws of nature (with the exception of the second law of thermodynamics) is its close connection with hierarchical rather than lateral forces. On the one hand, because the time scale, or if you wish, its flow, at different levels of the hierarchy in the real world varies by orders of magnitude, the structure of time moments (the structure of the present) [41,45] on the upper level leads to the irreversibility on its lower level. On the other hand, the reversibility at the lower level, in conditions of low complexity, leads to irreversibility on the top one (Boltzmann's H-theorem). In both cases, one of the consequences of the irreversible complication is the emergence of Eddington's arrow of time. From this, in turn, follows the fourth and the main difference between our general law of complications and all of the others. This is their fundamentally different roles in the overall evolution of the universe, which will be discussed in the next and last section of this article. Before this, however, we briefly discuss two aspects of the world around us and the specifics of our new law in their interaction.

4.4. Two Worlds: Large, Complex and Irreversible and Small, Simple and Timeless

If we look around, we will see a world that is very large (about 10^{13} light-years), very complicated (except perhaps the planetary systems and the majority of the artifacts created by man) and evolving over time in all time scales. At first glance, this vision includes all of the material world in its whole diversity. Meanwhile, there is another world, a simple one, existing outside of time, and so small that, according

to Seth Lloyd [22], it can be placed on the back of an ordinary T-shirt. Moreover, it is this small world that defines all of the parameters, behavior and evolution of the large one.

We have in mind the world of the laws of nature. Though their origin is quite vague, their undeniable role and effectiveness have been confirmed by the very existence of science and technology based on them. As far as we can judge from astronomical data, these laws did not change over the lifetime of the universe, or at least for about 13 billion years, after cosmic microwave background appeared. They are very simple in the sense that they can be described very compactly (think of Lloyd's T-shirt) if, of course, there is a context that allows us to understand them. However, since the nineteenth century, because flourishing science demonstrated the quite obvious effectiveness of new laws of nature, their world has begun to be discussed more and more often [46–48]. In the article [46], written only two years ago, Michael Shulman connects this small world with a kind of “metareality” that includes some special “metatime” and “metaspace”. We cannot agree that there is enough evidence, in both science and philosophy, of the existence of such “metareality”. We rather consider laws of nature as a frame, or if you like, “boundary conditions” of the only reality, but his arguments for the existence of the second world look pretty weighty and persuasive. In any case, this discussion does not relate to the core topic of this article.

These two worlds not only exist side by side, but also interact closely, albeit asymmetrically. The world of laws dictates the acceptable behavior of the world of nature, while the world of nature fills the laws out with the specific content and allows them to realize themselves as real forces and processes.

The general law of complication undoubtedly belongs to the world of laws, but existing, as well as all other laws outside of time, belongs to the few of them that parameterize temporal and, hence, evolutionary aspects of the large material world. At the same time, the other and bigger part of the laws of nature, which are symmetric in time, play a special and no less important role in the evolution of our universe.

5. Evolution of the Universe, Its Engine and Its Driver

5.1. Diversity and Selection in Physical-Chemical Evolution and Laws of Nature as “Breeder”

Evolution as a scientific term appeared in biology in the nineteenth century thanks to Jean-Baptiste Lamarck and then Charles Darwin. Darwin also proposed the mechanism of biological evolution as the combination of variation, inheritance and natural selection, based on competition. Physics meanwhile believed the universe to be infinite in space and time and, therefore, unchanged as a whole. The only exception was the concept of “heat death”, formulated by Lord Kelvin (William Thomson) [49] and based on the second law of thermodynamics. According to this concept, the entropy of the universe as an isolated system increases, which will eventually lead to the complete dissipation of free energy, alignment of all of the gradients and “heat death” as the cessation of any activity. However, this concept was perceived rather controversially and did not belong to the mainstream of physical ideas.

Only in the middle of the twentieth century, after the formation of ideas on the Big Bang, which occurred about 13 billion years ago, and the subsequent expansion of the universe, did the concept of evolution of the universe become firmly established in physics. Then, when Arno Penzias and

Robert Wilson [50] revealed the cosmic microwave background, experimental data have allowed tracing back the history of our universe up to 377,000 years after the Big Bang. In this period, called the recombination, matter cooled down to such an extent that the formation of atoms from the soup of protons, neutrons and electrons started, and the universe became transparent, which allows us now to observe this radiation itself. Soon enough, it became clear that the cosmic microwave background is isotropic to within 0.001% [51]. In other words, the universe at that time did not have clearly-defined structure or ordering.

At first glance, it suggests the near to maximum entropy and complexity of the universe as a whole. This would be fair not taking into account gravitational forces and the high temperature at that time. As was noted by Isaac Newton [52], the homogeneous state of gravitational systems may not be stable, because the slightest perturbation of density under the influence of gravitational forces will continue to grow; therefore, a completely homogeneous and very hot state of matter at that time; this is, on the contrary, the state with the minimum entropy and maximum free energy [53,54]. Then, further development (evolution) of the universe has gone, and will go, in full accordance with the general law of complication, in the direction of increasing complexity, entropy, information and the number of degrees of freedom. Due to density perturbations and under the influence of gravitational forces, protogalaxies and protostars emerged. Then, gravity squeezed and warmed up the protostars so much that thermonuclear reactions ignited in them, and the stars “lit up”. Stars illuminated planetary systems formed around some of them, and there, in turn, chemical evolution began.

All of these irreversible processes of inorganic evolution defined the physical arrow of time developed due to the general law of complication, according to which the complexity and, consequently, the information of the universe irreversibly increased. As a result, the diversity of its elements determined by the same Shannon formula (2) increased, too. According to Seth Lloyd [55], the amount of information created and processed in a such way since the Big Bang is 10^{92} bits. However, what determined which variants of this diversity were stable and participated in the further evolution and which ones could not exist, were destroyed and eliminated? In other words, what is the mechanism of physical and chemical evolution?

We take the liberty to say that this mechanism is akin to natural selection in biological evolution, but much simpler and more primitive. Such selection rather rejects “bad” variants than selecting “good” ones. Material for it comes from the increasing variety because of complication, and selection itself is realized by the small, timeless world mentioned above, that is by the laws of nature. For example, if one proton meets one electron at a certain combination of the magnitude and direction of their velocities, they will be able to unite in a hydrogen atom, and if one electron meets the helium nucleus, no atom obviously will be formed, since the formation of a helium atom needs two electrons, and with opposite spins. The fact that the hydrogen atom is stable with one electron and a helium atom with two is determined by nothing but the laws of nature, in this case, the laws of quantum mechanics. Similarly, the collision of atoms or molecules may lead to new chemical bonds or may end up with nothing. This is determined, within the given collision parameters, exclusively by the laws of quantum chemistry, based on the same quantum mechanics. Gradually, this process led to the emergence of atoms, molecules, stars, galaxies, planets, minerals, crystals, *etc.*

In biology, the closest to this type of selection is stabilizing selection [56], which discards all variants that differ from the norm most adapted to current environmental conditions. It works effectively when these conditions do not change for quite some time. In our case, the laws of nature are immutable absolutely, existing outside of time. Accordingly, stabilizing selection is the most effective in the physical evolution. This highlights a similarity between physical and biological evolution at their very base level, which was already demonstrated by Leonid Martyushev [57]. In more complex cases of the chemical or, in fact, already prebiotic evolution, when two or more chemical reactions are going in parallel and competing for the same resource in the form of a certain substance or energy, selecting between them may even acquire the features of a competitive selection, a characteristic for biological evolution, where it, as we will see below, plays a decisive role. Thus, the general law of complication, leading to an increase in diversity and, therefore, accumulation of material for selection, plays the role of the engine of evolution; while selection of “viable” stable variants from all of this diversity is a kind of driver of evolution that determines its specific direction. The role of a “breeder” of this selection plays other, usually less general, laws of nature, which have remained unchanged since the Big Bang.

5.2. Evolution in a Steady Stream of Free Energy, Dissipative Structures and Their Selection

The situation changed significantly after the stars lit, and around many of them, planetary systems were formed. Each of these planets existed in a constant influx of free energy in the form of radiation from the star around which it rotated. Moreover, this influx is amazingly constant through billions of years. This happens because typical main-sequence stars do not essentially change their entropy production and, respectively, energy dissipation during such periods of time [58]. Free energy, being a value opposite of entropy, and, therefore, the information and complexity are essentially information killers. This is particularly evident in the form of a single impact of free energy, which can be regarded as natural disasters, which we discuss below. In the case of a constant steady stream, the free energy just does not allow entropy, information and complexity to achieve their maximum due to the ordering in the form of the appearance of dissipative structures [59]. Such structures appear in dissipative systems that are a subset of dynamic systems. Dissipative systems are open ones that are operating out of, and often far from, thermodynamic equilibrium in an environment with which they exchange energy and matter.

Respectively, dissipative systems and their structures cannot exist without a constant supply of free energy, but on a planet close enough to its star, that was not a problem. Their appearance drastically enriched the evolution and accelerated its flow. The structures that previously were eliminated because of their instability, now thanks to the constant influx of free energy, began to exist for a long time. Such existence, however, needs entropy production as a result of free energy dissipation, and the more dissipative structures became complicated, the more non-equilibrium and non-steady they were and the more entropy production needed. As a result, the maximum entropy production principle formulated by Hans Ziegler [60] in 1963 is a natural consequence of the general law of complication. Ilya Prigogine [61], on the other hand, formulated in 1945 a minimum entropy production principle, but as Leonid Martyushev [57] clearly proved, it is just a special case of Ziegler’s principle realized only for steady states not far from equilibrium. Therefore, the maximum entropy production principle is more general than Prigogine’s one, and it determines the particularity of the implementation the general law

of complication for dissipative structures. If in a physicochemical world with only stable structures, the first derivative of entropy with respect to time is positive, but either approximately constant or sporadically varies in time, after the appearance of dissipative structures, this first derivative itself began to increase irreversibly. This respectively led to the accelerated increase of information, complexity and degrees of freedom. At the same time, the competition for this free energy developed, resulting in the appearance of a competitive component of natural selection, which is specific for biological evolution. In addition, because these dissipative structures were able to self-organize [59,62,63], that is probably another consequence of the maximum entropy production principle, the growth of complexity, information, diversity and respectively variations for natural selection increased dramatically. All of this as a whole created preconditions for further prebiotic physicochemical evolution, which at some point started the primitive biological one [64]. However, even in these cases, the general law of complication played the role of the engine that moves evolution by delivering increasingly rich material for it, while selection the role of the driver that directs the evolutionary process towards, at first, admissible structures only, but then, with the advent of competition, to ones that use resources more efficiently.

5.3. Biological Evolution of Dissipative Structures and the Role of Catastrophes, Natural and Artificial, in It

With the emergence of prokaryotic life, evolution accelerated again due to two main factors: the dominant role of competitive selection and the emergence of genetic memory, *i.e.*, heredity. Competitive selection led to surviving of not just viable, but the organisms most adapted to the environmental conditions. Since the competition for resources has become a major mechanism of selection, the generalized entropy time of biological systems began to “split” into metabolic times corresponding to individual resources, as described in Section 3.3. Then, genetic memory allowed storing in recessive alleles a history of past failures, which can still be very successful in changed environmental conditions. Note that if the elements of a competitive selection, as mentioned above, were already at the late stages of chemical evolution, it was nothing like the genetic memory or heredity in physicochemical evolution.

Lloyd A. Demetrius and Volker M. Gundlach [65] suggested evolutionary entropy as a quantitative measure of fitness and formulated the fundamental theorem of evolution that asserts an increase in evolutionary entropy when resource composition is diverse and resource abundance constant. Respectively, as soon as the environment is diverse and resources are available, complexity in accordance with the general law of complication increases during evolution, too. This increasing of the complexity and concomitant rising of diversity went on as usual, increasing from several stable atoms at the beginning of the physicochemical evolution to about 10,000,000,000,000 proteins in the contemporary biosphere [66]. However, raising of the complexity and diversity by itself does not guarantee invulnerability. On the contrary, the more a system is complex, the more it is vulnerable to disasters, both external and internal.

External catastrophes include the unexpected and powerful impacts of free energy, to which the system is completely not adapted. The free energy as an information killer drastically simplifies the system and throws it back in its development. However, the complexity and information already accumulated by the system are not destroyed completely, as a rule, and the system according to

conservative or casual evolution, mentioned in Section 3.4, continues developing, not from scratch, but from some already achieved level.

Internal catastrophes are caused by ineffective links within the system, when complexity becomes excessive for a given level of evolution and leads to duplication, triplication, and so on, of relations, circuiting them into loops, nesting loop ones into others and, as a result, to the collapse of the system due to loss of coordination between the elements.

Such catastrophes inevitably occur from time to time, both as a result of natural complication and because of the not quite disastrous development at higher levels of the hierarchy, which should not be evaluated only negatively. Indeed, the increase of complexity, according to the general law of complication, leads to the achievement of a local maximum in the evolutionary landscape. This gets a system into a kind of dead end where the material for further evolution is essentially exhausted. Almost everybody is familiar with this, watching how excessive complexity (bureaucratization) of a business or public organization leads to the situation when it begins to serve itself and loses all potential for further development. The result can be either a bankruptcy due to a general economic crisis (external catastrophe) or, for example, self-destruction or decay into several businesses or organizations as a result of the loss of effective governance and, ultimately, competitiveness (internal catastrophe). However, dumping a system with such a local maximum, the catastrophe gives it the opportunity to continue the complication process and potentially achieve a higher peak.

During the physicochemical and prokaryotic evolution, catastrophes were purely random (with respect to evolution itself), which causes its pace to be relatively slow. However, with the emergence of eukaryotes and especially multicellular organisms, the situation changed radically again. Living organisms “learned” not only to use internal catastrophes in their evolution, but also to cause them artificially on a permanent basis. Of course, the word “learned” should not be taken literally; just that systems that survived catastrophes relatively easily won the competition until the appearance of systems that caused internal catastrophes themselves and had, respectively, a huge selective advantage. The most striking example of such an artificial internal catastrophe is the formation of germ cells in multicellular organisms, their fertilization, and subsequent development of a new multicellular organism from scratch, *i.e.*, from one cell (zygote) that has all of the genetic information accumulated by previous generations. In addition, such processes as mitosis in eukaryotes, meiosis in the formation of sex cells, pupation, changing life forms in insects or the formation of cysts and spores, hormonal changes at puberty in higher animals, *etc.*, can also be seen as internal catastrophes, but on a smaller scale.

Of course, the appearance of internal artificial catastrophes did not cancel external ones, which systems cannot control in principle. Therefore, after the Cambrian period, when life had mastered managing catastrophes and began to blossom, dramatically increasing its diversity and spawning countless new forms, there were at least five biosphere species extinctions due to one or another external global catastrophe. There were the Ordovician-Silurian extinction (60% of species died out), the Upper Devonian (75%), Permian-Triassic (>90%), Triassic-Jurassic (65%) and Cretaceous (75%) among them. Each of them reduced the diversity, information and, respectively, complexity of the biosphere more than twice; and every time, complexity was not only restored, but also became significantly higher than the previous maximum, ensuring the irreversibility of evolution and flow of the generalized entropy time.

Similar processes we observe in human society, where floods, ice ages, earthquakes, eruptions and other natural disasters play the role of external catastrophes, while revolutions, perestroikas and all sorts of adjustment, paradigm shifts, reorganizations, *etc.*, internal ones.

Thus, the question posed in the title of this article can be answered in three ways: (1) the arrow of time is flying thanks to the general law of complication; (2) its goal is the maximum complexity, maximum information, maximum entropy and maximum degrees of freedom under the given conditions; but (3) this arrow misses each time due to the laws of nature and the internal and external catastrophes, which provide inevitable and irreversible evolution.

6. Conclusions

1. The irreversibility of time is expressed as the increase in entropy, information, degrees of freedom and complexity, which rise monotonically with respect to each other.
2. Using the extremal principle and theory of categories, it was shown that the entropy of the system can be determined without any statistical assumptions and distributions as a generalized entropy and, consequently, information, degrees of freedom and complexity.
3. The increase in such generalized entropy, information, degrees of freedom and complexity can be considered as a generalization of the second law of thermodynamics in the form of the general law of complication and determines the direction of the arrow of time in our universe.
4. The engine of physicochemical evolution (as the implementation of the arrow of time) is the general law of complication, while its driver is the other laws of nature, which are very simple and exist essentially out of time.
5. The mechanism of physicochemical evolution is eliminating or stabilizing selection of structures admissible by the laws of nature (*i.e.*, stable) from the entire set of options generated by complication.
6. The emergence of stars, generated by the steady stream of free energy, allowed selecting not only stable, but also dissipative structures, competing for energy and material resources and maximizing entropy production.
7. The enrichment of selection with competition for resources and the survival of not just possible, but the fittest structures elevated evolution to the new, biological level, and this biological evolution was accelerated sharply, especially due to managing internal catastrophes by multicellular organisms.
8. Thus, the flight of the arrow of time constantly accelerates, and its aim is maximum at the moment of the complexity, degrees of freedom, information and entropy. However, this arrow never hits local targets, because of the limitations imposed by the laws of nature and, in particular, by catastrophes.

Acknowledgments

The authors deeply thank Gary Sorrell and Denis Vasilyev for their editing and constructive insight.

Author Contributions

Both authors contributed to the research and the writing. Both authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Boltzmann, L. *Weitere Studien Uber das Warmegleichgewicht Unter Gas Molekullen*; Springer: Wiesbaden, Germany, 1970.
2. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1963.
3. McArthur, R. On the Relative Abundance of Species. *Am. Nat.* **1960**, *94*, 25–36.
4. Hutcheson, K. A Test for Comparing Diversities Index Based on Shannon Formula. *J. Theor. Biol.* **1970**, *29*, 151–154.
5. Wiener, N. *Cybernetics, or the Control and Communication in the Animal and the Machine*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 1965.
6. Eco, U. *Opera Aperta: Forma e Indeterminazione nelle Poetiche contemporanee*; Bompiani: Milano, Italy, 2011.
7. Brillouin, L. *La Science et la Theorie de l'Information*; Editions Jacques Gabay: Paris, France, 1988.
8. Ben-Naim, A. *Entropy and the Second Law: Interpretation and Miss-interpretations*; World Scientific: Singapore, Singapore, 2012.
9. Ben-Naim, A. *Entropy Demystified*; World Scientific: Singapore, Singapore, 2008.
10. Gell-Mann, M. *The Quark and the Jaguar*, 3rd ed.; St. Martin's Press: London, UK, 1995.
11. Lloyd, S. Measures of Complexity: A Non-exhaustive List. *IEEE Control Syst. Mag.* **2001**, *21*, 7–8.
12. Solomonoff, R.J. A Formal Theory of Inductive Inference. Part I. *Inf. Control* **1964**, *7*, 1–22.
13. Solomonoff, R.J. A Formal Theory of Inductive Inference. Part II. *Inf. Control* **1964**, *7*, 224–254.
14. Kolmogorov, A.N. Three Approaches to the Quantitative Definition of Information. *Probl. Inf. Transm.* **1965**, *1*, 1–7.
15. Chaitin, G.J. On the Length of Programs for Computing Finite Binary Sequences: Statistical Considerations. *J. Assoc. Comput. Mach.* **1969**, *16*, 145–159.
16. Kolmogorov, A.N. Combinatorial Foundations of Information Theory and the Calculus of Probabilities. *Russ. Math. Surv.* **1983**, *38*, 29–40.

17. Cover, P.; Gacs, T.M.; Gray, R.M. Kolmogorov's Contributions to Information Theory and Algorithmic Complexity. *Ann. Probab.* **1989**, *17*, 840–865.
18. Lloyd, S.; Pagels, H. Complexity as Thermodynamic Depth. *Ann. Phys.* **1988**, *188*, 186–213.
19. Febres, G.; Jaffe, K. A Fundamental Scale of Descriptions for Analyzing Information Content of Communication Systems. *Entropy* **2015**, *17*, 1606–1633.
20. Grunwald, P.; Vitanyi, P. Shannon Information and Kolmogorov Complexity. **2004**, arXiv: cs/0410002.
21. Ladyman, J.; Lambert, J.; Weisner, K.B. What is a Complex System? *Eur. J. Philos. Sci.* **2013**, *3*, 33–67.
22. Lloyd, S. On the Spontaneous Generation of Complexity in the Universe. In *Complexity and the Arrow of Time*; Lineweaver, C.H., Davis, P.C.W., Ruse, M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 80–112.
23. Sant'Anna, A.S. Entropy is Complexity. **2004**, arXiv: math-ph/0408040.
24. Teixeira, A.; Matos, A.; Souto, A.; Antunes, L. Entropy Measures vs. Kolmogorov Complexity. *Entropy* **2011**, *13*, 595–611.
25. Wolpert, D.H. Information Width: A Way for the Second Law to Increase Complexity. In *Complexity and the Arrow of Time*; Lineweaver, C.H., Davis, P.C.W., Ruse, M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 246–275.
26. Levich, A.P. Time as Variability of Natural Systems. In *On the Way to Understanding the Time Phenomenon: The Constructions of Time in Natural Science, Part I*; World Scientific: Singapore, Singapore, 1995; pp. 149–192.
27. Levich, A.P.; Solov'yov, A.V. Category- Functor Modelling of Natural Systems. *Cybern. Syst.* **1999**, *30*, 571–585.
28. Levich, A.P. *Art and Method of Systems Modeling: Variational Methods in Communities Ecology, Structural and Extremal Principles, Categories and Functors*; Computing Investigation Institute: Moscow, Russia, 2012.
29. Levich, A.P.; Fursova, P.V. Problems and Theorems of Variational Modeling in Ecology of Communities. *Fundam. Appl. Math.* **2002**, *8*, 1035–1045.
30. Kauffman, S.A. Evolution beyond Newton, Darwin, and Entailing Law: The Origin of Complexity in the Evolving Biosphere. In *Complexity and the Arrow of Time*; Lineweaver, C.H., Davis, P.C.W., Ruse, M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 162–190.
31. Levich, A.P. Entropic Parameterization of Time in General Systems Theory. In *Systems Approach in Modern Science*; Progress-Tradition: Moscow, Russia, 2004; pp. 153–188.
32. Eddington, A.S. *The Nature of the Physical World: Gifford Lectures (1927)*; Cambridge University Press: Cambridge, UK, 2012.
33. Tribus, M.; McIrvine, E.C. Energy and Information. *Sci. Am.* **1971**, *225*, 179–188.
34. Cooper, L.N. *An Introduction to the Meaning and Structure of Physics*; Harper: New York, NY, USA, 1968.
35. Goldreich, O. *P, NP, and NP-Completeness: The Basics of Computational Complexity*; Cambridge University Press: Cambridge, UK, 2010.

36. Fuentes, M.A. Complexity and the Emergence of Physical Properties. *Entropy* **2014**, *16*, 4489–4496.
37. Lineweaver, C.H. A Simple Treatment of Complexity: Cosmological Entropic Boundary Conditions on Increasing Complexity. In *Complexity and the Arrow of Time*; Lineweaver, C.H., Davis, P.C.W., Ruse, M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 42–67.
38. Lloyd, S. Black Holes, Demons and the Loss of Coherence: How Complex Systems Get Information, and What They Do with It. Ph.D. Thesis, The Rockefeller University, NY, USA, April 1988.
39. Wolchover, N. New Quantum Theory Could Explain the Flow of Time. **2014**. Available online: <http://www.wired.com/2014/04/quantum-theory-flow-time/> (accessed on 8 July 2015).
40. Wolchover, N. February 1927: Heisenberg's uncertainty principle. *Am. Phys. Soc. News* **2008**, *17*, 3.
41. Mikhailovsky, G.E. Biological Time, Its Organization, Hierarchy and Presentation by Complex Values. In *On the Way to Understanding the Time Phenomenon: The Constructions of Time in Natural Science, Part I*; Levich A.P., Ed.; World Scientific: Singapore, Singapore, 1995; pp. 68–84.
42. Mikhailovsky, G.E. Organization of Time in Biological Systems. *J. Gen. Biol.* **1989**, *50*, 72–81.
43. Auletta, G.; Ellis, G.F.R.; Jaeger, L. Top-down Causation by Information Control: From a Philosophical Problem to a Scientific Research Program. *J. R. Soc. Interface* **2008**, *5*, 1159–1172.
44. Ellis, G. Recognizing Top-Down Causation. In *Questioning the Foundations of Physics*; Springer: Berlin, Germany, 2013; pp.17–44.
45. Prigogine, I. *From Being to Becoming: Time and Complexity in the Physical Sciences*; W.H. Freeman & Co: San Francisco, CA, USA, 1980.
46. Shulman, M. Kh. Entropy and Evolution. **2013**. Available online: http://www.timeorigin21.narod.ru/eng_time/Entropy_and_evolution_eng.pdf (accessed on 8 July 2015).
47. Wigner, E.P. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. Richard Courant Lecture in Mathematical Sciences Delivered at New York University, May 11, 1959. *Commun. Pure Appl. Math.* **1960**, *13*, 1–14.
48. Yanofsky, N.S. *Outer Limits of Reason: What Science, Mathematics, and Logic Cannot Tell Us*; The MIT Press: Cambridge, MA, USA, 2013; p. 403.
49. Thomson, W. On the Dynamical Theory of Heat. *Trans. R. Soc. Edinburgh. Excerpts* **1851**, [§§1-14 and §§99–100].
50. Penzias, A.A.; Wilson, R.W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophys. J.* **1965**, *142*, 419–421.
51. Freedman, W.L. Theoretical Overview of Cosmic Microwave Background Anisotropy. In *Measuring and Modeling the Universe*; Cambridge University Press: Cambridge, UK, 2004; pp. 291–309.
52. Newton, I. *Philosophiae Naturalis Principia Mathematica*; Royal Society: London, UK, 1684.
53. Liddle, A.R.; Lyth, D.H. *Cosmological Inflation and Large-scale Structure*; Cambridge University Press: Cambridge, UK, 2000.
54. Mukhanov, V. *Physical Foundations of Cosmology*; Cambridge University Press: Cambridge, UK, 2005.

55. Lloyd, S. Computational Capacity of the Universe. *Phys. Rev. Lett.* **2002**, *88*, 237901.
56. Charlesworth, B. Stabilizing Selection, Purifying Selection, and Mutational Bias in Finite Populations. *Genetics* **2013**, *194*, 955–971.
57. Martyushev, L.M. Entropy and Entropy Production: Old Misconceptions and New Breakthroughs. *Entropy* **2013**, *15*, 1152–1170.
58. Martyushev, L.M.; Zubarev S.M. Entropy Production of Main-Sequence Stars. *Entropy* **2015**, *17*, 658–668.
59. Prigogine, I. *Time, structure and fluctuations*; Nobel lecture, 1977, 23.
60. Zeigler, H. *Some Extremum Principles in Irreversible Thermodynamics with Application to Continuum Mechanics*; Swiss Federal Institute of Technology: Lausanne, Switzerland, 1962.
61. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes*; Interscience: New York, NY, USA, 1961.
62. Haken, H. *Information and Self-organization: A Macroscopic Approach to Complex Systems*; Springer: Berlin, Germany, 2006.
63. Nicolis, G.; Prigogine, I. *Self-organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*; Wiley: Hoboken, NJ, USA, 1977.
64. Sharov, A.A.; Gordon, R. Life before Earth. **2013**, arXiv:1304.3381
65. Demetrius, L.A.; Gundlach, V.M. Directionality Theory and the Entropic Principle of Natural Selection. *Entropy* **2014**, *16*, 5428–5522.
66. Kauffman, S. The Emergence of Autonomous Agents. In *From Complexity to Life (on the Emergence of Life and Meaning)*; Gregersen, N.H., Ed.; Oxford University Press: Oxford, UK, 2003; pp. 47–71.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).