BioCosmos

SPONTANEOUS APPEARANCE OF LIFE AND THE SECOND LAW OF THERMODYNAMICS

David W. Snoke

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Abstract

It is often argued both by scientists and the lay public that it is extremely unlikely for life or minds to arise spontaneously, but this argument is hard to quantify. In this paper I make this argument more rigorous, starting with a review of the concepts of information and entropy, and then examining the specific case of Maxwell's demon and how it relates to living systems. I argue that information and entropy are objective physical quantities, defined for systems as a whole, which allow general arguments in terms of physical law. In particular, I argue that living systems obey the same rules as Maxwell's demons.

Many people have the experience of looking at a living system, or an ecological system, and concluding, "This is all too amazing to just have happened by chance." Such impressions are often viewed as subjective, however, with no objective physical content. Some authors give this thermodynamic language by invoking the Second Law of Thermodynamics to argue that order cannot come from disorder, but quantifying what is meant by "order" is often ambiguous. A periodic crystal, for example, is orderly, as can be a periodic pattern of convection cells, but both of these periodic structures appear spontaneously due to normal thermodynamics when a natural length scale arises in a system.

The Second Law can be used to make an argument against the spontaneous appearance of life and mind, but such an argument requires careful attention to definitions of terms like information and entropy and understanding the physical quantities involved.

Entropy as a physical quantity

We live in the age of information, but the definition of information is surprisingly controversial. Often, the definition of information is connected to *communication*, the transmission of messages between intelligent beings (or machines designed by them). This entangles the discussion in issues of the nature of intelligence, consciousness, and purposive intent. It may be fruitful to have a secondary discussion on the concept of "semantic" information, which explicitly accounts for the meaning of messages between intelligent agents, but if our definition of information is tied to this,

then we must say that there is no information in biological systems at microscopic levels where there is presumably no intelligent communication happening. Yet the use of the term "information" in regard to biological systems is ubiquitous.[1,2] Rather than rejecting this terminology, we can adopt it as justified if we can find a definition of information content derived entirely from the physical properties of a system.

In defining information as a physical quantity, we can follow the example of statistical mechanics. Physicists define a number of quantities which belong only to the system as a whole, and not to any of the parts individually. For example, heat Q can be defined as a sum of the kinetic energy of the internal motion of the particles of a system relative to the center of mass of the whole system. The distinction between the "internal" motion of heat and the center-of-mass motion can only be made for the ensemble as a whole. Heat is an *extensive* property, which means that it gets larger as the system gets larger, and therefore we can have more or less of it. It is also *fungible*, which means that it can exist in different forms and can be transferred from one system to another.

It makes no sense to say that heat is not a "thing" because it can only be defined as a property of a collection of many things, and has no existence apart from those things. It is a real entity in the same way that a wave on an ocean is a real entity: it can only exist when the many molecules of water are there, but it is not equal to the water. As confirmation of the reality of heat, we have direct perception of it through our skin, an experience separate from the feel of the roughness or smoothness of something.

sciendo

Similarly, entropy is an extensive, fungible property of a whole system, defined as[3]

$$S = k_B \log \Omega, \tag{1}$$

where Ω is the total number of macroscopically equivalent states and k_B is Boltzmann's constant. Here, "macroscopically equivalent" involves some notion of course graining, in which only average properties of an ensemble, such as temperature or pressure, are important. Semiclassical quantum mechanics gives a natural way to define and count "states," but entropy can be defined without reference to quantum mechanics, e.g., by imposing a lattice of cells in real space and momentum space and counting the number of gas particles in each cell. The macroscopic properties of a gas do not depend on the details of the geometry of the cells, as long as the cells are not too large.

This definition of entropy in terms of the number of macroscopically equivalent states makes a connection of entropy to "disorder." This is based on our experience that things we care about ordering have just a few ways to be made, while random systems can be made in many ways. For example, a junk pile has many equivalent states: rearranging a junk pile would still leave it a junk pile. A neatly organized room has many fewer equivalent states; everything must be in its place.

Entropy can also be defined in terms of probability, as

$$S = -k_B \log p, \tag{2}$$

where p is the probability of a particular state being selected. If there are Ω equally possible states, then the probability of being in any specific state at a given point in time is $1/\Omega$, and (2) becomes

$$S = -k_B \log\left(\frac{1}{\Omega}\right), \tag{3}$$

which is equivalent to (1). If the states are not equally likely, a weighted average can be done, namely

$$\begin{aligned} S &= \sum_i p_i S_i = \sum_i p_i (-k_B \log p_i) \\ &= -k_B \sum_i p_i \log p_i, \end{aligned} \tag{4}$$

where p_i is the probability of state i in the total set of macroscopically equivalent states. This is called the "von Neumann" entropy.[4] In this paper we will not concern ourselves with this type of weighted average; we will assume all all the states in an entropy sum are equally likely.

Like heat, entropy is defined for a system as a whole, and as a fungible extensive quantity, we can have more or less of it, and it can flow from one place to another. It can also be traded for free energy—the thermodynamic equation $F = U\text{-}TS$ allows generation of usable energy F from internal energy U and entropy S, for a system at temperature T.

It has been argued[5] that entropy has no objective meaning because there may be hidden degrees of freedom that an outside observer cannot see. So, for example, a gas may contain a set of oxygen atoms of different isotopes, O_{16} and O_{18}. If these isotopes are sorted into different sides of a vessel, this corresponds to low entropy for this parameter of the gas, even though an observer may not see it. This argument is specious, however, because it ignores the possibility of physically real but unobserved subsystems. For example, suppose that there is a sealed, insulated canister of gas at 1000C submerged in an opaque fluid at room temperature. If a person uses a thermometer to measure the temperature of the fluid, and reads 300C, that does not imply that heat is not a real thing because the canister is invisible to the observer. This thought experiment actually has a direct analog in solid state systems; it is quite common for the phonons in a solid (the vibrational modes of the atoms) to have one temperature, carrying a certain amount of heat, while the electrons in the same system can have a very different temperature, much hotter, and carry a different amount of heat. On short time scales, these two subsystems (which represent different degrees of freedom of the full system) can be decoupled, so that an observation may detect only one or the other of them.

In the same way, entropy can belong to two subsystems separately. In the case of the gas of isotopes, we can count two extensive quantities for two different types of degrees of freedom. One is the degree of freedom corresponding to the center-of-mass motion of the atoms, while the second is the degree of freedom of the isotope number. The total entropy of the system is the sum of these two. If an observation is only sensitive to one type of degree of freedom, then only the entropy of that subsystem will be recorded. The situation is no different fundamentally from the case of a hidden canister of hot gas embedded in a larger system and decoupled from it. Not knowing for sure whether there are hidden insulated canisters in a tank of water does not negate the reality of heat.

The Second Law and entropy

The Second Law of Thermodynamics says that the total entropy of a closed system can only either stay the same

or increase. There has been extensive debate over the years, from the time of Boltzmann, in justifying the Second Law; most of the argumentation has focused on probability of random events in semiclassical systems.[6,7] Rather than reviewing that debate, here we can start with the modern understanding of the Second Law in quantum mechanics and infer its implications for statistical approaches.

The real world we live in is, of course, quantum mechanical. In particular, it is not even semiclassical; semiclassical physics envisions fixed numbers of particles like billiard balls in discrete quantum states that act like baskets. The basic theory of quantum mechanics is quantum field theory, which treats not just the number of particles in a given state, but the phase relationships between a vast number of possible many-particle states.

Recent work[8] has shown that the Second Law can be derived entirely within the framework of quantum field theory. That work showed that the concept of randomness is not needed at all; the quantum many-body wave function of a macroscopic system evolves deterministically and irreversibly toward equilibrium. What is needed for irreversibility is the erasure of microscopic details from the macroscopic, coarse-grained description of the system. In the case of quantum field theory, Ref. 8 showed that this occurs in a normal macroscopic system because the phase correlations of the many-particle quantum states rapidly decay toward zero. In principle, there is a Poincaré cycle time for recovery of this phase information, but when the size of the system is infinite, the Poincaré cycle time becomes infinite; in other words, there is never a return to the starting point. This can be understand by analogy to a wave disturbance, such as the plop of a pebble creating rings that move outward on the surface of a pond. In an infinite body of fluid, even if it is perfectly energy-conserving, the waves will move outward forever, dissipating as they spread energy over a vast space and many modes of oscillation. The system will repeat its behavior only if the system has finite boundaries that the waves can reflect from, so that they return inward.

In the argument of classical statistics leading to the Second Law, the role of the erasure of the microscopic details is played by the the assumption that at every time step, the distribution of the particles is replaced with a typical, probable one.[9] Both the loss of phase in the full quantum field-theory calculation and the removal of outlying, improbable distributions in the classical statistical approach have the effect of allowing a description of the system only in terms of average occupation numbers; once this is done, Boltzmann's H-theorem follows immediately.[8] The classical statistical approach can be unsatisfying, however, because one can always ask, "Yes, but what if a really improbable state occurs?" That question simply doesn't arise in the many-body quantum field theory approach.

The physical description of our world is very much like the infinite system considered in Ref. 8; everything on earth is coupled by electromagnetic radiation to the larger universe, which as far as we know has no boundaries that reflect back waves. (Even things with insulating walls have quantum mechanical tunneling to the larger universe). One can therefore assert that the Second Law really is a law, not just a likelihood; every system has irreversible radiation to outer space, ultimately. Thus no system on earth is a finite system with a finite Poincaré cycle time; all of the things we know are in a dynamic balance known as a "driven, dissipative" system,[10] with incoherent energy input from the sun and dissipation of energy into the rest of the universe. In such a system, unlike the scenario envisioned by Poincaré, there is scrambling of phase correlations that effectively removes memory of past states.

None of this precludes the usefulness of a statistical mechanics description. The quantum field calculation shows that with the loss of phase correlations, the description of the system becomes entirely semiclassical, in which case only the occupation numbers of the states matter. This is precisely the scenario assumed by statistical mechanics. Therefore, all of the statements of standard statistical mechanics for systems with a discrete number of atoms still apply.

We can apply this to the example considered above, a gas with two isotopes. The Second Law says that such a system will move toward a fully mixed state of isotopes in the same way that it moves toward thermal equilibrium. An outsider observer might not see this evolution toward a mixed state, but it will happen anyway. If no one measures the temperature of an object, heat will still flow through it, and in the same way, the entropy of the isotope degree of freedom can indeed change in a system if no one measures it.[11]

This example does point out an important feature of the Second Law, however. The Second Law does not operate only at some observer-defined high level, such that entropy may decrease locally in some subsystem, as long as the total entropy budget of the whole does not decrease. The Second Law says that *for every sufficiently large subsystem*, entropy must stay the same or increase. Thus, in the case of the mixture of isotopes given above, the isotopes cannot spontaneously sort themselves as long as the overall gas gets hotter, to keep the total entropy from decreasing. The total entropy of the system is the sum of the entropy in the

motional degrees of freedom and the entropy in the atomic-mass degree of freedom, and each of these entropies must obey the Second Law separately.

The reader may immediately think of the case of a refrigerator, in which entropy is locally reduced in some region, as an objection to the statement that the Second Law holds for every sufficiently large subsystem. A refrigerator is a machine, however, which I argue below is a type of Maxwell's demon. We will consider the role of the Second Law in the case of Maxwell's demons at that point.

Information as a physical quantity

It is natural to define information as an extensive physical property similar to heat and entropy. This approach has been used in the physics community for many decades, following the work of Szilard,[12] Landauer,[13] and others.

Information can be defined as the elimination of possibilities.[14] The more possibilities that are eliminated, the more information that is gained. Thus, in a system with Ω equally possible states, if we know that the system is definitely in one of them, we have gone from Ω possibilities to just one. It is therefore natural to use Shannon's definition of information,[15]

$$I = \log_2 \Omega_s, \qquad (5)$$

where Ω_s is the number of possible states that the known state was chosen from. This has the same form as the entropy defined above, and is sometimes called "Shannon entropy." In computational terms, it is equivalent to saying that the information content is equal to the number of bits needed to represent a state, because the total number of possible states that can be represented in an n-bit register is 2^n, and therefore the Shannon information is $\log_2 2^n = n$.

Although I have used the terminology of "knowing" here, the crucial property for information content here is not knowledge or communication, but *selection*. An information-processing system selects and acts on one of Ω possibilities. It is possible, of course, for an intelligent agent to perform selections, but a selection process does not intrinsically require the presence of intelligence; any physical system which has a macroscopic action one way in response to a particular selected state, and not the same way in response to others, may be said to engage in selection.

It is also important to note that Shannon information, which looks formally like an entropy, is not the same as physical
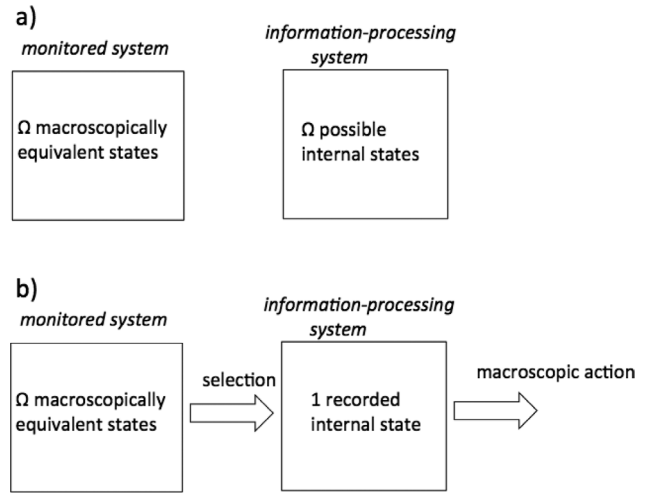

Figure 1. A monitored system and an information-processing system a) before, and b) after the selection process.

entropy. Instead, the physical entropy *before* the selection process gives the maximum possible Shannon information that can be extracted from the system. *After* the selection process, there are no longer Ω equal possibilities. The information-processing system has selected one state out of this set to act on (see Figure 1). If the information-processing system can perform an action with macroscopic effects, then this one state is no longer macroscopically equivalent to all the others. After this selection process, we may say that the information content in the detection system is $\log_2 \Omega$, but the physical entropy content is $k_B \log 1 = 0$. More generally, if N internal states generate the same macroscopic response of the information-processing system, the entropy content is $k_B \log N$, where N can be much less than Ω.

The reduction of entropy from $k_B \log \Omega$ to $k_B \log N$ is a real physical reduction of entropy, but it need not violate the Second Law of Thermodynamics. The question of how this can happen is exactly the question of Maxwell's demon, discussed below.

In defining information in terms of selection, we have assumed the existence of some information-processing system, which gives a macroscopically distinguishable response to certain states. We can therefore talk of information in terms of function. This leads to an alternate definition of information instead of Shannon information, which we may call *functional information*.

Functional information may be defined as the number of ways of altering the macroscopic function of a system. In common language, we might say that functional information content in something is the number of ways of breaking it. For example,

suppose that an information-processing system is designed to respond in some way to a specific pattern of 0's and 1's stored in its memory, and no other pattern gives the same action. Then if we change any one of the bit values, we have "broken" it, that is, altered its macroscopic function. The functional information value in this case corresponds to the Shannon information value, i.e., the number of bits. On the other hand, suppose that we have a junk pile, or a random sequence of bits which triggers no response. Then there is no function at all, and changing one of the parts will have no effect; the functional information value is 0.

Functional information differs from the Shannon information because the Shannon definition assigns high information value to large random strings. The functional definition therefore agrees much more with our common intuition of what information is. It has some aspects in common with *Kolmogorov information*,[16] which can be defined as the number of bits in the shortest possible computer program that will generate the sequence of bits in a string. Kologorov information therefore would assign low information content to a large sequence of alternating 0's and 1's. But it shares the disadvantage of Shannon information that it also assigns high information value to specific random strings, which must be written out exactly in a computer program. Functional information takes note of the fact that many random number strings are functionally the same, because they are effectively just "noise."

The functional definition of information is easily applied to biological systems in which there is no obvious intelligent communication happening. One can simply ask whether *selection* is happening, that is, macroscopically distinguishable changes of the system in response to special states. This clearly occurs in the signaling in biological systems in response to detected conditions in the environment or in the internal conditions of a cell.

The definition of function information also has the advantage that it can be applied to systems that do not obviously look like data strings. A memory register with a line of identical two-value bit-memory locations and a DNA molecule with a similar line of four-value locations obviously look like information-storage devices, but information can be stored in other ways. For example, the silicon-and-metal structure in a computer needed to read out a memory register also has information content. It may be translated into a string of bits, e.g., by creating images of the blueprints for a device and storing these images as binary data files, but the information need not be stored this way; it is already present in the device itself. This has become widely appreciated in biological research, as it is now understood that there is epigenetic information
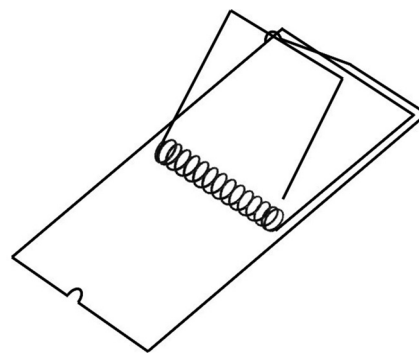

Figure 2. An information-rich system with a single part.

in the cell which is not transmitted through the register-like medium of DNA.[17]

Finally, the definition of functional information is also helpful when a system does not have obviously separable "parts." For example, the device shown in Figure 2 is full of functional information although it has a single "part;" it is a mousetrap made from a single bent wire. If we ask how many ways are to significantly alter its macroscopic function (that is, how many ways there are to "break" it), then it is easy to see that if the wire is bent differently in any one of a large number of places, it will cease to have a hair-trigger action. This corresponds to a large functional information content.

Above, we saw that entropy is definable as a physical quantity, but only as a quantity that is meaningful for a large system. The same is true of information, but it requires a special type of physical system, namely one with macroscopic selection activity. This selection or trigger activity need not be human knowledge; it could be the independent action of a digital computer, or the catching of a mouse. Thus, as with entropy, information can be taken as a real physical quantity like heat even if humans do not observe it.

Maxwell's demon

The canonical example of the interplay of energy and information is Maxwell's demon. In the standard scenario, Maxwell's demon is a tiny being that sorts the atoms in a classical gas by observing the positions of the atoms and then opening and closing a small door, to either reflect the atoms back or allow them to pass from one chamber into another. A tiny intelligent being may be imaginary, but the main features of this thought experiment are not; it is entirely possible to do this experiment in reality. A human can play the role of Maxwell's demon, if the particles are large enough and slow enough. For example, we can suppose that the "particles" are large, massive objects floating in a zero-
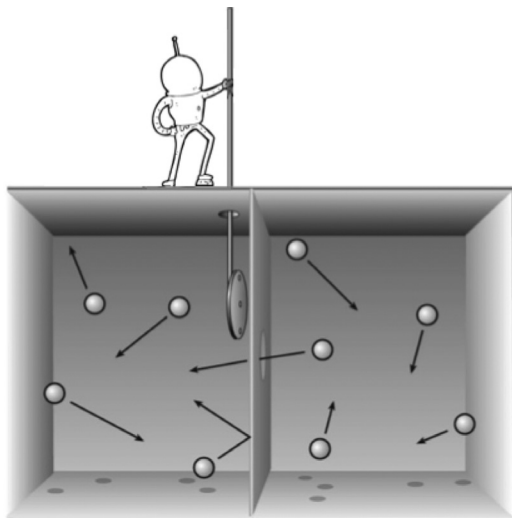
Figure 3. A real Maxwell's demon experiment performed by a human with large objects floating in a zero-gravity environment.

gravity environment in outer space, as shown in Figure 3. We can always make a classical system large enough so that a) the energy cost of observing the positions and speeds of the particles is negligible, and b) the energy cost of opening and closing the sliding door is negligible. (Although this experiment is unlikely to be really done in outer space, it has been simulated as a video game used by many physics teachers which has all the same relevant features).

This system can be turned into a machine generating usable work with just a few small changes. The central wall between the chambers can be allowed to slide back and forth depending on the relative pressure difference between the two chambers. The demon (i.e., the person, in the zero-gravity implementation discussed above) is able to clamp or unclamp the sliding wall with negligible energy cost. In this case, a cycle can be set up. The demon first sorts the atoms to get higher pressure on one side of the wall than the other, while the wall is clamped. Then the wall is unclamped, and the pressure difference moves the wall, doing work on the wall, which can be connected to an external device. The wall is then clamped again, and the particles are re-sorted to give higher pressure on the opposite side. When the wall is then unclamped, it will be pushed back the other way. The oscillating slides of the wall in response to the sorting can drive any type of cyclical machine, such as an electric generator. The energy to perform this work comes from the kinetic energy of the floating particles, which can be restored at constant temperature by the vibrations of the walls of the chambers. This general scheme is known as a "Szilard engine."[12] The Szilard engine shows that information and free energy are fungible; that is, the information processing
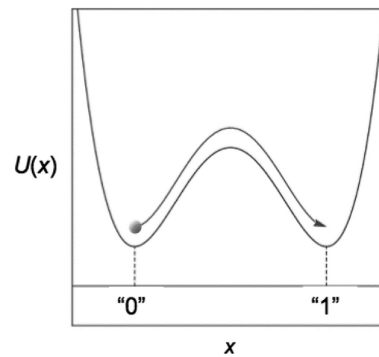

Figure 4. A single-bit storage device.

agent (in this case, the human) can convert information into usable work.

The action of the Szilard engine does not violate the Second Law of Thermodynamics. One might initially think so. For example, we could replace the human with a computer/robotic system that performed the same sorting action. This computer could be run by an electric generator. Since the Szilard engine can run an electric generator, it can supply energy to the Maxwell's demon computer. Could this then be used to make a perpetual motion machine? If not, why not?

The proof that it cannot is attributed to Landauer.[13] The argument focuses on the cost to store and reset the information storage. The simplest memory storage device is illustrated in Figure 4, namely a system with two energy minima and a barrier between the two. To store one bit of information in a system at a temperature T, the barrier height must be large compared to $k_B T$. Otherwise random thermal excitations could cause the state of the system to hop over the barrier. To write information, that is, to put the system definitely into one of the two possible states, energy greater than the barrier height must be put into the system to lift it over the barrier, and then this energy must be dissipated to restore the system to a state well below the barrier. Each definite act of setting the value of the bit therefore dissipates energy greater than $k_B T$.

Let us suppose that the Szilard engine is at the same temperature as the memory storage device. One bit of information corresponds to identifying whether one particle is on the left or the right of the central barrier. Allowing this particle to hit the central partition and move it transfers an average energy of $k_B T$ from the gas to the barrier, which can be turned into usable work. But the cost of recording one bit of information, as we have seen, is larger than $k_B T$. Thus the system cannot operate as a perpetual motion machine.

The information processing system (demon) and the gas of particles do not need to be at the same temperature, however.

Let us return to our example of a human at temperature of 37C sorting large massive objects in a zero-gravity environment. The effective temperature of these objects is set by their average kinetic energy, which may be, say, $E= \frac{1}{2}mv^2$ with m = 1 kg and v = 1 m/s. This corresponds to T ~ E/k_B ~ 10^{23} Kelvin, far above the temperature of the human sorter. After a very long time, they might come to equilibrium at the same temperature as the human, but that could be an extremely long time. We can therefore think about the general case of two different temperatures for the demon and the sorted system.

In the general best-case scenario, for every bit recorded and used in the Szilard engine cycle, an energy of $k_B T_2$ is gained for usable work from the gas of particles at temperature T_2, and $k_B T_1$ is dissipated into the environment at temperature T_1 of the information processing system. If the two systems are held at their different temperatures by contact with two different heat reservoirs, this system will convert the temperature gradient into work in a continuous cycle. The efficiency of the process is defined as the work done divided by the heat input. For N bits of information converted, the heat input is $Nk_B T_2$, while the work output is $Nk_B T_2$ minus the amount needed to run the information processing system, which is of order $Nk_B T_1$. We thus have

$$\eta = \frac{Nk_B T_2 - Nk_B T_1}{Nk_B T_2} = \frac{T_2 - T_1}{T_2}. \tag{6}$$

This is the same efficiency as an ideal Carnot engine[3] using the difference between two heat baths to produce work. The similarity of the two is not surprising, because a Carnot engine is actually an example of a two-bit information processor, that is, a two-bit Maxwell's demon. In the standard Carnot process, the system detects four combinations of pressure and volume and operates two switches by which it can connect and disconnect heat flow to two baths at different temperatures in response to the information it has about the state of the system. Rather than binary information about which side of a barrier an atom occupies, the Carnot engine gathers binary information about whether the pressure and volume exceed or fall below certain thresholds.

The standard scenario of Maxwell's demon which we visualize is quite similar to the case discussed above, with a person who can observe the position and velocity of massive objects floating in zero gravity. This implicitly assumes a high temperature difference, which in turn implies nearly perfect efficiency, when the cost to set a bit of information $k_B T_1$ is negligible.

We can define a "machine" as any device that acts like a Maxwell's demon to use information to change the macroscopic state of a system. We can then define a continuum from a simple Carnot engine with two bits of actionable information, to a full computer processor, to a living system doing the same type of stimulus-response selection. Even in this limit, however, the Szilard engine, like the Carnot engine, does not violate the Second Law in its normal operation, as the work performed is ultimately driven by heat flow from the hot bath.

Living systems often operate without a temperature gradient. However, they often use another type of imbalance, namely a concentration gradient,[18] to have the same effect. Like a Szilard engine, they generate usable free energy by selecting different macroscopic responses to different states of the system.

Information and entropy in the existence of Maxwell's demons

We have seen that both entropy and information can be defined as real, physical properties of systems as a whole. They are not the same as each other, although they involve similar math of counting macroscopically equivalent states. The distinction between the two can help us to think about the origin-of-life problem.

Much past work has shown that the *operation* of an information-processing machine, for example a refrigerator, does not violate the Second Law of Thermodynamics. It is easy to see that every subsystem of the whole refrigerator locally obeys the Second Law. Heat always flows from hot to cold, and gas always flows from high pressure to low pressure. The work of Jeremy England[19] is similar, showing that the operation of living things, including in the reproduction process, does not violate the Second Law, even though they operate far from equilibrium.

But there is a second question that can be asked, namely, what is the likelihood of the *existence* of the refrigerator? We all intuitively know that it would be highly unlikely to arrive at the planet Mars and find a fully operational electric refrigerator there, produced by the action of the wind and water on Mars. If we did, we would find that it obeyed the Second Law while operating, but that wouldn't reduce our surprise at finding it. In the same way, it is not the daily operation of living things that surprises us, but their origin.

We can quantify this distinction by talking of refrigerators, machines, and living systems in terms of information, not just entropy. Namely, we can quantify the degree of surprise we feel in terms of the amount of information processing that

occurs. A four-bit processor such as a refrigerator surprises us somewhat; a living system which selects raw materials from the environment to construct copies of itself in a process with hundreds of switches, surprises us more.

The Second Law tells us that entropy cannot decrease. Is there an equivalent law for information? Dembski has proposed,[20] as an axiomatic assertion, that information can never spontaneously increase. Can we do better, to create an information principle based on the Second Law?

Let us consider the initial state of the Szilard engine scenario discussed above. We suppose that the gas is thermally disconnected from the outside world, so that no heat flows in or out. At the start, the gas is in a maximal entropy state distributed between the two chambers. After the information processor (demon, computer, or person) has acted for some time, the gas is sorted so that nearly all of the particles are on one side of the partition.

After this sorting process, the entropy of the gas is greatly reduced. The information processor has dissipated some energy, raising the entropy slightly, but as discussed in the previous section, this amount can be made negligible if the processor is at low temperature compared to the gas. The net entropy of the system therefore appears to have been reduced dramatically. Does this violate the Second Law?

If we do not allow that the Second Law has not been violated, then it must be the case that the entropy of the whole system was already low at the start. This is obvious when we think of the definition of entropy in terms of the number of macroscopically equivalent states. To treat the initial state of the gas as having high entropy, we counted a large number of states as equivalent. But they were manifestly *not* equivalent if the information processor could distinguish between them and produce a different macroscopic result for different states. The existence of the information processor which makes the states *not* macroscopically equivalent means that the entropy of the whole system was already low before the sorting started.

For the Second Law of Thermodynamics to hold, we must adopt the viewpoint that the initial state, consisting of the information processor and gas together, already had low entropy. Since the gas by itself was the same as any other gas in the same volume, we must say that existence of the information processor caused the whole system to have low entropy. We may therefore say in general terms that a system with an information processor is a low-entropy state. This is in accord with our experience that information processors, even simple ones such as two-bit information processors,

e.g., refrigerators, do not appear spontaneously, but in our experience are always generated from other information processors. Refrigerators, machines, and computers are generated by humans. We never find even simple refrigerators popping up spontaneously from nonliving matter.

The quantify this, one can say that *the probability of an information processor appearing in a system is the same as the probability of a fluctuation in that system which reduces the entropy of the initial state to the amount apparently removed by the machine*. Fluctuations with entropy flow in the reverse direction of the Second Law are known from statistical mechanics to occur,[21] with a probability that deceases exponentially with the amount of entropy involved. In the case of gradual increase of information, we can say that at any stage, the probability of moving upward in information content is the equal to the probability of going down in entropy by the amount apparently removed by the new parts of the machine.

Let us apply this to another example. Suppose we have a system consisting of two chambers, with a one-way door between them. If a particle hits the door from one side, it will push the door open and pass through into the other chamber, after which the door springs closed again. If a particle hits the other side of the door, it bounces back, and cannot pass through.

After some time, the gas in the two chambers will be mostly in one chamber. This has the action of a Maxwell demon, although it is not cyclical—the process stops when the number of particles on one side is so great that it is likely that some particles flow backwards through the door when a particle opens the door from the other side. The process could also be stopped if the heat dissipated in the door causes enough vibrations that the door flaps open and shut randomly.

There seems to be no action of an intelligence or information-processing system in this case, but the entropy (when looking only at the gas) has been reduced just as with Maxwell's demon. As argued above, we can say that the system must have started with low entropy. This becomes obvious when we realize that existence of the one-way door has very low probability. If a human did not intervene to make this system, it would be highly unlikely to arise on its own. First, a dividing wall between two chambers must arise. Then, a hole in this divider must arise. A door must arise next to this hole which can open only when pushed from one side. This door must be attached to the dividing wall, and it must have a spring to pull it closed again after it has been opened. It also must have damping to dissipate the energy once it is closed, so that it does not just bounce

open again, but it must not have too much damping, or the damping will prevent the door from being opened when it is hit by an incoming particle on one side.

Living organisms have systems which are analogous to this one-way door, e.g., pumps which transport ions in one direction across a membrane.[22] These systems can also be incorporated in larger, cyclical systems which operate similar to a heat pump. Therefore we can say from the above considerations that the spontaneous appearance of a living system is indeed a low-entropy, low-probability state. If we are to preserve the Second Law of Thermodynamics, we must say that the probability of a one-way door arising spontaneously is no greater than the probability of a gas spontaneously sorting itself into the same low-entropy state created by the one-way door.

Cyclical machines. In the example of the one-way door, the action of the machine stopped after a certain point. Many machines have cyclical action, however, so that they are reset and work again many times. Intuitively, this is clearly a step higher in sophistication and therefore expected to be less probable to occur spontaneously.

As discussed above, the action of a machine in its normal operational cycle does not involve a net entropy cost; machines obey the Second Law. However, the appearance of a new part of the machine that takes the state of the system from one step in the cycle to the next is equivalent to the appearance of a new machine. We may therefore say that the probability of the full cyclical machine appearing is the same as the probability of the first state of the system appearing spontaneously in the absence of the machine, and the second state of the system appearing spontaneously from the first state in the absence of the machine, and the third state appearing spontaneously from the second state in the absence of the machine, and so on for as many steps there are in the cycle. Since the probability of a set of multiple events is given by the product of the individual probabilities, in the absence of known correlations, this can make the probability problem of a cyclical machine's appearance quite severe.

Of course, if a cyclical machine is made by another machine, there is no entropy cost; the machine doing the making obeys the Second Law in dissipating energy while it makes another machine, as pointed out by England.[19] But this raises an obvious regress problem: whence the first machine, which made the second one? Following the same approach, we must say that the existence of a machine-making machine is itself a low-entropy state, with a probability given by the product of the probabilities of each of its steps.

Natural pattern formation

The above considerations have been subject to much debate following the work of Nobel prize winner Ilya Prigogene[23] and others who have pointed to the existence of spontaneous pattern formation in nature. In the language of this paper, spontaneous pattern formation acts as a 1-bit information processing machine; when a critical threshold is crossed, e.g., the temperature difference between two plates of a convection chamber like the one illustrated in Figure 5, a macroscopic response occurs, in this case the appearance of macroscopic convection cells with visible boundaries. Similar pattern formation occurs in many cases in nature, such as regularly spaced ridges of sand on the sea shore, or regularly spaced clouds in the sky. It has widely been taken that the import of this work is that information-processing machines (in the terminology of this paper), which are the basis of life, can appear spontaneously in systems that obey the Second Law but are not in equilibrium.

Let us analyze this case using the approach of the previous section. Since the system acts as a one-bit information processor, we can assume that one bit of information processing has been front-loaded into the system. We can easily identify where. A natural direction (up-down) has been defined by gravity and the placement of the plates with the temperature inversion. There is then a natural length scale in the system that responds to this vector, namely the convection cell size, which depends on the viscosity of the fluid and the spacing of the plates. Condensed matter physics shows
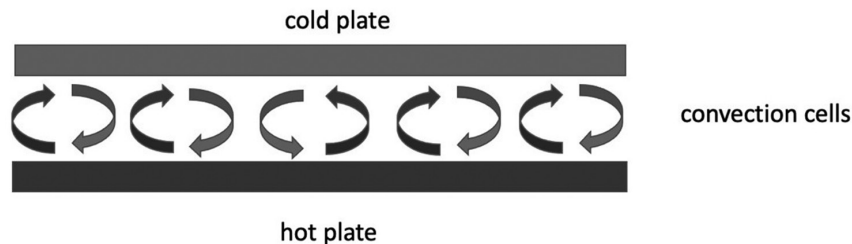

Figure 5. Spontaneous pattern formation in a convection chamber.

that when a natural length scale arises in a system, there is generally an instability to cause structures with that length scale. The same thing can happen in the case of natural time scales, which lead to the spontaneous appearance of natural clocks.

In the case of an experimental apparatus, a human has designed and built the system, and is the agent of the front-loading of information. In the case of natural patterns on earth, these ultimately stem from the up-down vector of the earth's gravity in conjunction with the properties of fluids, which in turn stem from the law of gravity that favors highly compact planets with fluids on their surfaces.

This shows that a single-bit processor can arise spontaneously in nature, but does it follow that processing of any number of bits, up to the millions of triggers used in biological systems, can arise by the same process? Many decades of research on this have largely failed; for example, a single lipid layer can form spontaneously under properly front-loaded circumstances, but putting a second lipid layer inside a lipid bubble, facing the opposite direction, as is the basis of all living cells, is beyond any known science and engineering capabilities.[24] To make spontaneous information processing with additional bits, one must create a system with two, three, and more natural length scales analogous to the single natural length scale of convection cells discussed here. While such might be accomplished in the lab by carefully designed processes, for it to occur in nature would require that the laws of nature have multiple natural length scales written into them which can play the same role as gravity in a convection chamber, but which are different from gravity. While the convection chamber does show spontaneous information processing, it has exhausted all of the front-loaded resources it has, and cannot generate any more degrees of freedom for triggered macroscopic responses.

For example, as biochemist David Keller has pointed out,[25] just one biological machine, DNA polymerase, has 90 design parameters including 15 length scales that must be fine tuned to match the (independently produced) DNA structure, which the polymerase holds with a "hand and glove" structure; if each of these parameters does not have a specific value, the machine will not work (it will "break"). None of these length scales are natural length scales that would appear spontaneously; all of the length scales of this machine are produced by other molecular machines based ultimately on information stored in the genome. Indeed, a major feature of the DNA transcription system is that there are no natural length scales to which structures are forced to conform, which allows protein machines of many different shapes and sizes

to be built. Furthermore, the DNA replicase structure cannot be generated by natural selection, at least in its present form, because it plays a crucial role in replication, which must exist for natural selection to work. The DNA replicase information is highly conserved in all living systems;[26] if mutations occurs in this structure, the creature simply dies.

Conclusions

The main conclusions of this paper may be stated as the rejection of several claims.

- Entropy and information are not intrinsically tied to human knowledge; there are ways to define them entirely in terms of physical properties of a system. However, they are not single-particle properties, but instead are properties of an aggregate whole system. Information can be defined functionally in terms of a selection mechanism that gives a macroscopic response to some subset of a larger set of possibilities.
- A Maxwell demon need not dissipate as much energy as it creates in free energy; although this must be the case if the system and the demon are at the same temperature, they need not be at the same temperature. More generally, a perfect Szilard engine based on a Maxwell demon operates with the same efficiency as an ideal Carnot engine. Indeed, a Carnot engine can be viewed as an example of a Maxwell demon; both are examples of a broader class, namely machines that have the action of selection leading to macroscopic action, which we may generally call information processing.
- The Second Law does not just hold for the total entropy of a system, but for every sufficiently large subsystem. Therefore entropy loss in one subsystem cannot be traded off for entropy gain in another system, in the absence of an information-processing machine.
- Showing that information-processing machines, e.g., Maxwell's demons, do not violate the Second Law in their regular operation does not say anything about their origin. From basic considerations, we can assume that the spontaneous appearance of an information-processing machine has the same improbability as a negative entropy fluctuation equal to a negative entropy fluctuation equal to the apparent loss of entropy in obtaining the machine's initial action without the presence of the machine.
- The spontaneous appearance of 1-bit information processors in natural pattern formation does not point the way toward spontaneous formation of million- to billion-bit processors as are common in living systems (such as human brains). For each new bit to be processed, another natural process with its own natural length

scale or natural time constant must be posited. There is an upper bound to how many of these exist in nature, presumably of the order of the number of free parameters in the laws of nature themselves.

Living systems are to all intents and purposes equivalent to Maxwell's demons, in that they are information processors that perform selection processes. We may therefore conclude that there is a fundamental entropy problem with the origin of life.

References

1. E.g., E. Lutz and S. Ciliberto, "Information: From Maxwell's demon to Landauer's eraser," Physics Today **68**, 30 (2015).
2. For a review, see D. Snoke, "Systems biology as a research program for intelligent design," Bio-Complexity **3**, 1 (2014).
3. F. Reif, *Fundamentals of Statistical and Thermal Physics*, (McGraw-Hill, 1965).
4. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, 1955).
5. G. Sewell, "On 'compensating' entropy decreases," Physics Essays **30**, 1 (2017).
6. J. von Neuman, "Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik," Zeitschrift für Physik **57**, 30 (1928).
7. H.R. Brown, W. Myrvold, and J. Uffink, "Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics," Studies in History and Philosophy of Modern Physics **40**, 174 (2009).
8. D.W. Snoke, G.-Q. Liu, and S.M. Girvin, "The basis of the second law of thermodynamics in quantum field theory," Annals of Physics **327**, 1825 (2012).
9. See, e.g., J. Uffink "Boltzmann's work in statistical physics," Stanford Encyclopedia of Philosophy, E.N. Zalta, ed., (2017).
10. See, e.g., D.A. Paz and M.F. Maghrebi, "Time-reversal symmetry breaking and resurrection in driven-dissipative Ising models," arXiv:2105.12747 (2021).
11. This presumes that one is not a solipcist, who denies that we can assume the reality of anything not observed by humans (i.e., a tree falling in the woods makes no sound). We can say that the laws of physics in this situation imply that isotopes will mix even if no one observes them; Occam's razor says that there is no good reason to assume the laws of physics are suspended when people don't look.
12. L. Szilard, "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings," Zeitschrift für Physik **53**, 840 (1929).
13. R. Landauer, "Irreversibility and heat generation in the computing process," IBM Journal of Research and Development **5**, 183 (1961).
14. R. Stalnaker, *Inquiry*, (MIT Press, 1984).
15. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, (U. Illinois Press, 1963).
16. A.N. Kolmogorov, "Three approaches to the quantitative definition of information," International Journal of Computer Mathematics **2**, 157 (1968).
17. N Carey, *The Epigenetics Revolution: How Modern Biology is Rewriting our Understanding of Genetics, Disease, and Inheritance*, (Columbia University Press, 2013).
18. See, e.g., "Concentration Gradient,'' Biology Dictionary, Biology-dictionary.net, 10 Jan. 2017. (https://biologydictionary.net/concentration-gradient/.)
19. J.L. England, "Statistical physics of self-replication," Journal of Chemical Physics **139**,121923 (2013).
20. W. Dembski, *No Free Lunch: Why Specified Complexity Cannot Be Purchased without Intelligence*, (Rowman and Littlefield, 2001).
21. D.J. Evans and D.J. Searles, "The fluctuation theorem," Advances in Physics **51**, 1529 (2002).
22. See, e.g., G. Karp, *Cell and Molecular Biology: Concepts and Experiments*, 6th ed., (Wiley, 2009), chapter 4.
23. I. Prigogene, "Thermodynamics of evolution," Physics Today 25, **23** (1972); *From Being to Becoming*, (W.H. Freeman, 1980).
24. T. Dingjan and A.H. Futerman, "The fine-tuning of cell membrane lipid bilayers accentuates their compositional complexity," Bioessays **43**, 2100021 (2021).
25. David Keller, University of New Mexico, talk at the Christian Scientific Society, Pittsburgh, 2008. For an overview of the structure of DNA polymerase, see M.H. Lamers, R.E. Georgescu, S.-G. Lee, M. O'Donnell, and J. Kuriyan, "Crystal structure of the catalytic a subunit of E. coli replicative DNA polymerase III," Cell **126**, 881 (2006).
26. S. Waga and B. Stillman, "The DNA replication fork in eukaryotic cells," Annual Reviews of Biochemistry **67**, 721 (1998).