Chapter 1
Life, Energy and Information

The principles we need to take into consideration while studying the evolution of how matter aggregates are quite simple.

To extract energy from a system, we need information about that system. We need to be able to predict how it will react, evolve. But to process information, in a brain, in a computer, in the DNA, we need energy.

Life means storing information to extract energy, and extract energy to store information. This chapter will analyse the concepts of energy and information, and how they relate to each other.

What Is Life?

Know thyself is a good starting point for someone who wants to study its origin. And in our case, we must start with the question: what is life, what does it mean to be alive?

Erwin Schrödinger's, one of the parents of quantum mechanics, gave this definition in his booklet *What is life* (1944):

> Life is organized matter which evades the decay to equilibrium through absorption of energy

If we do not eat, all our cells, and then molecules, and the atoms, will be scattered in the environment. If we eat, we keep our body organised and avoid that.

A definition does not explain why life has emerged, or how it evolves. But is a good starting point to investigate these questions. Schrödinger definition allows us to clearly define the object of our interest, and is therefore worth to further explore it.

The Decay Towards Equilibrium

The concept of equilibrium mostly derives from the work of one single physicist: Ludwig Boltzmann. Towards the end of the nineteenth century, on the basis of the work done by Maxwell and others, Boltzmann introduced probability into the explanation of the evolution of systems into states of equilibrium.

The question no one could answer was –why there is such a thing as the equilibrium? Why, if I have a box filled with gas, I will never observe all molecules of gas in a single corner, but instead will see them always filling the whole box? Why do I need energy to compress all molecules in a corner?

The question is not much different from –why do I need energy to keep all molecules of an organism together? Why if I don't provide energy to the organism, all its molecules and atoms will eventually diffuse in the environment, like the molecules of gas in the box?

We can better comprehend Boltzmann's reasoning by simply imagining a gas with two molecules, one blue and one read, in a container with a semi-divider wall, as in Fig. 1.1.

If the two balls are not coloured, the states B′ and B″, with a ball on both sides, appear to be the same, so an external observer might call both states "macrostate B". If we shake the box, the macrostate B is more probable than L and R, because it is actually made of two *microstates*.

If there are one million balls (as many as are molecules in a cube of air with a 0.1 mm edge) the probability of obtaining states with approximately half the balls in each region is about $10^{300,000}$ times higher than that of obtaining states in which all the balls are on one side.

This is why we say that the system's equilibrium is with the balls, or molecules of gas, evenly distributed in the box. Unless we do something, the system will spontaneously evolve with the balls on both sides. Not always the same balls, but as far as we can say, the balls are evenly distributed.

Fig. 1.1 A box, a wall and two coloured balls. There are four microstates, i.e. states which we can identify as different thanks to the color of the balls. When the balls are indistinguishable, like molecules in a gas, we only identify three macrostates
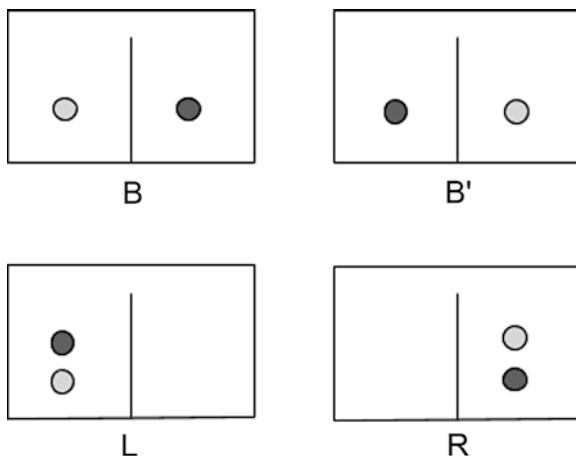
Similarly, if living systems do not absorb energy, they decay towards equilibrium in the sense that all their molecules tend to occupy the most probable macrostate, which is the one with all molecules diffused, like the ones in the gas. The gas does that quickly, our molecules slowly, but the process is the same.

To avoid the decay to equilibrium, we need energy. With some effort, we can move the balls from one side and put them on the other side, so to remain in a non-equilibrium state. Similarly, with some energy we keep our organism… organised, and all its molecules together.

Energy Extraction Requires Information

The real breakthrough for Boltzmann was linking the concept of equilibrium to the one of *entropy*, a physical quantity related to order and the ability to extract energy from a system.[1] The more a system is far from equilibrium, the more energy we can extract. Let's see what it means.

Around 80 years after Boltzmann's studies on statistics and equilibrium, one of the most brilliant minds of the twentieth century, John von Neumann[2] (1956), linked entropy and information with that definition:

> Entropy … corresponds to the amount of (microscopic) information that is missing in the (macroscopic) description.[3]

Entropy, says von Neumann, is a lack of information about the system. We have microscopic information when we can identify each ball (e.g. with colour), macroscopic when not (all balls look the same). The less we can discern different microstates, and aggregate them into less informative macrostate, the higher the entropy.

Let us put this definition together with the one given by James Maxwell (1902), one of the founders of thermodynamics:

> The greater the original entropy, the smaller is the available energy of the body.

---

[1] Boltzmann defines his famous definition of entropy, engraved on his tombstone as a fitting epitaph: $S = k \bullet log\ W$, where W is the number of microstates equivalent to the macrostate observed and S is the dimensional constant, known as the Boltzmann constant.

[2] von Neumann, a Hungarian, one of the many Jews who fled from inhospitable Europe to America in the 1930s, was considered the most gifted man in the world by the likes of Enrico Fermi, who after working with him on a problem, said "I felt like the fly who sits on the plough and says 'we are ploughing'" (Schwartz 2017).

[3] The complete quote reads: "The closeness and the nature of the connection between information and entropy is inherent in L. Boltzmann's classical definition of entropy … as the logarithm of the "configuration number". The "configuration number" is the number of a priori equally probable states that are compatible with the macroscopic description of the state – i.e. it corresponds to the amount of (microscopic) information that is missing in the (macroscopic) description"

Maxwell says that low entropy means being able to extract energy. Von Neumann that low entropy means having information about the system. Therefore, when we have information about a system, we are able to extract energy from it.

If we think about that, it is quite obvious. If we want to extract wealth from the stock market, we need to study it. We need to be able to know how it evolves. If we want to extract energy from a liter of fuel, we need to know the laws of thermodynamics, so that we can build an engine.

In order to get energy, we, like any other living system, must have information about the environment. This allows us to absorb the energy which allows us to escape equilibrium.

Having defined life, we ended up with the idea that living organisms are systems which collect information about the environment. They use this information to extract energy from the environment and keep themselves in a state far from equilibrium. Before asking ourselves why they do so, we need to define information.

Defining Information

If we want to study living systems, which store information on how to extract energy from the environment, we want to have a clear definition of what information is.

> Acquiring information on a system means becoming able to predict how that system evolves with less uncertainty than we did before.

For those keen on a bit of mathematics, below we define uncertainty as a function of probability, and information as a function of uncertainty.

To do this, all we have to do is define the level of surprise for an event. Surprise is a function of the probability p*, where* p *indicates how strongly, in a 0–1 range, we believe that an event is going to happen. Surprise should therefore be big in the case of a small* p *(we are very surprised if something we think as improbable happens) to zero in the case of* p = 1 *(we are not surprised if something we consider to be inevitable happens).*
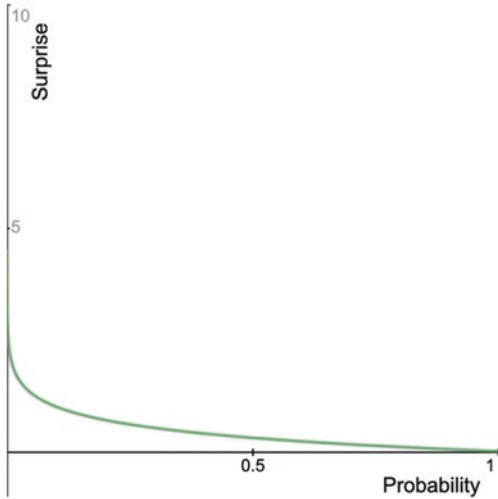
For those interested in a bit of math, the function that satisfies this relationship between p and surprise is the logarithm. As in Shannon (1948), we consider the logarithm in base 2:

$$Surprise = -log_2(p).$$

From this definition, we define uncertainty as the average surprise. Which makes intuitive sense: if we are surprised very often of what happens around us, it means we don't know much about the world we live in.

To understand the concept, we can take a very simple system –a coin.

In a coin, head and tail have the same probability. Let us imagine that for some reason we believe the coin to be biased. We believe that heads comes up 80% of the time and tails 20%. This means, each time head comes up our surprise will be

$$surprise(heads) =- log_2(.8) = 0.32 \ bit$$

and each time we see tail:

$$surprise(tail) = - log_2(.2) = 2.32 \ bit$$

Because the coin is actually not biased, we will have a surprise of 0.32 bit 50% of the times, and of 2.32 bit the remaining 50% of the times. On average, our surprise will be

Average_surprise(we believe biassed coin) =
0.5 • 0.32 bit +0.5 • 2.32 bit = 1.32 bit.

If we had believed that the coin was fair, as it was, our surprise for both head and tail would have been

surprise(head or tail) = − log₂(0.5) bit = 1 bit

average surprise would have been lower:

average_surprise(we believe fair coin) = 0.5 • 1 bit +0.5 • 1 bit = 1 bit

This will always be true: the average surprise is minimum when the probability we assign to each event is actually the frequency with which the event will happen.

More formally, we can say that if the system can be in N possible states, with an associated probability of p_i and a frequency of q_i, our uncertainty S for the system is

$$S = \sum_{i=1}^{N} - q_i \cdot \log_2 (p_i)$$

According to the Gibbs' inequality, S has its minimum for $p_i = q_i$, i.e. when the probabilities we associate to each event, p, is the one we will actually observe, q.

In this sense, acquiring information on a system means knowing how to predict the frequency of each result.

If we have a good model describing the solar system, we'll be able to predict the next eclipse and not be surprised when the sun will disappear.

A lion – a carnivore who is one of the laziest hunters in the animal kingdom – like any other living system works on minimising uncertainty, to get more energy (food) from hunting. Of various paths used by prey to get to a water hole, the lion studies which are the most probable (Schaller 2009). The lion minimise the uncertainty on the prays' path, and therefore increases the probability of extracting energy from the environment.

Note that there is no such a thing as absolute information. While the frequency with which events happens are not observer-specific, the probability we associate to them are. As Rovelli (2015) writes: "the information relevant in physics is always the relative information between two systems" (see also Bennett 1985).

Information Storage Requires Energy

If the pages of this book looked something like this
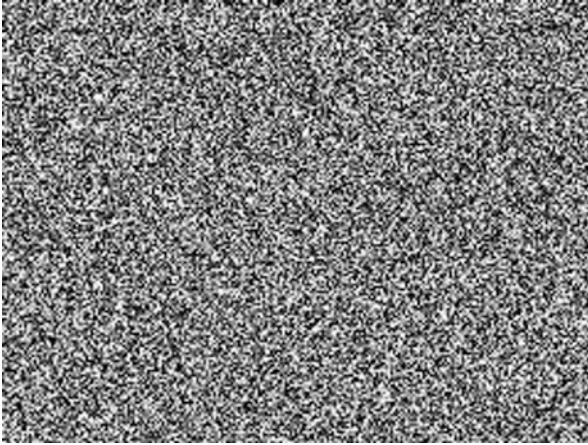
few people would believe there was any message at all –rightfully, because this is just a randomly generated image.[4] It looks like the ink was just spread, randomly, into this state and then dried.

Viceversa, the Jiahu symbols, drawn in China 8600 years ago, are widely seen as one of the first example of human writing (Li et al. 2003).

Our intuition tells us that someone made an effort to create signs with a meaning, and that this drawing is not random. We don't know how valuable the information was, but we do know that someone wanted to communicate something, to pass information.

---

[4] https://upload.wikimedia.org/wikipedia/commons/f/f6/White-noise-mv255-240x180.png

To store information we need to set the system in a low-probability macrostate.[5] If we use as a communication tool a box with 10 balls and 9 walls, it would not be wise to assign a meaning to the state where all balls are evenly distributed: the system would often decay, if not controlled, to this state. A recipient would see nothing more than the "normal" state of the system.

This means that **information also needs energy**. Because we use low-entropy states, we need some energy to avoid the decay to equilibrium of the memory.

Random Access Memory (RAM) in a computer, for example, requires continuous energy input in order not to lose the information stored. Our brain dies in a few minutes without energy, and so on.

~

The last two sections bring us closer to understanding how a living system might work. They tell us that in order to absorb energy, the living system must store information about the environment. And that to store information, the system needs energy.

Information and energy

Storing Information

We took as examples of information-storing devices writing, computers' memory and the brain. Only the last is (part of) a living organism.

---

[5]The opposite is not necessarily true, and there are indeed drawings, which we will analyse later, which are not universally recognised as symbols. This, because sometimes a system can settle itself in an apparently not-random state, to which we can assign a meaning, although no information was put there. For example, Italy, one of the largest exporters of footwear in the world, has the shape of a perfect boot. Nonetheless it's hard to believe that someone in the universe created the peninsula with such a shape to advertise the Italian shoe industry.

Nonetheless, for us humans, information is linked to language. We use language to store information, and to exchange information. Is the way we store information in language so different from how we living beings store information? Not really.

If we are asked about the structure of a brain, the first thing which comes to mind is probably "a network". Language is no different, as it is a network (Wood 1970). When we express concepts, we connect words. The information is stored by identifying patterns inside the network of language.

We could think of the book as a path in the network of the 6837 unique words[6] used to write it:

"The" ⇒ "amazing" ⇒ "journey" ⇒ "of" ⇒ "reason" ⇒ "\new_line" ⇒ "from"… and so on.

One characteristic of language is that part of the information is in the microstate – each word is different from the others– and part of it is in the macrostate –words have different meanings depending on which words they are connected to. Therefore part of the information is in the nodes of the network, part in the network itself.

What also is of interest, is that the complexity of the language network emerges as we use the language. The more we use language to describe new concepts, the more connections appear, the more complex the network becomes. One hundred years ago the word "so" was never used before the word "cool" and after "ur". Today, with usage, a new path has appeared.

Something similar happens in the brain. The memorisation process in the network of the brain occurs connecting nodes: "Pathways of connected neurons, not individual neurons, convey information. Interconnected neurons form anatomically and functionally distinct pathways" (Kandel 2013). As for language, information and complexity are inseparable in the brain.

Though the process is complicated (the human brain has various kinds of memory: short, medium, long-term, spatial, procedural), the memorisation mechanism involves the definition of new paths in the neural network, called Hebbian learning after the psychologist who came up with the theory in 1949.

Donald Hebb understood that memorising the association between two events (like in Pavlov's experiment, in which a dog associates the sound of a bell with food) occurs when two neurons fire several times simultaneously, so the activity of the same often appears to be related: the connection between these two neurons is reinforced, leaving an information trail (Kandel 2013).[7]

In the case of long-term memory, the link is definitive, in the case of short-term memory it isn't, and only remains if used frequently (by neurons that fire periodically). In general, with experience, the brain reinforces certain connections ("sensitisation through a heterosynaptic process", Kandel 2013) and weakens others ("synapses that are weakened by habituation through a homosynaptic process)". The less-used connections disappear, while the more-used ones are consolidated. Exactly in the same way as with language.

---

[6] Case-sensitive

[7] Specifically, "E3 – Mathematical Models of Hebbian Plasticity"

To sum up, networks are complex systems where their complexity, their internal structure, arise from storing information. In this sense, every network which stores information can be considered complex according to Herbert Simon (1962) definition of complexity: "Roughly, by a complex system I mean one made up of a large number of parts that interact in a nonsimple way. In such systems, the whole is more than the sum of the parts".