Life & Complexity

Levien van Zon (levien@gnuritas.org)

This is the text of a talk I gave at the Chaos Communication Camp 2007.
For slides and more information see: http://events.ccc.de/camp/2007/Life_&_Complexity

Life & Complexity. You may wonder, what does this have to do with security? Well, it has nothing to do with security. It doesn't have a lot to do with hacking. It does have a bit to do with computers. But mostly, it has to do with yourself, and with life in general.

I want to address the fundamental question: why is there life, and why is it the way it is. If you think about it, it is strange that we know more about the state of the universe a picosecond after the Big Bang than we know about this curious creature, a *cellular slime mold*.

Of course, one can always invoke one's deity of choice to explain the existence of life. I will however skip over that type of explanation here, as it has been invoked many times before and you can read about it elsewhere, if you wish.
I do however believe that if God had created the world in 7 days, He would not have had much time for debugging, which would explain a great deal.

Next, you could ask your friendly neighbourhood Darwinist why we're here, and he (or she) would probably say that it's all due to evolution and natural selection, which of course still doesn't tell you why there is something to evolve in the first place.

So, you go to a bookstore and pick up a cutting edge science book. And it would tell you that it's all due to emergence, and that it's all very complex. It will then probably continue about how butterflies cause thunderstorms and the way in which ants look for food, and in the end you still wouldn't know why we are here and how we got here.

And then you could ask me, and I would say “I don't know. But... I have a nice story I can tell you. It won't explain everything, but it might give you a glimpse of an answer.”

I will conveniently start at the beginning, about a picosecond after the Big Bang (which doesn't really go “Bang”, as there's no matter yet). It takes a few minutes for hydrogen and helium nuclei to form, then 380,000 years for atoms to form, and during the next billion years or so small perturbations get amplified and grow into the first galaxies and stars. As these stars die, supernova's eventually produce the heavier elements that make up the rocky (and mushy) bits of our Solar System, and after about 9 billion years our Sun and its planets have been formed from a large cloud of gas.

As we shall see, the energy from the Sun is to become very important in sustaining life later on. But as far as many scientists now believe, life first originated around volcanic vents at the bottom of the early oceans, around 4 billion years ago. These vents were probably very similar to the so-called Black Smokers we still find at the bottom of the ocean. The vents provide plenty of sulfides that yield energy when broken down, as well as other molecular building blocks and a relatively protected environment.

Now, with so much energy and so many building blocks available, it is not hard to imagine the appearance of the first simple metabolisms: chains or networks of chemical reactions that are driven by breaking high-energy chemical bonds. As these reaction networks became more numerous and complicated, some might have arisen that could catalyse the formation of their own constituents. In other words, in theory you could get a set of reactions that “eat” high-energy molecules and that can produce itself from the available building blocks. We call this an autocatalytic set, and you could consider it to be the most basic form of chemical “life”.

Well, once you have a bunch of these self-reproducing chemical systems, you can get amplification, competition, natural selection, and with it fairly rapid evolution. It is thought that after a while several of such systems arose, some based on proteins, some based on RNA or DNA, and probably lots of others we don't know about because they have long since gone extinct. Some of these reproductive systems eventually got linked, and then got enclosed in a protective bubble of lipid molecules, and so the first living cell might have been born.

As I told you, the first cells were driven by chemical energy derived from “eating” certain molecules, but shortly afterwards cells must have appeared that were capable of using sunlight, in a process called photosynthesis. These organisms literally changed the planet. They are responsible for the oxygen atmosphere we have today, and photosynthesis is still the process that provides the energy to support almost all ecosystems on earth, including most of human society.

But... It's a long way from the single-celled organisms that arose 3.5 billion years ago to human society. In fact, the first multi-celled organisms didn't arise until about 1.4 billion years ago, and the first mammals (and marsupials) appeared “only” 0.2 billion years ago. So how do you get from a single cell to something like a Wombat? That's not exactly a trivial question. Let us start by looking at the basic ingredients of life and self-organisation in general.

As it turns out, self-organisation is the easy part. This may seem a bit counter-intuitive, as the Second Law of Thermodynamics states that disorder (or entropy) always increases. But that is only the case for closed systems. The Earth is an open system, it produces heat internally and it receives energy from the Sun. Lots of energy. And as long as this is the case, it will not reach thermodynamic equilibrium (the state of maximum entropy), which means that stuff will happen. What kind of stuff? Well, in the presence of a large energy gradient, matter will start to organise itself to get rid of this gradient. You will get circular flows of matter that transport (or *dissipate*) the heat. This is better known as *convection*, and these flows then organise themselves into hexagonal structures known as *Bénard convection cells*. Such convection cells are a basic example of dissipative structures, that is, structures that spontaneously form to get rid of an energy gradient. And this, in turn, is an example of self-organisation.

Now, instead of being dissipated by convection, solar or geothermal energy can also be dissipated along other routes. It can, for instance, be stored in chemical bonds between atoms, and then slowly be dissipated though a series of exothermic reactions. Which, not coincidentally, is exactly what happens in living systems. In other words, you could see life as a very large and complicated set of chemical reactions that dissipate solar energy into heat.

The opposite of disorder is of course order, and order represents information. This is why information is sometimes called negative entropy or “negentropy”. An energy gradient (or an *energy potential*) is an ordered state (– it has a high side and a low side), and it therefore contains information. This “thermodynamic information” is slowly lost as the gradient is dissipated. But, because coupled chemical reactions can transfer free energy, the thermodynamic information contained within the energy gradient can be temporarily converted into other kinds of information. This is the principal driving force behind the formation of complex structures and processes. It implies that structure and “complexity” can arise and exist as long as long as there is an input of energy. But without energy, the structure and information contained in it is slowly (or quickly) lost. Just think about what happens if you power down your computer. Or if you stop eating for a few weeks.

As the famous physicist Erwin Schrödinger put it: “Thus a living organism continuously increases its entropy – or, as you may say, produces positive entropy, which is death. It can only [...] keep alive, by continually drawing from its environment negative entropy, which is something very positive as we shall immediately see. [...] the essential thing in metabolism is that the organism frees itself from all the entropy that it cannot help producing while alive...” (From: E. Schrödinger, What Is Life? 1944).

So, looking at it this way, living systems and ecosystems are basically just very complicated cycles of matter, driven by the huge energy gradient created by the sun. The matter is constantly being recycled, and it's just temporarily stored in living structures. Thermodynamically speaking, it's not so strange we have life, and complexity.

So energy is essential. But there are also other processes that shape the patterns we see around us, including those of life. Probably the most important of these are feedback processes. Especially positive feedback. As most of you will know, positive feedback is when a process has a positive influence on itself, so that it keeps amplifying itself. Through the amplification of randomness or small differences, positive feedback is responsible for things ranging from the formation of galaxies to the process of natural selection. And just as positive feedback is useful for amplification, negative feedback can keep things from getting out of hand. Negative feedback is when a process has a negative influence on itself, so that it slows down or stops. Feedback processes seem to occur quite naturally and therefore frequently, and they are responsible for much of the non-linearity in nature.

The combination of positive and negative feedback can produce a wide range of patterns. In time, it can produce oscillating patterns and threshold behaviour such as nerve pulses (*action potential*) and memory effects (*hysteresis*). Much of this was first formalised in 1952 by Alan Turing, whom I expect most of you will know. He worked with a combination of mathematical equations, that basically represent positive and negative feedback in space (i.e. with diffusion). In two or three dimensional space this makes that you can get static or oscillating so-called Turing patterns (spots or stripes), as well as moving waves or spirals. And because feedback processes are very general, you see such patterns everywhere in nature.
Spatial patterns like sand-dunes are caused by feedback. Tropical cyclones are a positive feedback pattern, as are Ice-Ages. The spots and stripes of some organisms may be created by an combination of feedbacks. Neural patterns are created by feedbacks, and the “wiring” of the average gene regulation network (that is, the network of interactions between genes) includes thousands of feedback loops. The development process of plants and animals, for as far as we understand it, makes heavy use of feedback processes. Even the establishment of dominance hierarchies in societies of animals ranging from bumblebees to chimpanzees is the result of feedback. Not to mention the 'self-organising' foraging behaviour of ants, which also has an important positive feedback component.

So, just as energy provides the thermodynamic driving force behind structure and organisation, feedback processes provide one of the main pattern generating mechanisms. In fact, natural selection itself can be considered just another example of amplification by positive feedback. And feedback helps shape many of the processes and structures on which evolution can act. But still, we need a bit more if we're to explain multicellular organisms. Or, as Alan Turing himself is once said to have remarked, regarding the stripes on a Zebra: “The stripes are easy, but the horse part is harder to explain.” So in other words, pattern formation is a bit more than just just feedback.

The next piece of our puzzle involves the relationships between *replicators* (things that replicate/copy themselves). I guess you could call such relationships feed-forward processes, although that term usually means something a bit different.

The nature of dissipative systems makes that food chains are a natural thing to have, as energy can be degraded along the chain (or usually a web, but let's keep things simple for now). There are of course the familiar chains, in which plants capture sunlight, are subsequently eaten by animals, which are eaten by yet other animals and so forth, until the last animals in the chain die and are consumed by decomposers. At each step in the chain, around 90% of the energy is lost as heat, illustrating the dissipative nature of this chain.

Besides these trophic relationships ("eating and being eaten"), there are competitive relationships, for instance competition for space or food. Such relationships are the main drivers behind natural selection, as the fastest-reproducing lineage or species dominates the other ones over time (i.e. they cannot coexist in equilibrium).

Finally, there are the "other" relationships, which biologists like to give difficult names like mutualism, commensalism, amensalism, and a few other -isms. These basically include all ways that one organism can influence or depend on another, other than for food. And believe it or not, these are actually interesting, for several reasons. The first reason is the fact that such relationships are not specific to the usual ecosystems containing plants and animals and stuff. You can have such relations between chemical reactions. You can also find many such relationships in, for instance, communities of bacteria. They even arise in communities of digital organisms. A nice example is the well-known "virtual world" Tierra (and its derivatives, such as Avida). Tierra was created by the tropical ecologist Thomas Ray in the early 90's . Tierra is basically a virtual machine, in which small assembly programs can evolve and compete for CPU time and memory. Within a few thousand generations you get all kinds of dependence relationships between the programs, including parasites, mutualists and hyper-parasites. When mutation is turned off, they even invent a kind of "sexual reproduction"! So the kind of relationships and processes you see in living systems, can also appear in systems we don't usually regard as "living".

At least as interesting are the simulations that are performed by Nobuto Takeuchi at our group in Utrecht. He is trying to simulate the early stages of molecular evolution by mutating virtual strings of RNA, and then calculating their two-dimensional structure and letting them interact. Even in such a "simple" system (biologically speaking), he already gets a basic "ecosystem" of two coexisting "species" of RNA-catalysts, as well as two types of parasites. An that's with a "stripped-down" version of RNA – the real molecules also have a 3-dimensional structure and can do a lot more.

So, in evolving systems that offer the possibility of relationships between "things" that replicate, you will quite easily get all kinds of interactions and relationships. And each interaction may create possibilities for new interactions. This may lead to quite complex networks of interactions that "grow" more or less "automatically".

Besides being a natural form of complexity, there is another reason why these relationships are interesting: They can help explain the increasing (mostly hierarchical) complexity in replicating systems over time, such as the increasing complexity of single-celled organisms and the appearance of multicellular organisms. The basic mechanism is that some relationships may prove beneficial, which means that natural selection can act on them and they can get fixed over time. In effect, the interactions between replicators create a new level of organisation, and with it a new "level of selection". I will give some examples.

A very concrete example of this is in our cells, and in the cells of plants. These so-called "Eucaryotic" cells have all kinds of "specialistic" structures inside to perform different tasks. For instance, there is the cell nucleus that holds most of the DNA, the mitochondrium which is involved in oxidizing other molecules to provide energy, the chloroplasts which do photosynthesis in plants, and all kinds of so-called plastids that are used for storing stuff. There is genetic evidence that all of these structures (or *organelles*) were once separate single-celled organisms that were specialised at some function.

Once they started living together, this provided an advantage to the entire complex, and the collection of organisms eventually became a new organism. So although selection initially acted on a number of individual cells, the reproduction of the *collection* of cells became more important over time. This created a new level for natural selection, which after a while turned into a new level of organisation.

It can also be that another level of selection arises as a more unexpected side-effect of interactions. A nice example can be found in a computer-model of Maarten Boerlijst and Paulien Hogeweg (1991). They simulated a so-called Hypercycle in a spatial grid. A Hypercycle is a cycle of self-reproducing molecules that also catalyse each other. When you allow these molecules to catalyse each other on a spatial grid (in this case a cellular automata grid), they will form a pattern of spiral waves. And it turns out that these spirals then start to compete for space. “Who has the biggest spiral” is then more important for the reproducing molecules than their own reproduction speed. In other words, selection is no longer on how well these molecules catalyse their own formation, but instead is on how well their spirals rotate (“reproduce”) compared to other spirals. Of course this still depends on the properties of the catalysts. But the spirals have formed a new level of organisation and selection. So you can have multiple levels of selection and organisation, even in such a limited and abstract system.

You may have noticed that the creation of new levels (of organisation and selection) will lead to a more or less hierarchical form of organisation, in which each organisation level is composed of a collection of simpler entities, which in turn are formed from even simpler entities, and so forth. All the levels influence each other, and each level adds more complexity, and more things that can be modified by evolution. In fact, this hierarchical organisation and increasing complexity is exactly what we see in nature.

Let's return for a moment to the replicating molecules around 4 billion years ago. As I told you, a combination of replicating molecular systems probably formed the first cells, 3.5 billion years ago. Fairly soon, these started to live together in filaments, communities and even more “intimate” forms that finally merged into a single eukaryotic cell around 2.5 billion years ago. By 1.8 billion years ago the internal complexity the genetic system in cells had greatly increased, allowing for much more complicated dynamics at the genetic level, and also for sexual reproduction (the “mixing” of genetic material from two cells), which greatly speeded up evolution. By 1.4 billion years ago the first truly multicellular organisms had appeared, which consisted of a few different (eukaryotic) cell-types that had become specialised for different tasks. Shortly afterwards, these started to form bigger structures with more cell-types, capable of ingesting food and sometimes capable of organised movement. By 600 million years ago multicellular life had evolved into a diversity of forms, similar to today's jellyfish, sea anemones, worms and molluscs. These organisms lived off smaller organisms, and sometimes off each other. Then came the so-called Cambrian Explosion, in which hard-bodied organisms appeared, and after which multicellular life diversified into thousands of new crab- and insect-like species. Around 425 million years ago, the first fish and land-plants had appeared, quickly followed by the first insect- and crab-like land animals. It didn't take long (less than 100 million years) for the fish to crawl out of the water and evolve into the first amphibians and then reptiles, which branched off into mammals and birds around 200 million years ago. The reptiles went on to evolve into dinosaurs, which did pretty well for a while but got hit by a meteorite the size of a small city, 65 million years ago, after which the warm-blooded mammals and birds took over. Around 50 million years ago the first primates appeared, which then went on to form social structures, evolve language, invent agriculture, cities, writing and eventually nuclear weapons, global warming and Windows Vista.

Now, what is interesting in all this, is that you can still see some of these evolutionary stages back in our development from egg-cell to newborn baby. One would almost forget, but even the most complex multicellular organism starts off as just a single cell. And we have to get all the way to being multicellular and complicated, so somehow part of our evolutionary history is repeated in “fast-forward” during our development. Each developmental stage building on and extending the previous stage. Of course we're not actually repeating evolution, because the information for doing this is stored in our genome, a long string of molecules that can hold just under a Gigabyte of information. A complex life form can apparently be compressed into a few million bits of pure information, and some chemical and physical machinery that translate this information into chemical reactions.

Now I'll be honest with you, we don't really know how this works. We do know that the properties of the cell are important, as it's the basic building block for all multicellular organisms. The more complex the organism, the more types of cell it tends to have, each specialised for a certain task. And it's of course the interactions between cells that makes that certain tissues and organs are formed. And these tissues and organs can then again be modified by evolution to suit specific purposes.

We also know that the dynamics of the gene network are important. The genes themselves code for proteins, which then partake in chemical reactions that can for instance alter the behaviour of certain cells. But the genes also interact with each other, in a complex network of feedback and feed-forward processes. This creates a kind of genetic program, in which certain genes can amplify or suppress other genes. And this network can quite easily be modified by evolutionary processes, simply by duplicating or removing genes, and then changing these genes and their interactions. And finally, the interactions between cells are to a large extent dictated by the genetic program, but the interactions between cells also feed back on the genes by causing certain genes to be suppressed or amplified. In other words, the organisation levels of cells and tissues and the gene network all influence each other, and there is a kind of local communication between individual cells. There is also a more global form of signalling, provided by messenger molecules, which can convey messages over longer distances. The substances we call *hormones* are an example of this.

Now we are actually starting to understand how the elements I just mentioned can lead to a bunch of single cells forming a more complex organism with more than one cell-type. It is actually possible to simulate parts of this process. I would like to show you two computer models which are at the level of biological cells, and are based on the so-called *Cellular Potts-Model*. This model simulates cells as a collection of smaller cells on a two or three-dimensional grid, which follow a set of physical rules such as volume conservation, surface energy minimisation, deformation under pressure, etc. This way they can interact with each other and with the medium, in many ways these cells behave as real cells. And the nice thing is, you can add rules, to represent for instance parts of a gene network, or chemical reaction pathways, or extra physical constraints.

So what Paulien Hogeweg did, is she took this cell-model and she added a network of binary operators, which represent a simple gene network. This network then determines the adhesion between cells, so the “genes” influence the interactions between cells. But also the other way around, because communication with neighbouring cells is provided through two communication “genes” (nodes), so the interactions between cells can also cause bits of the “gene” network to get switched on or off. If you then start with one cell, add some randomness and let it grow and divide, you will end up with some multi-cell shape, and because of the interactions you get different cell-types, with different adhesion properties. And if you then let the gene-network evolve, and you select shapes for the maximum number of cell-types, you end up with shapes that actually resemble tissues and structures found in real organisms! Moreover, it turns out that many evolutionary paths lead to the same kinds of shapes, so you see frequent re-invention of the same shapes, something you also see in nature.

It's great that you can get these shapes by simply simulating some of the processes that we think are responsible for creating more complex organisms in nature. But of course, these shapes don't actually do anything, because we didn't select for that. But it is possible to get collections of cells to perform a function.

What I will show you now, is a very curious creature called a “cellular slime mold”. This species is called *Dictyostelium discoideum*, and it's basically a single-celled amoeba, similar to the first Eukaryotes. It happily lives in the soil, eating bacteria, but if there is a shortage of food something strange happens. The amoebae start to aggregate and together form a sort of small mound surrounded by slime, which then gets up and crawls away. This “slug” can then migrate over quite a large distance (a few centimetres), usually to the soil surface. When it gets to the surface, it stops and grows into a kind of stalk, with a so-called fruiting body at the top – it becomes sort of a mushroom. The fruiting body contains spores, which are spread through the air, and if they land somewhere they turn into amoebae again.

It's a crazy creature, but it turns out we can actually simulate it! Or a part of it anyway. The model is created by Stan Mareé, Paulien Hogeweg, and Nick Savill, and is again an extended version of the Cellular Potts-Model, but this time without a gene network. Instead, only a few chemical reactions are simulated within the cells. These reactions are known to be present in the real slime molds, and mostly involve a feedback mechanism that reacts to a substance called *cyclic-AMP* (*cAMP*). This reaction is influenced by temperature and light, which allows the organism to react to its environment. It turns out that the amoebae start to excrete cyclic-AMP when they cannot find enough food. This is then amplified by the other cells in the vicinity, and this positive feedback creates a spiral-wave pattern in space. The cells go toward the centre of the spiral, which has the highest concentration of cAMP, and there they form the “slug”, which has at least three different types of cells. The cAMP production causes oscillatory behaviour in the slug, which causes waves of cAMP through the slug, which causes pressure waves, causing the slug to crawl. Light or temperature differences cause the wave-front to be slightly slanted, and the slug to change direction and crawl toward light or a higher temperature. The pressure waves that cause the movement also cause the three cell-types to be nicely sorted out. The formation of the stalk is driven by the same kind of cAMP waves, which cause pressure waves that push one cell-type (the *stalk-cells*) downward, and because action is minus reaction, the rest goes up, and we end up with a nice fruiting body with spore-cells at the top.

These aren't exactly high-tech graphics, but I was completely amazed when I saw this for the first time. And it nicely shows that cooperation between cells and complex behaviour can quite easily “emerge” from a few basic physical and chemical processes, and interaction within and between multiple levels of organisation.

It also shows the importance of computer simulations for understanding these processes. It's not enough to just study the elements of a complex system at the various organisation levels, because most of the behaviour is actually generated by the interactions! In the computer you can just take a number of these processes and interactions, simulate them and see what happens.

Unfortunately most biologists are not very good programmers, so simulations of the type I just showed you are fairly rare. Also, people – also scientists (often *especially* scientists) are a bit afraid of complexity. They don't like it, as it's difficult to model and difficult to analyse. You see this in biology, but also in economics. Scientists prefer to stick to nice clean mathematical expressions, and try to explain everything in terms of equilibria and optimality. This works fairly well for the basic physical and chemical processes, but you can't really get the type of “messy” (but realistic) behaviour I just showed you. Moreover, nature is often not “optimal”. People talk about “survival of the fittest”, but what *is* the fittest? Is it the one that grows fastest? It can be, but usually it is also the one that doesn't immediately die as soon as the environment changes. And the environment, be it physical, biological or social, changes *all the time*. That's why evolution not only selects for growth, but also for diversity, for adaptability and for robustness.

People talk about fitness landscapes, and getting stuck in local optima. But fitness landscapes are actually seascapes, and *real* systems are in local optima all the time. There may be global optima, but they move all the time, and real evolution simply doesn't get there because it has to work with what's available. And that's great, because an “optimal” world would be very boring! You would have only one species – a very bored species because sex wouldn't exist either, that would be just a waste of energy.

Luckily we don't live in an “optimal” world. We live in a messy and complex world, in which evolution is driven by accumulating levels of organisation, and the accumulation of information. A world in which complexity creates even further complexity, because every new development creates a whole range of new possibilities. This positive feedback loop makes that evolution is actually speeding up at an almost exponential rate. It took over 2 billion years to get the first “modern” biological cells, but only about half a billion years to get multicellular organisms, and within another billion years this evolved into almost all the groups of plants and animals we still see today. It took only about 20 to 30 million years for most modern groups of mammal evolve from a mouse-like ancestor, and humans evolved from apes in less than 5 million years. The big leap forward for us came with the evolution of language, around 300,000 years ago. More or less modern societies emerged mostly after the invention of writing, or at least bookkeeping. The invention of printing eventually made education and scientific knowledge available to millions of people. And we all know what the introduction of computers and modern communication did to the speed of technological development. It seems to be the accumulation and exchange of information that drives evolution, first biological and now cultural and technological. And this process is in turn driven by a steady supply of energy. And who knows where it will end. Perhaps in chaos, perhaps not, I don't know.

What I do know is that it is important to gain a better understanding of the role of energy, information, feedback mechanisms, organisation levels and all kinds of interactions, in the creation and maintenance of complex systems. And I hope I've been able to show this to you.

It would seem that life is a bit more than just the sum of its parts. In fact, you could argue that life is defined more by the process of energy dissipation and especially by the interactions between its parts. And these processes and interactions occur not just in living systems, but all through nature and the Universe, and also in society, economics and all kinds of man-made systems. That explains why we see the same patterns in all off these systems. So I would say, it's time we started looking more at the interactions, instead of just the studying the parts.