# Entropy Maximisation, Life and Living Engines

## Video Transcript

Ok so, forget everything you think you know about entropy.

But this video isn't even about entropy: I'm not going to explain to you about what entropy is. In essence it's actually, I suppose, about living things like us, and the way that living things like us perhaps maximise our entropy (whatever that means); maximise our *irreversibility* (and we'll get into what exactly that might mean).

You see, this year I've been having a lot of trouble explaining an idea. And it's the kind of idea that I think moves in a lot of different ways. I think it can used to understand our own cellular biology and evolutionary origins. I think that it could be useful in considering the environmental consequences of life present in our geological record, and in anticipating the changes that will now be happening on the surface of this planet in the near future. I also believe we could use this thing in software engineering, and in the designing of particular kinds of computation machines, perhaps even in providing insight on some of the fundamental statistics within our brains.

But alright, okay, those are some very big claims and, to be fair, this video isn't even about any of those things particularly either. For now, please allow me to slow all this down. Let's just settle in, and listen, please, because honestly this has all been a bit of a bother to explain. If I could take you alongside for the last odd year of my life, where I've tried to find ways of talking all this through in a way that actually rounds up nicely. Because you see, since before my graduation last summer,[1] I've had this thing ringing in my ears, and it's been trying to find form, and little by little I've found ways to lay it out a bit, and now I think, there should be enough here for

[1] MASt in Physics from the University of Cambridge, and before that I received my MPhys in Theoretical Physics from the University of St Andrews.

some sketched sense at least of what it's all about. So that's where this story will start now. Let's step back to last summer and, well I suppose, let's see where this takes us.

So entropy is a complicated quantity, and it is often confused with order. But, I think I want to make the case to you that actually you shouldn't really be thinking about entropy separately to thinking about engines and the entropy associated with engines.[2]

You see, there are two different types of engines.[3] On one hand you have the engine-making that's involved with moving energy around, like the steam engine, or the engine in your car... ...on the other hand, you have engines that are more probabilistic, that are trying to build models of their environment.

And we'll talk about both of those two types of engines, but before we do I just wanna tell you about two things. The first is what an engine is, as far as I'm concerned. So when I say engine, I want you to think about some entity that enacts a precise sequence of states, in a very particular order, such that the state that it starts in, is also the state that it ends in and so it behaves cyclically, in a cycle. So when I say engine I want you to think of just some cyclical sequence of states.[4]

[2]So when I say "forget everything you think you know about entropy" I really mean it. The concept of entropy only came about in the first place through Clausius' analysis of the Carnot Engine, [1] and yet discussions on the interpretation of entropy typically take the quantity outside of its original context, and frequently apply supposed equivalence to the notion of 'disorder'. I would direct you to the work of Wright [2] and Lambert [3] who discuss various specific examples of the analogy breaking down. Despite applying well in certain cases, such an equivalence is frequently abused.

[3]This is a big simplification used for the purposes of this video. The idea I'm trying to lay out is that there are these two different ways of thinking about engines and entropy, and that it might be possible to unite them in some way when applying them to living things.

[4]I understand that this a fairly extreme simplification. You might be asking: What about fuel or gear-changing? Or adiabatic and isothermal surfaces? What about expansion of gas? Doesn't 'engine' mean different things in the context of computation and thermodynamics? For now just forget about all that and consider the principle property of an engine as the fact that it ends a series of distinct states and process in the same place that it started, ready to cycle all over again. It's think it's fine to be skeptical but I believe that this way of framing the problem is extremely helpful in the application of thermodynamics to a wide variety of interesting non-equilibrium problems.

Now the second thing I want to tell you about is something that actually I don't care about. It's called the Carnot Engine. [4] Now, if you are doing an undergraduate in physics then you'll know about the Carnot Engine. It's the engine you get taught in any thermodynamics course. It is the perfect engine, it is the most efficient engine; Victorian scientists came up with it a few hundred years ago.[5] Now this engine is so efficient precisely because it is reversible. And its reversibility, in entropy terms (and we'll talk about entropy more and more and more but this is I want the first thing I want to tell you about entropy), in entropy terms this engine is so efficient because it's entropy is constant across all of the cycles of sequence of states that it undergoes. So this Carnot Engine is efficient because it's entropy is not changing. And that makes it easy to think about mathematically, and it makes it very interesting when you're trying to design steam engines or more mechanical motors,[6] but it fails when it comes to talking about living organisms that we see in nature. And, why is that? Well, ultimately it's because the Carnot engine is boring. It's simple: it's in equilibrium, it's closed from its environment,[7] and it is reversible which means its entropy is not changing. And none of those things apply to living organisms. They are constantly in dialogue with their environment (they depend on their environment to persevere),they are not in equilibrium, they are out of equilibrium (but perhaps we'll talk about that further in a future video)[8], and their entropy is by no means constant.

So the language of thermodynamics that's very well established can't be applied to living organisms that we see in nature. But there is, perhaps,

---

[5] Sadi Carnot was, of course, French, but his work was developed by the likes of Lord Kelvin and James Joule during the Victorian Era in Britain.

[6] The Carnot Engine is, of course, *idealised.* In real-life situations efficiency is not perfect and entropy will always increase (although preferably as little as practically possible).

[7] Some sources consider a steady-state Carnot Engine where liquid vapour enters during one step, and steam exits during another, allowing mass exchange during individual steps (implying an open system) however when all individual process are taken as a whole cycle, no net mass is transferred across the system boundary and we effectively have a closed system.

[8] Living organisms are primarily, if nothing else, extremely effective dissipators of energy, thriving far away from equilibrium. Part of the difficulty in modelling life from a thermodynamic context is that the vast proportion of our understanding of thermodynamics applies only to so-called 'boring' *equilibrium* systems. This is the most important difference out of the three, because for living things, you need a fundamentally different way of thinking about a thermodynamic engine that hovers far away from equilibrium.

some semblance of a solution that we might be able to pursue.[9]

And such a solution lies, ultimately I think, in the unification between these two different types of engines that I alluded to earlier. The, the heat engine (although hopefully something a little bit less limited, less boring than that Carnot Engine) and the inference engine (which I'll elaborate on in a few moments). But you know this summer my mind was completely scattered in trying to explain all of this to my tutors, and friends and fellow students alike. So, I invited a friend of mine to come and visit me while I was living in Cambridge to chat all of this through. This is Ewan. Ewan is one of my oldest and closest friends, we grew up here in Dubai together and we have shared a ton of adventurers of the years that I've known him.

We've filled buses full of bean bags, we've snuck into sand-dunes and played music, he's just finished his Engineering Masters in Oxford and is now working to help make buildings and cities more carbon neutral and energy efficient and ecologically sustainable. He's got his head really screwed on straight so we took to the Cambridge river Cam and the college gardens too and I attempted to ramble a few million words right though to him:

*"It seems that the steady state that the atmosphere adopts... ...but that really is just talking about the different way that meridian fluxes in energy in... ...all the colliding collisions of air in... ...Mars and to Titan and they found the same results... ...if you have laminar flow of fluid... ...in the form of like volcanoes that go off, of maybe then there's more extreme weather events... ...large-scale perturbation and then the way that your system responds you can see evidence for it in the carbon record because we see that there are these... ...I feel like I ramble too much... ..it just feels like irrelevant, random bits of information to you doesn't it?"*

*"I think you have to find and isolate the bits that you want and just write everything down in stages, of like this is that first thing I need to talk about, the second thing I need to talk about, and how those two link together and then how it goes into the actual message."*

[9] I'm referring to the vast majority of traditional work on thermodynamics here, but there is nevertheless plenty of research in areas outside of this traditional context, and many do explore the underlying fundamental statistics of living things. The work of Jeremy England [5] from MIT is a brilliant example, and I implore you explore his research if you can. Apologies for the simplifications, I'm trying to broadly frame the problem in order to get onto talking about the Maximum Entropy Principle.

I figured he was right, so, this video is about the inference engine side of things (just the basic building blocks), I have loads more things I wanna say about this but, I have taken his advice and I'm just taking it down and telling you about this really cool law in information theory that I learnt about, that is widely unapplied and widely unknown by the physics community.[10]

So, over the last year I have learnt about the work of a scientist called Edwin Jaynes, who, in my opinion, is probably the most important scientist that you've never heard of.[11] Now, this guy worked with probabilities, he was an information scientist so he was concerned with constructing probability distributions based on incomplete information: he was a Bayesian scientist, which is all about having a limited amount of information on your environment, on your system in question, and having to just do with what you can to try and articulate and model the system in question. And, he came up with this principle and it's called the Maximum Entropy Principle. [6]

So this Maximum Entropy Principle is really just an algorithm, currently buried deep down in the obscure and abstract world of data analysis and Bayesian radio astronomy.[12] [8] Now, what this algorithm does is fairly

[10]It might be a little bit cheeky to say this, and I suspect the word I was really looking for is "under-appreciated". I'm sure that there are many physicists that know of Jaynes' work, and of course the general principle of systems tending towards states of maximum entropy is generally widely known. It is his formulation of statistics from pure probability theory via MaxEnt that is not something that many are made aware of, and the potential to apply the principle to a wide variety of problems is something that I believe is being currently overlooked by many in the scientific community.

[11]Jaynes was an absolutely incredible mind, with an extremely diverse range of expertise. His complete reformulation of statistical mechanics in terms of probability distributions derived by entropy maximisation [6] is probably the most beautiful piece of published work I have ever encountered, but he was also an expert on applied classical electrodynamics and a quantum electrodynamics alternative called neoclassical radiation theory. He challenged contemporary thought and questioned established dogma in a variety of refreshing ways, and if you would like to learn more about his life and work visit `https://bayes.wustl.edu/etj/etj.html`

[12]As far as I'm aware this principle *is* infrequently applied, and seems to be limited to heavy usage only in image reconstruction from noisy data, but there may be areas that I'm missing. Jaynes' work is not widely celebrated, although within certain circles it is highly revered. During my year at Cambridge there was only a single professor at the University that worked on it, and even then it was more of a side-interest than a full-on research subject. It should be noted that Jaynes himself saw massive potential to his theory, and even briefly discussed its potential application to biology [7]. However its application as a

straightforward: it generates the probability of something taking place, and it ensures that that probability is, well let's say, optimum. It basically makes probability distributions.[13]

Now you might ask: what is a probably distribution? Ok let me explain. If I've got a die, a cube with six sides and upon each side is a single number then the likelihood of any particular number coming up (let's say) is one-sixth and so my probability distribution looks fairly straightforward, it looks like this. If, though, my die is not fair, in fact it's weighted, it might look a little bit different: maybe there's more chance for a six than a three or a two than a one. Now, if the system that I'm looking at actually doesn't contain only six discrete options but perhaps it can occupy a whole continuous range of states, then the probability distribution is going to be a smooth lines across the whole range of possibilities, and if it's high up on the axis it's going to be a very likely outcome and if it's low down on the axis it's probably going to be rarely occurring.

Now, the thing about Bayesian Information Science and Probability Theory in general and in particular this Maximum Entropy Principle is that it is concerned with systems that don't just give you the probability distribution: they force you to figure it out based on measurements that you can make with instruments that are perhaps only sensitive to particular things (particular variables) and it allows you to construct probability distributions based on those measurements. [9]

I'm going to give you just one equation — just one equation right now — it's the equation for Informational Entropy.[14] If I unpack it for you perhaps you can see here that for every probability distribution there is a corresponding entropy and vise-versa.[15] Now the Maximum Entropy Principle states

fundamental property of living systems, I'm fairly certain, has not been considered before.

[13]It should be said that there has historically been some confusion over whether it can truly be said that the principle generates prior probability distributions: technically it does not, as the uniform distribution (which arises from the probabilistic normalization constraint) always lurks as a pseudo-'prior' in the background.

[14]This is a variation on Shannon's 'differential entropy': an extension for a continuous probability distribution case. I am using it here to aid an understanding of what the function does to the $p(x)$ in order to generate $S$, but it should be said that Jaynes actually showed [10] that this isn't technically the correct form of entropy for the continuous case. One should instead use the so-called 'limiting density of discrete points'.

[15]Apologies, this is not true. A probability distribution can in theory always be integrated over to obtain a unique entropy value, however an arbitrary entropy does not, of

that if I maximise my entropy (which is a mathematical operation),[16] I'm going to generate the probability, the corresponding probability distribution, that minimises my bias with respect the constrains that I maximise my entropy to [11], and the constraints that I maximise my entropy to are just the measurements, they're just the dies rolls that I see or the particular variables that my instruments are sensitive too.[17]

So if I've generated some probabilities this way, I then get a probability distribution and then I go out and measure some more information I can then go back and update my probability distribution, maximise my entropy with this extra constraint that I've just been measured and I'm gonna generate better odds,[18] and actually I can repeat that cycle again and again and again, until my probability distribution is as good as it can be and I can't measure anything better. in fact, even more interestingly, the probability distribution I construct can then inform me on where I go looking for the next bit of data that I use to maximise my entropy.[19]

And why I think this is so interesting is be because there's cycle here: this is an inference engine. The cycle is: I generate some odds; I use these odds to inform some bet that I make; I make that bet; I see how that bet plays out; and I use that result to then inform my next distribution of odds — I update my odds — and I continue this process in a cycle on and on and on and on and each time I do this my probability gets better and better and

course, have a corresponding probability distribution. What I'm meaning here is that any *maximised* entropy *does* correspond *uniquely* to a single probability distribution, because the process of maximising will specify an associated likelihood distribution.

[16]Typically performed using the Lagrange Multiplier method, and which also always generates a unique maximum. If you want to actually see this in action, visit the appendix where I have outlined what is probably the simplest example of the principle in action: the Boltzmann canonical ensemble case. By using it in conjunction with a very simple hypothesis space one can generate the correct associated probability distribution along with the corresponding thermodynamic entropy relationship to the partition function. The great advantage of Jaynes' method is that it allows you to extend the definition of entropy to systems that aren't even in equilibrium.

[17]Often in the form of observed expectation values (again, see the appendix for an example of this in action). In the formalism that I'm trying to outline, such expectations are ultimately to be extended to carefully defined event rates.

[18]Provided your experiment is set up appropriately at least.

[19]It can inform you in the sense that it allows you to test predictions and perform experiments, but obviously it can't tell you anything about the inherent bias in your choice of sampling.

better and better.[20]

Ok, let's take a second to summarise shall we? I hope I haven't lost you too much. Just try and let's go back and think about the main point here, which is that there are these two different types of engines: there's the inference engine, and the heat engine.

I've tried to explain about this inference engine: I've tried to explain about how you can think about this entity that is working in a cycle and optimising to a principle and that principle I've tried to explain is this Maximum Entropy Principle, a way of generating probability distributions, you know, it's widely unapplied at the moment but that's just a suggestion. And then you have this heat engine, I've haven't explained too much but I've given you the example of the Carnot engine earlier in this video, this engine that is the 'golden standard' for thermodynamic physicists. It's this engine that has an entropy that is not changing and is therefore reversible which means that as you go through the steps of the cycle you can actually go backwards either way: it doesn't matter, it's reversible, you can step forwards you can step backwards it doesn't make a difference.[21]

Now this inference engine that I've alluded to is actually very different to this Carnot engine when you think about it. Far from it having an entropy that is constant, this guy is maximising it's entropy, it's trying to increase this entropy as much as possible over each cycle. And so just as the Carnot engine has got a constant entropy and therefore is reversible, the inference engine here is maximising its entropy and as it's doing so its actually maximising its

[20]If you're skeptical of this explanation: you should be. The cycle I broadly lay out here is heavily simplified and in some cases actually unhelpful when trying to apply the theory. We'll get into a much clearer and specific cycle if I get to talk to you about the molecular biology. Statistically speaking, the general idea is that by using MaxEnt, you generate a new partition function and a corresponding new free-energy landscape. Thermal fluctuations may then be harnessed to adiabatically drive the state of the system into the one that minimises free-energy, and then by once again spontaneously exchanging energy with the environment your system is kicked away from equilibrium, ready to cycle all over again.

[21]In practice systems are kept reversible by placing them alongside a thermal reservoir, and moved infinitely slowly, so that the system is constantly in a series of equilibrium states and entropy is unchanging. If you're new to this stuff and would like a clearer idea of this Carnot Engine in action, there are plenty of sources online that dive into it. On YouTube, I could suggest CrashCourse's video "Why We Can't Invent a Perfect Engine" for a neat and very simple step-by-step animation. `https://www.youtube.com/watch?v=2B81W6nNds0`

irreversibility, so each step of the cycle is a one-directional thing: you can't go back once you've done that.[22]

Now I'm having a bit difficulty now in this part of the video because I'm worried that I'm losing focus... because I could talk lot about what the implications to irreversible systems are and why it might make sense to think about some systems as wanting to maximise their irreversibility, especially probabilistic systems that are trying to learn and not forget. And I'm also really aware that I could go for a long time about how this inference engine probably does have a heat engine counterpart that looks quite different to the Carnot engine but yet still extracts heat from from an environment, just like the Carnot engine, performs work on the energy extracts, just like the Carnot engine, and also deposits its heat back into the environment, which is a really crucial part of engines.[23]

I don't want lose focus though, so I think we're gonna have to leave those two things to a future video and I will try and explain those in way more detail. I think that I could go at great, great, great length about them. But I think what we're gonna do for the rest of this video is to try and talk about how this inference could apply to living things and why this principle, this Maximum Entropy Principle, is a good principle to base a living organism's optimisation processes from.

So let me talk to you about another conversation I had with a friend of mine last year, at the end of summer. I spoke to a friend of mine called Ryan.

This is Ryan, I apologise on behalf of Ryan for his lack of decency here, you see, Ryan grew up in the Highlands of Scotland and so actually whenever there's the sun in the sky he feels absolutely obliged to remove his shirt. I went to university with him up in Scotland as well: he used to turn up in

[22] This isn't a great explanation but it will make more sense when it comes to the molecular biology example that I'm so desperate to talk about.

[23] I wish I could go on and on about this because I think an explanation here would really clear up a lot of the reasonable confusion that may arise from watching this video. I've set up the problem as needing to unite the concept of a heat engine and an inference engine in the context of a living organism, but the video now takes a turn away from this framing, and doesn't have the opportunity to return to it. It is the action of an inference engine handling work and heat from a probabilistic perspective (of energy in degrees of freedom that are in and out of control respectively) that will clear up the comparison between these two different types of engines.

shorts and a t-shirt on the coldest days of the year.

But we love Ryan, he's definitely one of the cleverest people I've ever had the opportunity to meet and speak to. He's now doing his PhD in Phenomenology, which is the study of what the basic building blocks of the universe are, in physics as well, in high-energy physics.

I actually went to visit him up in Scotland once at his hometown. He helped build his house with his father, he grew up building and making and tinkering and coding and fixing tractors and, well breaking tractors actually when I went up we did break a tractor.

Anyway, we took to the country park that was close by, we took our bikes out, he fixed my bike for me. Thank you for fixing my bike, Ryan. And while we were there I sat down with him and I asked him to explain evolution to me, to try and give me some kind of physical principle for the process of evolution.

And this is what he had to say:

*"It's like you have a massive experiment that you just keep running over and over again and every iteration you have things that do better and things that do worse. The things that do worse die, and the things that do better survive, therefore propagating the traits that perform better."*

Fair enough. So how then does that individual interact with different individuals? How do they all depend on each-other or their overall environment? I guess I wanted to know whether the environment was the master of the organisms or if the organisms themselves were ever the masters of their own environment.

*"This reminds me of the classic biology example of the predator and prey cycle, where they start off the same, so prey gets eaten and the predators diminish in number because they have less food - which means they eat less prey, so the prey goes up... ...and they trail each-other... ...for example, there was an early time, when life was forming where..."*

so then we moved on to talking about how early life had some really big and significant impacts on the state of the planet.

*"But what I'm trying to get at is that you have this massive body of organisms who are changing their environment, and they're changing the environment for everybody."*

And then we remembered something called "The Great Oxygenation

Event", which was the first main big extinction event, which occurred when our atmosphere filled up with Oxygen because of single-celled organisms.

*"When bacteria first came up with photosynthesis, oxygen filled up the atmosphere and, like, basically everything died. Because there was so much oxygen and no-one had become accustomed to dealing with the oxygen.", "So that's what happened, it literally killed the stuff that was producing the oxygen. Okay", "Yeah, exactly. And this dynamic is so interesting because it's like, these guys, you think of oxygen as their waste product and there's no evolutionary requirement to even have to address the waste product that they're producing. And then it fills up all of the air in the room and they suffocate."*[24]

It seemed really interesting that by depositing something seemingly so insignificant; some by-product that's simply unaccounted for, the whole planet eventually shifted in response, and it made me wonder about the underlying principles in probability theory that could be involved in the process of natural selection.

*"Because I haven't really read anything about that concept of living or-*

[24]For the sake of clarity, I've cut the conversation with Ryan to the bare bones, and so it is extremely simplified. Here is the actual story: Life on Earth at this time was single-cellular and generally anaerobic (meaning that it metabolised completely without oxygen). The highly reactive nature of the element meant that oxygen was extremely toxic to organisms that hadn't evolved to process it. When cyanobacteria came along and started to produce oxygen as a bi-product of the useful work generated via photosynthesis, nothing much happened for a long time... for 400 million years the oxygen in the atmosphere didn't increase substantially. There are two main hypotheses as to why this might have been. The first [12] is that there were some sort of geological processes that acted as an oxygen sink (such as the oxygenation of volcanic gases or continental crust, or both) that absorbed the global increase in oxygen. Eventually those sinks filled up and the atmosphere started storing it instead. The other suggestion [13] is that there were areas where oxygen was extremely abundant but contained inside an outward anoxic environment, and that it took time for the oxygen to leak globally. Regardless, by the time oxygen flooded the atmosphere, cyanobacteria had had the opportunity to evolve to regulate their own oxygen waste and survived the event. Some had even evolved to harness the reactivity of oxygen to unlock more energy for metabolism. The rest of their anaerobic fellows however were either wiped out, or banished to the dark anoxic corners of the globe and deep oceans. It's worth adding here that the Great Oxygenation Event actually eventually triggered a global glaciation (an ice age). The idea that by dissipating heat into a highly feedback-driven non-equilibrium system (such as the climate) and ultimately driving it to 'burn-out', resulting in a big-freeze is something that I'm extremely interested in, and I hope we'll get the chance to return to it in the future.

*ganisms optimising according to some rule — firstly there's no real rule that things should be optimising to that I have ever read.*[25] *Like, what's the rule? Is the rule Maxmimum Entropy?"*

In the end we were left wondering whether an organism that actually optimised according to maximum entropy would have an advantage over ones that didn't.

So this concept of bias is really important. If you're an investor, and you're making your investments based on some odds, you might ask: who's generating those odds? Maybe let's say it's the bookie: the book-keeper.

Now the book-keeper's job is to obtain all the relevant information needed to consider one particular event taking place, and calculate the likelihood of that event taking place based on that information. But that's a pretty hard job, and no bookie can be perfect. But! In an ideal world, the perfect investor has the perfect bookie next to them,[26] and the perfect bookie is the one that minimises its own bias: it's not adding any extra, superfluous information to inform their distribution of outcomes.[27]

So, the best investments are corresponding to the minimum bias in the bookie.[28] And the minimum bias corresponds to the maximum entropy,

[25] I'm yearning here for some underlying optimisation rule for evolution. Natural selection is a key mechanism of evolution, but does not qualify as a fundamental evolutionary law or key probabilistic principle of the process.

[26] Crucially, in the case of living organisms, the investor and the bookie are the same agent.

[27] This is equivalent to maximising uncertainty: hedging bets as much as possible given the data you've got to hand. A nice way to graphically picture what MaxEnt does is to think about it as 'flattening' the probability distribution as much as possible. The flatter the likelihood distribution, the less any one outcome stands out as more probable. So, if you flatten the distribution as much as you can whilst still staying true to your observed expectation values, you're maximising your informational entropy with respect to those observations. In other words, it ensures that the probabilities you've constructed contain the information you've obtained in your observations *and nothing else.*

[28] **This is just conjecture**, and I suspect it actually only applies to a particular class of evolutionary scenarios. The supposition is that in the case of a living organism operating in a steady environment with bets working out dependably, the probabilistic description has tended to one of minimum bias (and thus maximum entropy) in order to be most sensitive to fluctuations about the norm. In a scenario where competing organisms are fighting to gain a foothold in a wildly fluctuating and novel environment, I am much more skeptical of this claim here that "the best investments correspond to minimum bias". I suspect that the tending towards a state of maximum entropy often ultimately arises through the regulation

informational entropy, subject to the information that book-keeper knows. So there's this connection between maximisation of entropy and best bets.

Well okay, how is this relevant then you might ask. Well, I don't know for sure, but the more I think about it, the more I realise that Nature's pretty good at making bets right? I mean the trees are counting on the soil beneath their feet, the dandelions, here here we have one — they're counting on the wind moving it around, putting it and finding it in some different part of the ground to plant soil, to distribute its seed.[29]

When you look at ecosystems, they seem to be these environments where all these different agents are betting on each other working to pass. They've all tuned to become accustomed to…

…their environment. They've all tuned to become accustomed to each-other. In fact it leaves one to wonder, I think, upon the consequence to perturbation, to some kind of adjustment.[30] Brought on by changing winds,

and long-term feedback imposed upon the system by the surrounding environment to the organism. I hope we will talk more about this when exploring the Shuram and Early-Triassic excursions from the past.

[29]To clarify (although it almost should go without saying), when I speak of an individual organism as 'betting', I'm not assigning it any degree of actual meta-cognitive awareness. When a tree flowers is it consciously aware of the bees that come to pollinate? Of course not, but it still produces sugar-rich nectar exclusively for those pollinators, and counts on them visiting in order to transfer pollen to other plants. When that same tree grows its fruits is it 'aware' of the animals that come to transport its seeds? Of course not. But the process of making fruit in the first place is something that I would comfortably classify as a bet: the investment of a valuable resource deposited into the market with the aim of ultimately receiving a return. These transactional processes of living organisms can be perfectly described in this language without any anthropomorphising at work. I hope that's clear enough.

[30]This concept of kicking a system out of steady-state and watching the way it responds is absolutely integral to the kinds of things I would like to explore with you in the future. Specifically in regard to the following question: in such a scenario, what kinds of processes have the most significance in the overall state-trajectory of the whole system? To tackle it we'll first explore a very intriguing atmospheric analogy, and then look to a few events in geological history. The news articles that I use here (referenced in the original video description) refer to a small selection of the currently observed consequences to our own historic atmospheric interjection of fossil fuels and our rapid industrial expansion upon its surface. My opinion is that current predictions (as dire as they are) of ecological consequences of human activity have been significantly under-estimated. My main focus will be on the analysis of long-past events in order to provide context to the situation today.

or warming tides, or sweltering summers, or springs more like late autumn winters. This highly regulated market of bet-makers betting interchangeably, how this economy adjusts when the bets suddenly stop paying out like they used to, in a way that they've never done before.[31] Makes one wonder doesn't it?

So, that's where we are right now. I think that my plan to make this clear and concise maybe hasn't quite manifested (a couple of the people that I've shown this to have said that the message isn't particularly clear!). But it's hard to see just how widely relevant and applicable this all is without specific examples.

Questions such as: What is a bet? How do you define an event? How exactly is a living organism maximising its entropy? What does that practically mean? This stuff is going to make a lot more sense when applied to molecular biology. There's a protein that I really wanna tell you about. It's really important to the way that a cell functions, and I think that there's a fairly straightforward way of modelling this protein using MaxEnt (Maximum Entropy = MaxEnt).[32]

And I think once you see how this biological engine is extracting work from a thermal environment using, let's say, calculated investment or bet-making, then I think it will become apparent just to the number of different scales and levels you can apply this theory.[33]

And then there's the stuff I was talking about with Ryan and Ewan. Stuff that's really more of a step-back and looks at the bigger picture. Concerns itself more with thinking about: if you've got a collection of living organisms that are all operating in a similar way, how does that have an influence on the overall environment? How does that have significance when you think about quantities like entropy production, and specifically I wanna talk to

[31] Something I'd like to make a little more explicit is that I believe that this probabilistic formulation gives you a foothold to use the language of 'markets' and 'debt', 'recession' and 'stock exchange' alongside that of 'heat', 'work', 'steady-state' and 'environmental reservoir'.

[32] The inside of a cell is a perfect example of a system where bets are working out repeatedly and dependably. It is these systems where I postulate that Maximum Entropy will best apply.

[33] Just as with the rest of non-equilibrium statistical mechanics, things start to make sense when you actually *calculate* something, and I hope we'll get the chance to do that for the example I have in mind.

you about 3 periods in our geological history that I think that you can use to shed light on some of this stuff.[34]

The first is during a very famous period of time when life exploded in terms of its diversity and volume on the planet during the Cambrian period. The second period of time that I wanna talk you about is just after a big extinction event in the early-Triassic, then I also want to talk to you about, (in the relatively recent mid-Cretaceous) there were these big long periods of time where the deep ocean was completely empty of oxygen for some reason.

And then there's the computation, the statistics, the neuroscience. I think that InshāAllāh this will come in time, but specifically I think I can tell you about this really cool theory by a man called Karl Friston, who is a very well respected neuroscientist. He's come up with this theory [14] which says that the brain falls into states of minimum free-energy and I think that in certain cases there's a very intimate relationship between minimum free-energy and maximum entropy.[35] But this will come in time, I think: I think if you're interested in this straight away, Karl Friston did make a video about his theory [36]. And it's a good video and you can understand it (if you watch it a few times and pause after everything that he says to think about what he means). I'll give a link to that in the description.

Actually, speaking of the description, there's going to be a lot in there, lots of further reading, references, that kind of thing. I'm going to write a transcript of everything that I've said in this video and I'm probably going to write a bit more rigorous references to some of the things that I've said. And also I'm going to add some footnotes, because I've made some pretty extreme simplifications in some cases and I want the opportunity to justify those a little bit.

---

[34]Unlike the molecular biology, these examples won't involve any calculations. They consist of piecing together evidence into a story that provides perspective on the importance of living feedback processes on the planet *in very specific instances*. The short of it is that I think that geologists traditionally underestimate the significance of relatively small, yet highly non-linear dissipative process within the carbon cycle, particularly in response to global state-stressors. The focus will be on mixing in the oceans and authigenic carbonate formation.

[35]The intimate relationship between maximum entropy and minimum free-energy is widely known. Differences are concerned with which variables you hold fixed and which constraints you apply. Hopefully it is something we will get into. My intention is also to to sketch out an interesting way of considering the design and action of a quantum mechanical inference engine.

[36]by Serious Science: `https://www.youtube.com/watch?v=NIu_dJGyIQI`

# References


[1] W.H. Cropper. Rudolph clausius and the road to entropy. *Am J. Phys*, 54:1068, 1986.

[2] P.G. Wright. Entropy and disorder. *Contemp. Phys.*, 11:581, 1970.

[3] F.L. Lambert. Disorder - a cracked crutch for supporting entropy discussions. *J. Chem. Ed.*, 79:187, 2002.

[4] S. Carnot. *Reflections on the Motive Power of Fire and on Machines Fitted to Develop that Power*. Bachelier, Paris, 1824.

[5] J. L. England. Statistical physics of self-replication. *The Journal of Chemical Physics*, 139, 2013.

[6] E.T. Jaynes. Information theory and statistical mechanics. *Physical review*, 106(4):620, 1957.

[7] E.T. Jaynes. Where do we stand on maximum entropy? *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus (editors), M. I. T. Press, Cambridge, MA, 1979.

[8] G. J. Gull, S. F.; Daniell. Image reconstruction from incomplete and noisy data. *Nature*, 272:686–690, 1978.

[9] E.T. Jaynes. *Probability theory: the logic of science*. Cambridge university press, 2003.

[10] E.T. Jaynes. *Information Theory and Statistical Mechanics*. Benjamin, New York, 1963.

[11] W. T. Grandy. *Entropy and the Time Evolution of Macroscopic Systems*. Oxford university press, 2008.

[12] J. Kasting. The rise of atmospheric oxygen. *Science*, 293:819–20, 2001.

[13] D. Y.; et al Sumner. Antarctic microbial mats: A modern analog for archean lacustrine oxygen oases. *Geology*, 43:887–890, 2015.

[14] K. Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11:127–138, 2010.

**Appendix: The Boltzmann Canonical Ensemble** The Maximum Entropy principle arises from an analysis of the Shannon entropy for a set of variables $i$ and with probability distribution represented by $p_i$:

$$S_I(p) = -k \sum_i p_i \ln p_i \tag{1}$$

where $k$ is some positive constant, the value of which is completely determined by the choice of units via the selection of an appropriate logarithm base (we take it here as Boltzmann's constant $k_B$). One way to think about $S_I(p)$ is to consider it as the uncertainty represented in the probability distribution: a measure of the further information required in order to achieve full certainty.

In the canonical ensemble case[37] we analyse a system in contact with a thermal reservoir such that it may exchange energy (but no mass) with its surroundings.

So suppose we make an observation by performing some energy measurement and yield $E_T$. If we presume that this total energy is the average energy of each individual particle $\langle E_i \rangle$ summed over the total number of particles, we may say:

$$E_T = \sum_i \langle E_i \rangle = \sum_i p_i E_i \tag{2}$$

Additionally we must insist on the normalisation condition, namely that our probability distribution is an exhaustive set of mutually exclusive states:

$$1 = \sum_i p_i \tag{3}$$

These two suppositions now define what is called our **Hypothesis Space**. It is in this space that we place our past observations in order to use them a basis for our Bayesian analysis. (2) and (3) are used as constraints in the maximisation procedure. By combining with (1) and following the typical Lagrange multiplier method of maximisation, we subsequently obtain

$$p_i = \frac{e^{\frac{-\lambda E_i}{k_B}}}{Z} \quad ; \quad E_T = \frac{\sum_i E_i e^{\frac{-\lambda E_i}{k_B}}}{Z} \tag{4}$$

[37]This entropy maximisation derivation for the canonical ensemble has been performed many times over, often without mentioning Jayne's principle at all, despite correctly applying it here in this simple case.

with Lagrange multiplier $\lambda$ and the partition function $Z = \sum_i e^{\frac{-\lambda E_i}{k_B}}$.

If now we were to perform infinitesimal variations $dE_t$ and $dS$, and compare the result to the Standard 1st Law of Thermodynamics we would find

$$\lambda = \frac{1}{T} \tag{5}$$

And so, after defining $\beta = \frac{1}{k_B T}$, we can rewrite our probability distribution in the more familiar form:

$$p_i = \frac{e^{-\beta E_i}}{Z} \quad ; \quad Z = \sum_i e^{-\beta E_i} \tag{6}$$

and by inserting this back into (1) we obtain

$$S = \frac{E_T}{T} + k_B \ln Z \tag{7}$$

These are the standard results for the traditional Boltzmann canonical ensemble.[38]

[38]If you are unhappy with this derivation or would like to dive into this more, I've found a similar but more detailed derivation in the lecture notes of a Statistical Physics course from Edinburgh University: `https://www2.ph.ed.ac.uk/~mevans/sp/sp2.pdf`