On "Statistical Physics of Self-Replication"

Noah MacAulay

January 13, 2017

In this brief note I will present a simple counterexample to the main inequality found in "Statistical Physics of Self-Replication"[2](henceforth referred to as "the paper"). I will then present a slightly altered version of the inequality, discuss its significance for the main findings of the paper, and mention some related work.

1 The Original Inequality

First, some terminology. The paper concerns physical systems which can be found in some set of 'microstates'. A microstate could be thought of as a specification of the location of all particles in a system and their velocities, for instance. Interaction with an environment causes these microstates to transition from one to another in a stochastic fashion. We can think of the collection of microstates as a Markov chain – for every pair of microstates i, j, there is some probability $\pi(i \rightarrow j)$ of transitioning from i to j in a given interval of time(this probability can be zero) In addition, during each transition between microstates i, j the system releases an amount of heat ΔQ into the environment. This amount of heat can vary randomly, although we require the 'local reversibility' condition $\frac{\pi(j\rightarrow i)}{\pi(i\rightarrow j)} = \langle exp[-\beta\Delta Q]\rangle_{i\rightarrow j}$. β is the inverse temperature of the environment.

We can now define 'macrostates'. Macrostates are collections of microstates which are associated with some measurement outcome, e.g. 'there is one healthy cell in the environment', if our physical system consisted of an environment for growing cells. Then for each macrostate I, we can define a probability distribution over all microstates in I, $p(i|I)$, defined as the probability that the system is in microstate i after we observe I (after following some agreed-upon procedure to prepare the system) The entropy of a macrostate I is the Gibbs entropy of the probability distribution defined by that macrostate, $S_I = \sum_{i\in I} -p(i|I)ln[p(i|I)]$

Given two macrostates I and II, we can define some quantities related to the transition from I to II. $\pi(I \rightarrow II)$ is the probability of transitioning from the macrostate I to II over some fixed interval of time. The entropy change ΔS_{int} of a transition $I \rightarrow II$ is defined as $S_I I - S_I$. The average heat released is $\langle \Delta Q \rangle_{I \to II}$, where the average is done over all paths from I to II weighted by their probability and the probability of their starting states.

Now we can state the inequality: for arbitrary macrostates I and II, and a transition from I to II,

$$\beta \langle \Delta Q \rangle_{I \to II} + ln[\frac{\pi(II \to I)}{\pi(I \to II)}] + \Delta S_{int} \geq 0$$

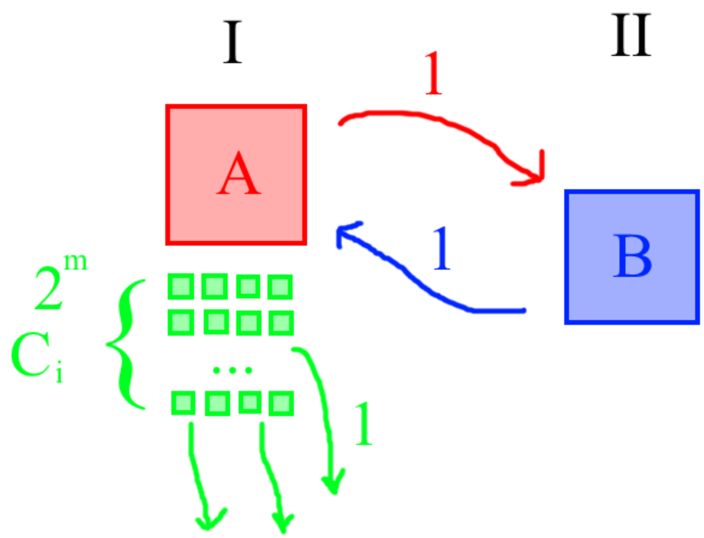

Figure 1: The counterexample.

2 Counterexample

To see the counterexample, let I consist of:

-a single microstate A with $p(A|I) = \frac{1}{2}$

-2^m microstates $C_1, ...C_{2^m}$ with $p(C_i|I) = 2^{-(m+1)}$.

And let II consist of a single microstate B with $p(B|II) = 1$.

Define the transition probabilities between microstates:

$$\pi(A \to B) = 1$$

$$\pi(C_i \to B) = 0$$

$$\pi(B \to A) = 1$$

From this, it is easy to calculate the quantities needed for the inequality. We have

$$\pi(I \to II) = \frac{1}{2}$$

$$\pi(II \to I) = 1$$

$$ln[\frac{\pi(II \to I)}{\pi(I \to II)}] = ln(2)$$

$$S_I = \frac{1}{2}ln(2) + \frac{1}{2}ln(2)(m+1)$$

$$S_{II} = 0$$

$$\Delta S_{int} = -\frac{1}{2}ln(2) - \frac{1}{2}ln(2)(m+1)$$

The only transition between I and II is $A \to B$, so
$\langle \Delta Q \rangle_{I \to II} = \langle \Delta Q \rangle_{A \to B}$
So we have

$$\beta \langle \Delta Q \rangle_{I \to II} + ln[\frac{\pi(II \to I)}{\pi(I \to II)}] + \Delta S_{int}$$

$$= \beta \langle \Delta Q \rangle_{A \to B} + ln(2) - \frac{1}{2}ln(2) - \frac{1}{2}ln(2)(m+1)$$

$$= \beta \langle \Delta Q \rangle_{A \to B} - \frac{1}{2}ln(2)m$$

Since $\beta \langle \Delta Q \rangle_{A \to B}$ is not dependent on m, we can always choose m large enough to make the above negative. So the inequality is violated.

3 The Altered Inequality

The problem in the derivation of the inequality came between equations (7) and (8) in the paper. Equation (7) states

$$\left\langle e^{-ln[\frac{\pi(II \to I)}{\pi(I \to II)}] + ln[\frac{p(j|II)}{p(i|I)}} \langle e^{-\beta \Delta Q_{i \to j}} \rangle_{i \to j} \right\rangle_{I \to II} = 1$$

By applying Jensen's inequality twice, we can obtain

$$\beta \langle \Delta Q \rangle_{I \to II} + ln[\frac{\pi(II \to I)}{\pi(I \to II)}] + \langle -ln(p(j|II)) \rangle_{I \to II} - \langle -ln(p(i|I)) \rangle_{I \to II} \geq 0$$

This is almost the desired inequality. The only problem is the last term, $\langle -ln(p(j|II)) \rangle_{I \to II} - \langle -ln(p(i|I)) \rangle_{I \to II}$. whereas, $\Delta S_{int} = \langle -ln(p(j|II)) \rangle_{II} - \langle -ln(p(i|I)) \rangle_{I}$.

From the definition of II and $I \to II$, we know that $\langle -ln(p(j|II)) \rangle_{I \to II} = \langle -ln(p(j|II)) \rangle_{II}$. However, in general, $\langle -ln(p(i|I)) \rangle_{I}$ will not be equal to $\langle -ln(p(i|I)) \rangle_{I \to II}$. The latter is an altered probability distribution, where states with a high probability of transitioning to II have a higher weight. Therefore, if there are some states in I which contribute a lot to the entropy but have a proportionately lower probability of transitioning to II, the original inequality can be violated, as in the example above. If we define the 'conditional entropy change'$\Delta S_{int|I \to II} = \langle -ln(p(j|II)) \rangle_{I \to II} - \langle -ln(p(i|I)) \rangle_{I \to II}$, we can write our new inequality in the form:

$$\beta \langle \Delta Q \rangle_{I \to II} + ln[\frac{\pi(II \to I)}{\pi(I \to II)}] + \Delta S_{int|I \to II} \geq 0$$

This is very close to the original inequality. The only difference is that the change in entropy is now a conditional change, where the distributions being summed over to calculate the entropies are conditioned on the fact that the transition took place.

4 Significance

Is this alteration significant for the main results of the paper, the application of the inequality to self-replicators?

It is not entirely clear. In the context of the paper, the state I is used to represent a single self-replicator, while II is a state containing two self-replicators. Then the inequality is used to bound the maximum possible ratio of duplication to decay events in terms of the heat released into the environment and the internal increase in entropy. It appears that the self-replicator could now 'cheat' this inequality, allowing it to have a smaller increase in internal entropy for a given amount of heat released plus ratio of growth to decay. But, the altered version of the inequality shows that it would have to do this in a very particular way – any gains over the original inequality would be bounded above by $S_I - S_{I|I \to II}$, which is essentially the knowledge gained by an observer about the initial state I, given that the transition to II took place. As in the example above, there would need to be a great number of states in I that contribute a lot to the entropy but are relatively unlikely to transition to II

What can we say about the structure of these states? Let's call the states that contribute a lot to the entropy but are unlikely to lead to a duplication 'impotent'. First, note that the degrees of freedom that lead to impotent states are unlikely to reside in the self-replicator itself, as the state II would presumably contain twice as many of these degrees of freedom, counteracting the apparent decrease in entropy production. But perhaps they could exist in the environment surrounding the replicators, if we are including that in the definition of the 'system' we are considering. For instance, perhaps there is some molecule in the environment that could be in a variety of configurations, and the self-replicator will only begin duplication if it encounters a particular configuration of the molecule. Then, the knowledge of an observer about the initial state would be increased, as the observer will now know that the molecule was more likely to be in some configurations than others.

I don't know if this sort of scenario is biologically realistic, or if it has any chance of applying to the systems considered in the original paper. But it seems at least conceivable that this alteration of the inequality could be exploited by self-replicators to exceed the bounds given in the paper.

5 Other Work

In their paper "The Bayesian Second Law of Thermodynamics"[1], Carroll et al. derive a useful general framework for thinking about problems of this sort. They derive the altered version of this inequality in section 7 of this paper.

References

[1] Anthony Bartolotta, Sean M Carroll, Stefan Leichenauer, and Jason Pollack. The bayesian second law of thermodynamics. *arXiv preprint arXiv:1508.02421*, 2015.

[2] Jeremy L England. Statistical physics of self-replication. *The Journal of chemical physics*, 139(12):121923, 2013.