Pre-read for Monday, April 15, Killing Laplace's Demon

Matt Salganik
Spring 2024

Note to readers:

- Given the purpose of this document will intentionally not have as much precision as a real piece of research, and I'll tend the caricature things to, I hope, make contrasts easier to see.
- Feel free to add comments, questions, and disagreements using the comments feature.
- After class, I plan to turn this pre-read into a series of blog posts.

In 2015, John Brockmann published a book with a provocative title: *This idea must die*. The book was a collection of essays by leading academics in a variety of fields responding to this prompt:

> **"Science advances by discovering new things and developing new ideas. Few truly new ideas are developed without abandoning old ones first. As theoretical physicist Max Planck (1858-1947) noted, "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it." In other words, science advances by a series of funerals. Why wait that long?**
>
> ## WHAT SCIENTIFIC IDEA IS READY FOR RETIREMENT?
>
> **Ideas change, and the times we live in change. Perhaps the biggest change today is the rate of change. What established scientific idea is ready to be moved aside so that science can advance?"**

I've got one idea that I think is ready for retirement after about 200 years of shaping people's thinking: Laplace's Demon. This thought experiment was proposed in 1814 by the great scientist Pierre-Simon LaPlace, and I think it shapes many people's thinking about predictability. But, it gives us exactly the wrong intuition about predictability, and it leads us down lots of pointless rabbit holes. So, in this set of readings, we are going to kill LaPlace's Demon. And, because I believe that to kill an idea you need to replace it with a better idea, I'll propose the best things I've seen so far: philosopher's Nancy Cartwright's idea of the "dappled world".

Be warned that the reading this week is pretty philosophical. My hope is that this will challenge and shape your worldview rather than give you some specific skill or piece of knowledge. Naturally, worldviews are hard to change so don't be surprised if this feels like a waste of time. But I would guess that for some people this reading could be transformation. This class is about taking intellectual risks that might have big payoffs. That's how I view this set of readings.

First, you will start by reading a Wikipedia entry describing Laplace's Demon. Then you will see that it quickly devolves into rabbit holes related to quantum mechanics and Cantor's diagonalization argument. In some sense, this seems like a pit of confusing ideas that people like to talk about late at night in their dorm room. For our purposes, we are not going to try to prove or disprove Laplace's Demon. Nor are we going to try to understand its relationship to free will or quantum mechanics. Instead, let's just suppose that Laplace's Demon is true. I'm still going to argue that it is irrelevant. And it is worse than irrelevant because it gives us the wrong mental model for how to think about predictability.

Laplace's Demon postulates an all knowing intelligence (today we might call it an artificial intelligence) that has perfect information about the state of matter, perfect information about the laws of the universe, and infinite compute. I want to argue that we are so far from these things that it is useless to consider what such an all-knowing intelligence might do. We've already talked a lot about not having all the data we need (think back to the Fragile Families Challenge and the "dark matter" interviews).

In the next reading by Frigg et al. you will see that having even slightly incorrect information about the laws of nature can cause big problems for predictability. In reality, we only have an approximation to the laws of nature so this might create a fundamental limit. For the purposes of this class, you should focus on their specific example; you don't need to understand how general that example is nor do you need to fully understand their proposed fix of "non-probabilistic odd". Naturally you are free to explore those ideas if they are interesting to you.

Next, you'll read something by Wolfram about computational irreducibility. In brief, he argues that for some systems the only way that we can see how they will evolve is just to let them evolve; there are no shortcuts. We have talked in class about strong views weakly held, please note that Wolfram makes some very strong claims with relatively little evidence. So you should not read these as facts, but rather as interesting conjectures. Also, upon further reflection I don't think this is related to Laplace's Demon as much as I would like, but it is certainly related to the limits of certain kinds of predictions. In class we will talk about rule 30.

The final reading will be the Introduction to Nancy Cartwight's book *The Dappled World: A Study of the Boundaries of Science*. To be honest, I'm not sure how to think about the argument (weak views, strongly held!), but this is the closest thing I've ever seen to a worldview that replaces Laplace's Demon with something coherent and more in line with empirical evidence about predictability.