

## Computational Complexity as Anthropic Principle

*"Give me a place to stand and with a lever I will move the whole world."*

Archimedes

*"I am a strange loop."*

Douglas Hofstadter

In the early 19th century the polymath and physicist Pierre Simon de Laplace imagined a demon. Laplace posited that a super-intellect in possession of all the data of nature, with the application of a single equation, could possess complete knowledge of not only the past, but the entirety of the future as well. What we have learned in the two centuries since he conceived of his demon is that its existence is almost certainly impossible. The laws of physics themselves prohibit the existence of any such super-intellect. And while that fact, in and of itself, may seem like a trivial observation there may be a way in which it is telling us something deep about the universe. Rather than merely serve as an ultimate limit to the human pursuit of knowledge we might use the impossibility of Laplacian demons as a tool to infer aspects of nature beyond what we currently know.

However, I will first need to explain how we learned Laplacian demons are impossible in the first place. Let me do so in the most engaging way I know possible- in the form of a fantastical story.

### Laplace builds a demon

He needed to know if the Terror would come for him. Already it had taken the lives of his friends Condorcet and Lavoisier, would he be next? In anticipation he had fled Paris along with his family but deemed it too dangerous to leave France itself. Being caught at the border would only confirm in the eyes of fanatics that he was a traitor. Only if he knew for certain that they would come for him here would he risk flight to safety in England or Prussia, but what was certainty?

It was during the height of his anxiety that Laplace stumbled upon the scroll. Into the hands of a scholar at the Académie des Sciences had come a curious document from the Ottoman land of Crete that had proven indecipherable to all who tried to crack its code. It appeared to be a sort of blueprint for a kind of mechanism, but the exact nature or purpose of the machine was utterly mysterious. It certainly was written using some form of sophisticated mathematics, though it utilized the geometric reasoning of the Greeks rather than the modern algebra, and for this reason in their desperation they had turned to Laplace.

The fear growing inside him was the force that let him see it. The device the scroll described was a sort of predictor, a means of calculating the future based upon a sort of

geometric calculus. And was this not a matter of great resonance? The ancient Greeks like the moderns had replaced magic with science but found the prospect of divining the future irresistible. Here was their attempt to transform the superstition of the Oracle at Delphi into something real and based on the true laws of nature. If such a device would work he might at last be free from the anxiety, not of fate, but of not knowing exactly what his fate entailed.

At once he set about building the instrument. He sent orders for a multitude of bizarre gears and levers to the finest machinists in Paris. It took him the better part of a year to complete it- punctuated by what proved only temporary frustrations and failures. When done the mechanism almost filled a large room, the greatest extent of which was taken up by a large sort of pendulum, in its center a thin needle, which gently glanced a circular layer of fine sand.

Questions to the oracle were posed by manipulating a series of knobs upon which were inscribed the script of the ancient Greek alphabet with letters used for numbers in the Milesian style. Pulling a lever returned the circle of sand to its prior, pristine state- smooth, symmetrical, featureless- and from the viewpoint of any particular grain of sand- perfectly random.

What perplexed and frustrated him was that the answers to his queries came back in the form of inscrutable scribbles drawn upon the sand by the pendulum. Laplace cursed himself for having done something wrong, but then he noticed the regularity of the drawings upon the sand.

He had an idea. He took a pair of dice and asked through the dials for a prediction of their outcome before he threw them. The instrument then drew lines in the style of tallying. Its predictions always perfectly matched the number of the dice that Laplace would throw. If it drew three lines and a space followed by two lines his roll would inevitably be a three and a two.

It was when Laplace asked for the square of the numbers he would roll that this tally system no longer worked- instead the device went back to carving nonsensical symbols. After doing this for a time he was reminded of an image he had seen in Diderot's *Encyclopédie*. It was an image describing the numerical system of the ancient Egyptians, and he realized to his delight that the answers to his questions were coming back in the form of hieroglyphs.

Harried by fear, Laplace called upon the assistance of the Egyptologist Jean-François Champollion, who fortune would have it, had fled to the exact same village as himself. Upon explaining his wondrous contraption, the two began the task of interrogating the

oracle (Jean-François called it The Demon) about the course of the future events of the world.

After a series of demonstrative experiments, he set about his task. Should they flee to England, Laplace asked by turning the knobs of Greek letters and waited as the device sketched out its answer in the sand as Jean-François looked on in astonishment.

“What is its answer?” Laplace barked, sometime after the drawing had finished with his assistant silent.

“I believe it says that it depends on the weather,” the Egyptologist responded.

“If I should fear storms in the channel, then, I’ll just ask it to predict the weather.”

Laplace replied, pulling the lever to clear the sand, and spelling out his question with the nobis. “And what weather does it predict?”

The device slowly drew out a series of shapes which Jean-François translated as “incalculable” followed by what he at first thought was an ankh symbol, but then realized it wasn’t a hieroglyphic at all, but a picture of a butterfly. “Unknowable because of butterflies? I do not know what it means.”

“The problem of three bodies,” Laplace exclaimed.

“What?”

“The position of three orbiting bodies cannot be calculated exactly.”

“There are limits to reason?” Jean-François wondered aloud.

“Might it at least calculate the best way for me to make my escape, given that the location of the Jacobins is unknown and always changing?” asked Laplace again pulling at the lever and nobis. Once again the oracle responded by drawing out symbols that meant unknowable.

“Is this unknowable at all times and places or only for you?” Laplace asked. To which the machine responded.

“For I.”

“Perhaps it’s a matter of size, if the machine were bigger...” Jean-François began at which Laplace immediately began to input the question.

“How much bigger would you need to be?” The machine’s answer to this Jean-François found to be nonsensical.

“It’s saying its circle would need to have a diameter of over three million river units.”

“What is that in kilometers?”

“Ten and one half.”

“Damn it, that’s over 30,000,000 kilometers!” Then Laplace remembered that he had recently stumbled across a similar number while researching his *Exposition du système du monde*. This figure was 250 times the diameter of the sun, a mass which he had surmised was so great that even light could not escape the pull of its gravity. There was almost panic in his voice, which prompted Jean-François to blurt out a long chewed over thought.

“I have often wondered how we might have avoided a revolution such as this. What conditions might have prevented the whole of society from crumbling like a castle made of sand. It was poverty that caused it. And what causes poverty but the want of surplus in society, a bare minimum left over from basic needs seized upon by those with power

so that they might live a true human life? Only a vast increase in the surplus would free us from this state, which might be fulfilled by machines with perpetual motion. Ask your demon whether such machines are possible.”

Laplace posed the question using the nobs. Slowly the oracle drew out its hieroglyph, which Jean-François translated as.

“Impossible.”

Laplace grimaced with pain and began. “Even should we never know the future course of events it is comfort enough to know at bottom that they have been determined in advance- that the world is structured according to the laws of some universal intellect far beyond human comprehension. Tell me oh ancient oracle- is the world at bottom determined or random?” And the machine answered:

“Random.”

He felt himself destroyed, as if the entire world he had constructed for himself had proven a lie and was disappearing through his fingers like a mirage. And then he remembered Epimenides.

“How can this cursed machine know that for all time what we ask of it is unknowable? It is true what they say- all Cretans are liars!”

## **The science that destroys demons**

In my imagined tale, Laplace’s attempt to build a super-intellect is undone by five of the deepest discoveries of science over the last two centuries: deterministic chaos, computational complexity, the Black hole information paradox, and the paradox of self-reference at the heart of both Gödel’s Incompleteness Theorem and Turing’s notion of uncomputability. I will discuss each in turn.

In the 20th century Edward Lorenz formalized the notion of deterministic chaos, which had been circulating at least since the genius Henri Poincaré had shown the three body problem could not be solved exactly. “Shadows” of possible trajectories would have to suffice. <sup>1</sup> The world might be deterministic, but we are unable to predict its exact course for any but the simplest systems. For dynamic and complex phenomenon prediction over long time scales requires infinite precision in terms of data. Seemingly inconsequential differences in measurement have a way of rippling through systems giving rise to radically different outcomes as Lorenz discovered in his investigations of the weather. <sup>2</sup>

Yet the limitations on a super-intellect imposed by the need to gather ever more precise data pale in comparison to the question of whether that data can even be processed in the first place. It’s here where we encounter the discoveries of computer science, especially since the 1970’s. <sup>3</sup> Computational complexity is formally a branch of mathematics, but is perhaps unique among the branches in that it takes the reality of the constraints of time and space seriously.

The problem of finding the best escape route that my imagined Laplace poses to the demon is a variant of what's called "The Canadian Traveler's Problem" (itself a harder version of the infamous "Traveling Salesman Problem.") <sup>4</sup> It is not that these and other examples of NP complete problems are unsolvable, it is that the time needed to solve them grows exponentially with the size of the problem itself.

Not even quantum computers are expected to be able to solve NP complete problems in polynomial time, let alone problems above that complexity class. The computer scientist Scott Aaronson has gone so far as to propose what he calls "The No Super-Search Conjecture", that states "There is no physical process to solve NP complete problems in polynomial time." <sup>5</sup>

The implications of computational complexity become extremely important for prediction (and therefore for physics) in the case of black holes and thermodynamics. As Aaronson points out, it's tempting to think we can solve NP complete problems in polynomial time by just throwing more resources at them. We can just use a bigger or faster computer. Yet both of these solutions merely swap one exponential for another. Exponential time for exponential energy. At some point the energy needs get so great that the concentrated mass of the computer exceeds the Schwarzschild Radius and collapses to form a black hole. <sup>6</sup>

If the physics of black holes is a unique environment where some combined version of quantum mechanics and general relativity becomes necessary, then for over a generation, physicists have been uncovering clues that such a unification might come in part from applying lessons learned in the field of computational complexity.

Since the work of Jacob Bekenstein we've known that black holes contain the maximum information density. The Bekenstein Bound means that no machine we build will ever have more than a finite amount of memory. <sup>7</sup> As much later pointed out by Patrick Hayden and John Preskill, limits on computation may be a way to solve the famous Black Hole Information Paradox. The information content of a black hole is equally inaccessible both inside and outside the horizon. The inside because of gravity and the outside because the information has become so scrambled <sup>8</sup> it would be impossible within the lifetime of the black hole to decode it. <sup>9</sup> Perhaps not truly random, but certainly pseudo-random. <sup>10</sup>

Still we don't need to go to the level of black holes to confront instances where uncomputability plays a deep role in physics. The other famous demon in physics was that of James Clerk Maxwell. Were a Maxwellian demon that could effortlessly separate hot and cold particles in a box possible, we would have created a version of the perpetual motion machine dreamed of since antiquity.

The reason why such a machine was impossible wasn't understood until Ralph Landauer re-conceptualized the paradox in terms of computation. Any Maxwellian demon would need to erase its memory, which requires energy and in the process would cancel out any energy gained from effortlessly separating its particles.<sup>11</sup> Ideas such as these borrowed from computational complexity may ultimately help physics to resolve the asymmetry of time in the classical and quantum worlds.<sup>12</sup>

We haven't yet even really touched upon quantum mechanics where not only is computational complexity informing physics, but physics is revolutionizing computer science itself. In just the last decade quantum computers have gone from science fiction to working models. More sophisticated versions of these machines, or even the discovery that we are incapable of building them, may finally allow us to resolve the dispute between the different versions of quantum mechanics- Schrodinger's, Heisenberg's or Feynman's- all of which make identical predictions but are built on radically different ontologies.<sup>13</sup> Quantum computers may also give us some indication whether the Many Worlds Interpretation is real or a mirage.<sup>14</sup> Computational complexity has already indicated that Bohmian Mechanics is intractable, and thus might give us a clue that hidden variable theories must be wrong.<sup>15</sup>

More astounding is the impact the study of quantum mechanics has had on computational complexity itself. The realization that quantum entanglement can be used as a resource for computation in the form of Multi-Prover Interactive Proofs (MIP\*) has resulted in an exciting marriage of the two fields of study the most profound of which seems to show that many forms of uncomputability might be overcome with the resource of infinite entanglement.<sup>16</sup> Unfortunately, this is not a resource found in the universe in which we live.

There have by now been so many instances of where the impossibility of Laplacian demons has allowed us to infer aspects of the universe that physicists might want to start thinking about using computational complexity as a way to infer constraints in the way they now use the anthropic principle. Any theory that would permit  $P = NP$  as in faster than light communication or time travel (closed time like curves), or which gave rise to memory or speed capacities beyond the Bekenstein Bound, could be bracketed as unlikely for that very reason.

Still none of this will answer what is perhaps the ultimate question when it comes to the role of computation in our world, the question that lies at the heart of our discovery of computation itself- is the world itself at bottom deterministic or random (in the sense of Wheeler's "law without law")?<sup>17</sup> Everything so far would appear to suggest that it is a finely tuned balance of both.<sup>18</sup> What makes the question perhaps impossible to answer is the dilemma of self-reference which is not a problem we can solve. We try to look at the universe from an Archimedean point outside of it and yet are inescapably trapped within. If fate exists then here is where it can be found, for we are forced to confront some version of Gödel's Incompleteness or Turing's Halting Problem in the radically

different worlds of black holes, quantum measurements, or even when we try to understand what it means to be a self. <sup>19</sup>

---

<sup>1</sup> R.C. Hilborn, *Chaos and Nonlinear Dynamics: an Introduction for Scientists and Engineers* (Oxford University Press, New York, 2006).

<sup>2</sup> Lorenz, Edward N. (1963). "Deterministic non-periodic flow". *Journal of the Atmospheric Sciences*. **20** (2): 130–141

<sup>3</sup> S.A. Cook, (1971). "The complexity of theorem-proving procedures" Proceedings of the Third Annual ACM Symposium on Theory of Computing - STOC 71

<sup>4</sup> C.H. Papadimitriou; M. Yannakakis (1989). "Shortest paths without a map". *Lecture Notes in Computer Science*. Proc. 16th ICALP. **372**. Springer-Verlag. pp. 610–620.

<sup>5</sup> S. Aaronson, (2014) Lens of Computation on the Sciences (2014). Institute for the Advancement of Science. <https://www.youtube.com/watch?v=hJibtaTmLtU>

<sup>6</sup> Ibid.

<sup>7</sup> J.D. Bekenstein, (1973) "Black holes and entropy" Phys. Rev. D. 7, 2333

<sup>8</sup> L. Susskind (2016) "Computational Complexity and Black Hole Horizons" Fortschritte Der Physik **64**, 24

<sup>9</sup> P. Hayden and J. Preskill, (2007) "Black holes as mirrors" Journal of High Energy Physics **2007**, 120

<sup>10</sup> Isaac H. Kim, Eugene Tang, John Preskill (2020) "The ghost in the radiation: Robust encodings of the black hole interior" <https://arxiv.org/abs/2003.05451>

<sup>11</sup> Rolf Landauer (1961), "Irreversibility and heat generation in the computing process", *IBM Journal of Research and Development*, **5** (3): 183–191

<sup>12</sup> Jayne Thompson et al. (2018) "Causal Asymmetry in a Quantum World" Phys. Rev. X **8**, 031013 Published 18 July

<sup>13</sup> S. Aaronson (2014)

<sup>14</sup> G. Jaeger, *Entanglement, Information, and the Interpretation of Quantum Mechanics* (Springer, Berlin, 2009).

<sup>15</sup> S. Aaronson, (2005) "Quantum Computing and Hidden Variables" Physical Review A **71**

<sup>16</sup> Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, Henry Yuen (2020) "MIP = RE" <https://arxiv.org/abs/2001.04383>

<sup>17</sup> John Archibald Wheeler "On recognizing 'law without law,'" Oersted Medal Response at the joint APS AAPT Meeting, New York, 25 January 1983

<sup>18</sup> T. Austin, (2018) "Measure concentration and the weak Pinsker property" Publications Mathématiques De L'IHÉS **128**, 1

<sup>19</sup> D.R. Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid* (Basic Books, New York, 1979).