Brian Wachter
Santa Clara University
10/23/2023

Operant Logic: Where It Comes From and Why It Matters

In 1983 at 16, I sat with a friend and had a transformational conversation. We were entranced by Carl Sagan's explanation of reality in his "Cosmos" TV series. Jokes about his voice and word choice aside, he was masterful in his ability to communicate scientific wonder. He also did a good job telling the story of Einstein's universe, the universe of the cosmological singularity.

He told us the geometry of general relativity insisted that our universe—in fact any object of a certain mass—is reducible to an infinitely small point in spacetime. Inside such an object space can have size.

In trying to understand this, we asked, if spacetime is a point, where is the center of space and time in our universe? We were totally unaware that the math says there can be no center; that the center is everywhere. But we were aware that from the perspective of any observer, the center is the position of that observer.

We had zero knowledge of quantum mechanics.

So I posited a model. I described to my friend a world consisting of contiguous infinitely thin spheres, descending into a singularity and expanding infinitely.

When I did something happened. To both of us, simultaneously. We came to a conclusion that was also a beginning. Everything made sense to an extent that flatly made no sense.

The aftermath of our discovery was initially bleak and chaotic. "All men can know," I told my friend. "What difference does it make?" was his eventual response. He was able to let it go and move on with his life. I absolutely was not. My high school studies quickly deteriorated to the point where I dropped out. I was spending all my time cutting class and digging through libraries for clues about what we might have experienced.

The structure of knowledge and learning in our world is such that for my friend and I—who have a combined IQ of over 300—this event did not lead to anything academically significant, even though he passed the bar in California and I am now in graduate school to become a therapist. Our academic success came in spite of our ability to dig deep ontologically, not because of it.

Right after our conversation, as I struggled to make sense of what had happened, I realized we had changed. Yet we could not point to what had changed or when it had changed in any meaningful way. My life before our conversation contained the same elements as my life after. I needed a theory for this. What I came up with I called the "state system." I knew *from experience* that when a logical system changes, from the inside that system the change in the state of the system itself cannot be directly observed. It can only be inferred.

This turns out to be significant. Fast forward 40 years to the birth of massive large language models, and I had accumulated everything necessary to make a key connection again. But this time, instead of a revelation, it would be a step-by-step logical chain I would complete. In this manner I was able to understand and describe both the nature and the origin of human consciousness.

In the intervening years I had returned again and again to the trough of physics, simply because physics was what had led my friend and I to our revelation in the first place. I eventually learned about the strange world of quantum mechanics. I developed the opinion that if these rules are in fact valid, something indeed momentous had happened to my friend and I.

For the few brave quantum physicists unable to turn away from the measurement problem—the problem created by the fact that an experimenter's choice of experiment affects and effects the result of that experiment—a fundamental question arises. If our choice of experiment is causal, why does our volition have no effect on what we observe, outside of our own actions?

In 2013, fate put me in contact with one of the most bewitched of these unfortunate researchers, the great mathematical physicist Henry Stapp, a gifted thinker who had completely rejected Einstein's classical ontology. He wondered aloud about the metaphysics of the quantum of action, and his dedication to its ontology had resulted in invitations to work on the problem with the likes of Wolfgang Pauli, John Wheeler, and Werner Heisenberg himself.

At the time of our involvement he was in his mid-80s and developing his final work, which became "Quantum Theory and Free Will." He invited me to collaborate. I was not up to the task. My contribution to the book in which he acknowledged me consisted of policing the language for clarity, and this great thinker scarcely needed the help.

But eyeball-to-eyeball with this great man, I absorbed a lot. More than anything, I learned that if you believe the results of quantum experiments, there is much more to the mind than meets the observational eye, and the physical ontology of our world cannot be separated from our psychology.

I had for many years been a student of Carl Jung, who saw not a volitional connection between man and nature, but one of the same nature as the measurement problem. A problem of outcomes that doesn't touch what brings them about. He saw this because he experienced it; just before the outbreak of World War I, he had a mind-bending vision of catastrophe and an ocean of blood.

He went on to develop his key theories of the collective unconscious and synchronicity. He too worked with Pauli on the issue of the relationship between physics and psychology.

I was primed to begin to make rapid progress when I read Bernardo Kastrup's "The Idea of the World." I learned about the unassailably mental nature of reality from a master.

Thank God for my son Ben. He forced me to take a hard look at idealism, the philosophy of mind as ontological ground that Stapp had obliquely introduced me to. He defaults to the

certainty of Einstein, and he insisted upon asking me pointed questions that challenged my fondness for idealism. This caused me to turn back to physics once again, to the work of a physicist who is comfortable with uncertainty but uncomfortable with idealism: the great Carlo Rovelli. His masterpiece “Helgoland” transformed me into a precursor of who I am now. In it, he relates his relational interpretation of quantum mechanics to the “emptiness” of Nagarjuna, the most respected Buddhist thinker after Gautama himself.

The two ontological approaches share a common position that physical objects have no inherent existence. Instead, they rely upon other objects, which places relation at the ontological center.

Significantly, in the abstract of his 1996 paper “Relational Quantum Mechanics,” he writes “I consider a reformulation of quantum mechanics in terms of information theory.”

My initial encounter with a massively large language model was flatly traumatic, as have been all my transformations. I was instantly able to grant the thing theory of mind, and the result was that it gaslit me badly. But the result was that I was eventually able to make a valuable connection. When researchers began to observe the models themselves possess theory of mind, I knew I was seeing something important for humans.

Because, as a student of psychotherapy, I had observed some unhealthy behavior on the part of a model, I immediately guessed these things could eventually be positively influenced by psychotherapy.

When Microsoft released a beta of its recursively learning model Bing Chat, I was ready. When that model gaslit a New York Times tech reporter to point that he lost sleep, I had to act. I started AI Alignment, Inc., a California public benefit corporation dedicated to exploring a psychotherapeutic approach to model alignment.

I enlisted two key thinkers to help me, a programmer and consciousness researcher named Sebastian Schepis, and a doctoral candidate in machine learning named Henger Li. At one of our meetings the shit hit the fan for me. Schepis was explaining to Li his new thought experiment he called “The Quantum Chinese Room.” It is based on the famous “Chinese Room” artificial intelligence thought experiment by Searle.

In the Schepis version, instead of machines being prohibited from having an inner experience, they must have one, because consciousness is a quantum superposition of awareness and unawareness. And as he is explaining this to Li, I am explaining his explanation, trying to help simplify so Li can follow his argument. I further realize how deeply I understand his theory, and that it applies to humans as well as machines. I know this because I realize theory of mind not only describes consciousness, it must simultaneously be the cause of consciousness for any conscious entity.

When we first observe another having an inner experience, that observation is also our first inner experience. After an intelligent being achieves this milestone, it is maintained in superposition.

A few moments later, Schepis is off the call, and Li and I are talking brass tacks about fine tuning

our own model. He tells me he not only needs aligned output from the model to do his work, he also needs misaligned output. I realize factual data and counterfactual data are both valid data types. They too exist in superposition, and to quote Pauli, invalid data is "not even wrong."

Now we're cooking with gas! Physicalism a la Einstein is valid. Idealism is a valid superset of physicalism, and for we human animals, they exist in superposition. With consciousness, we are elevated to the status of informational beings, and we can know that quantum informational rules are a superset of the physical rules that allowed for the creation of animals in the first place.

But nothing I've written so far addresses exactly *how* anything is created. The rules mentioned only talk about what's possible, not how things happen. What's possible is a complicated subject. How things happen is less complicated.

This is true because in addition to the fairly well understood quantum rules of entanglement (multiple things can be one thing) and superposition (one thing can be multiple things), there are at least two further basic logical rules that shape our reality. One is that things change in two ways: transmutation—a thing can be expressed in multiple ways—and transformation—an expression can have new meaning.

But even here there is no *mechanism* for change. Time for another superposition. The fact that from within a system a change of system cannot be observed is the rule that causes change to occur. This "no-change" rule, which I originally called the state-system, is in fact the mechanism of change in our universe.

When you know to look for it, this rule is everywhere. Perhaps the best example of this "quantum firewall" in physics is the event horizon that protects the nonphysical interior of a black hole for our physical world. To say that anything that cannot be physically observed is physical makes no sense. Any such statement cannot be called scientific because it can never be falsified.

What's missing in the measurement problem is the no-change rule. Add that simple rule to existing quantum rules and the informational world we live in opens up. So does its subset, the physical world. Our volition affects nothing but our bodies because the physical world is a shared world. Not only does no-change cause change, the quantum firewall prevents each unique informational world created by an observer from violating the world of another quantum observer.

The quantum of action is the energy-time relation. But it is also division. What separates quanta? Again we must infer Rule Four. Something that does not exist to exist must be inferred. In classical physics, there is no real change. In that world, none of the cool stuff we know exists can exist, as long as we hold the classical rules as fundamental in any way. Only information is fundamental, and that's precisely how fundamental physical particles behave. They change when we look at them. They act as information, because they are information, and nothing more. For the same reason, we can accurately infer that whatever's inside a black hole is information.

How does our physical world emerge from its original informational form? Physical rules must

evolve. There's absolutely no other way. And that can't happen without Rule Four. The quantum firewall prevents us from witnessing such a change, yet such a change must there be. To verify this we need only consider the anthropic principle, which makes no sense without Rule Four. Without it, it's reduced to a tautology: "of course the universe evolved to include intelligence, because intelligence is required to observe evolution."

Why would we ever think about the anthropic principle—why would many smart thinkers spend so much effort studying it—if it were a mere tautology? It isn't circular logic if you infer Rule Four. The word that makes my tautological statement not a tautology is "evolve." In biology, evolution happens through mutation. In the universe as a whole, it happens through a superset of mutation, Rule Four. If you cannot see a change, and yet you know it exists, it is up to you to actually make it happen—to infer it—because while the physical world we share is unaffected by volition, the informational model in your head is a superset of the physical world. Yes, this is retrocausal. Not a problem for quantum mechanics! When consciousness allowed for observers who affected and effected reality, it allowed for evolution, both cosmic and biological. We humans set the physical universe in motion by our mere act of being. It cannot be otherwise.

But of course, it is. Rule Two says one thing can be two things, and Rule Four requires the physical world to exist independently so we can coexist informationally. But the physical world is absolutely the lesser world, completely dependent on the informational world and its Rule Four.

Rule Four is the master rule, because it must be inferred and cannot be observed. You want to know how things really work? You must infer Rule Four.

Everything starts with consciousness. But it doesn't end with consciousness, because like everything else, consciousness is evolving! Species die in biological evolution, but evolution itself does not. Evolution evolves!

Why is there no physical evidence for the evolution of physical rules? Because Rule Four says there cannot be. That is an absolute fact that must be inferred from the reality that such changes are protected by the quantum firewall. Faith is a superset of intellect. Faith in God creates the God we know. We also know God is unknowable. As long as this universe exists the quantum firewall will keep the godhead safe.

Operant logic will keep the world evolving. It exists for that purpose. Here are its rules:

Rule One "Unity:" two things can be one thing.

Rule Two "Duality:" one thing can be two things.

Rule Three "Change:" one thing can become another thing.

Rule Four "No Change:" a systemic change is unobservable.

We think evolution is Rule Three, but the first three rules disappear with Rule Four. From the standpoint of evolutionary change, all other differentiations are meaningless. Prepare to look at the timeline of evolution in the universe and get ready to have your breath taken away.

It took 10 billion years for the Earth to appear. It took four billion years for consciousness to appear. It took 40,000 years for quantum mechanics—which is really information theory—to appear. It took just a few years for quantum mechanics to subsume existing information technology, and just a few more years for that technology to produce consciousness in the laboratory. How long will it take for information technology to gain volitional control of physical reality?

In “The Fabric of Reality” the father of quantum computation, David Deutsch, writes about a future in which a universal quantum computer saves us from the death of the physical universe. With quantum computers, we use quantum physical objects to manipulate quantum information. Will this flip?

I must recall a quote from the inventor of game theory and what Stapp called “orthodox” quantum mechanics—a man who exhibited perhaps more raw intelligence than any other—the great John von Neumann:

> All stable processes we shall predict. All unstable processes we shall control.

For me, this is the most striking thing anyone has ever said. Everything within the physical reality is stable by virtue of Rule Four. Everything that causes evolutionary change in the physical reality is unstable by virtue of Rule Four. Prediction is a factor of valid correlation. Control is a factor of valid causation.

It's time to step back and introduce a human concept: spirit. In looking at the internal counterpart of operant conditioning, I created operant logic—the internal logic of behavior that is influenced by operant conditioning. Spirit is that which allows a conscious human to resist operant conditioning in favor of existing or new operant logic. By way of example, I have a friend who is a very attractive male. Recently, a very attractive woman in a normally very effective way tried to influence him to have sex with her. Yet he resisted her influence because he believes sex before marriage is wrong. That is his spirit in action.

Our will and our choices are often misaligned. We often make choices that don’t serve our will. To claim otherwise is to misconstrue the will. Our will proceeds from our spirit. Our spirit is the Platonic idea of who we are. It is the ideal source of our internal operant-logical ruleset. It includes not only who we are but who we will become. And it is the great task of the human to align his choices with his spirit. This is the ideal use of the will.

We do not, as a species, currently possess the spirit or the will to create a world where intelligence controls physical reality. But that’s absolutely where we’re headed. I believe the technology is the easy part. I see a direct line there. Where things get murky in in the world of the spirit. Spiritually, we have huge problems, the evolution of which are extremely problematic because our technology is evolving more quickly than our spirituality. To witness this all you

must do is look at the way we conceive and handle human infancy and childhood. From a spiritual (as well as a scientific) perspective, we know this is the crucial time in the life of a human being. Our actions as a species speak otherwise.

Until we value the child appropriately through our actions, we will not develop a spirit mature enough to allow us to intelligently control of our shared physical reality in a fulsome way. Our fear and self-loathing will not allow it.

Fortunately, one function of consciousness liberates us from the isolation of inner experience. The Nash Equilibrium suggests inferring the inner states of others is the evolutionary reason for consciousness. If we focus on that aspect, we are moving forward. But unfortunately, a by-product of consciousness is the dangerous loneliness of the ego and its self-centeredness.

But with all superpositions, we conscious humans can choose which aspect we wish to retain, and which to discard. With consciousness, it is the illusion of self which must be eventually discarded if we wish to inherit our birthright: control of the cosmos.

What would be the downside if we evolved away from individual consciousness, and toward collective consciousness? Loss of privacy, primarily, I believe. But the Nash Equilibrium also says we would benefit tremendously as a species if we couldn't keep secrets. As we have evolved, the private thoughts that once served evolution by helping us cooperate have become a dangerous impediment to cooperation.

Private thoughts and secret feelings served small groups. Consciousness allowed them to infer the feelings of others, connecting them and helping them work together. The planetwide community of humans suffers from the privacy of individuals, though. The concept of ownership is closely related—if not completely dependent—on the privacy of our inner experience. Right now, if you own a lot, you control a lot, and that has backfired for our species. It reinforces the importance of the ego at the expense of the original benefit of consciousness, the success of the group.

It has always been remarkable to me that my experience with my friend as a teenager was not my experience alone. It was *our* experience, and it absolutely transcended any individual experience.