

The Peer-to-Peer Hypothesis and a New Theory of Free Will

Marcus Arvan
University of Tampa

Nick Bostrom is well-known for arguing on purely probabilistic grounds that we are probably living in a simulation.¹ Somewhat similarly, David Chalmers has argued that we should consider the “simulation hypothesis” not as a skeptical hypothesis that threatens our having knowledge of the external world but as a metaphysical hypothesis regarding what our world is actually made of.² Finally, the simulation hypothesis is gaining some traction in physics.

In "A New Theory of Free Will"³ and "A Unified Explanation of Quantum Phenomena?"⁴, I argue that a new version of the simulation hypothesis--the ***Peer-to-Peer (P2P) Simulation Hypothesis***--is not only implied by several serious hypotheses in philosophy and physics, but promises to provide a unified explanation of a bunch of baffling physical and metaphysical features of our world.

I begin "A New Theory of Free Will" by arguing that we presently have some philosophical and scientific evidence in favor of each of the following hypotheses:

1. *Eternalism about physical objects and properties*: past, present, and future physical objects and properties exist “timelessly.”
2. *Mind-body dualism*: the mind is at least partly comprised by non-physical properties or substance(s).
3. *Subjectivity about the passage of time*: time’s *passage* occurs within consciousness alone.
4. *Only one timeline (ours) is actualized*, or consciously experienced by observers.
5. *The holographic principle*: the physical universe is simply a series of ordered two-dimensional information (i.e. 1's and 0's) “written” on the cosmological horizon.
6. *The multiverse hypothesis*: the observable physical universe is merely a small part of vast multiverse of alternative possibilities.

Although many of these hypotheses are extremely controversial, I believe that the model they jointly imply--the P2P hypothesis--promises to explain things about our reality that need explaining, and for which we currently have no good explanation. Allow me to explain.

In computer science, there are two types of online simulations: (A) dedicated server simulations, in which one computer on the network (the "server") serves as the definitive

¹ Bostrom (2003).

² Chalmers (2003).

³ Arvan (2013).

⁴ Arvan (2014).

representation of objects and properties in the simulation, and (B) peer-to-peer networked simulations, in which *no single computer* on the network serves as the definitive representation of objects and properties in the simulation, but in which, instead, "the" simulation is simply represented in parallel on the various interacting computers on the network.

Figure 1. A Dedicated Server Simulation

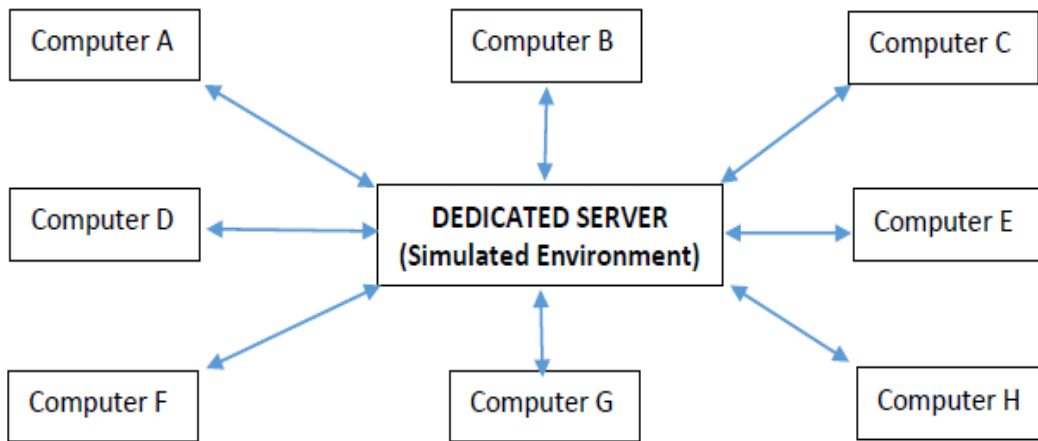
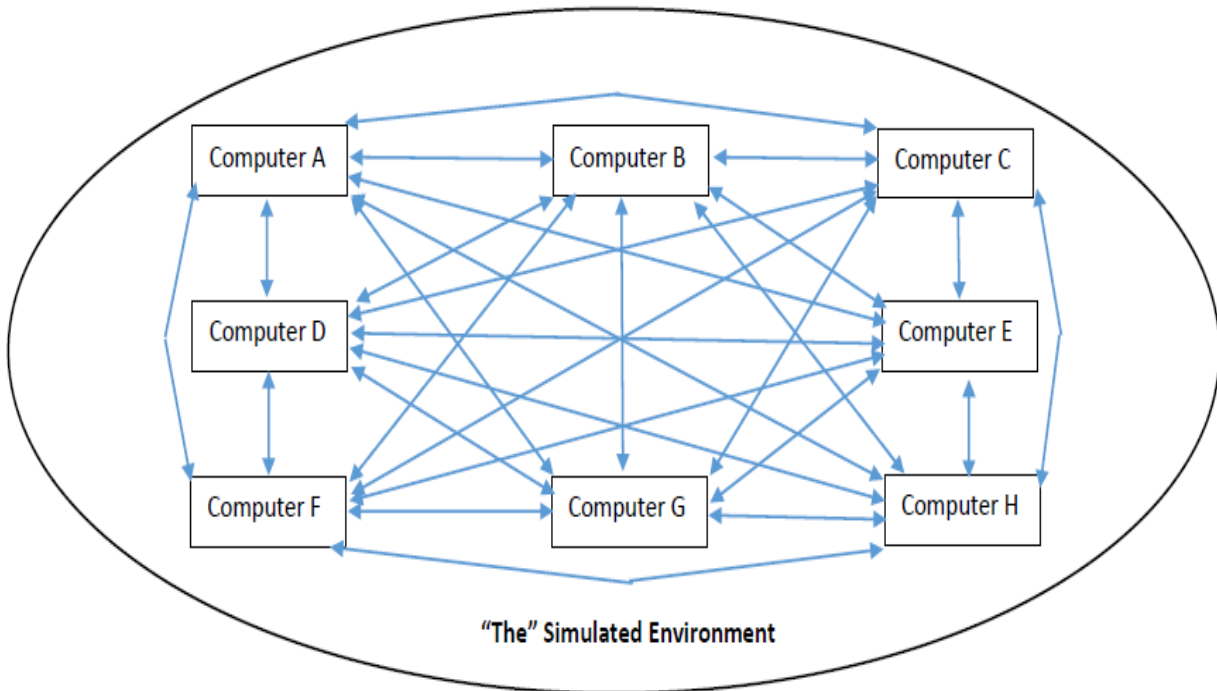


Figure 2. A Peer-to-Peer Network Simulation



Notice what a P2P simulation is. A P2P simulation *just is*:

- An array two-dimensional information (e.g. each computer's game program or DVD)

- Comprising a vast array of "possible pasts, presents, and futures" for the simulation,
- Being read in *real time*,
- By a multitude of *external measuring devices* (i.e. each computer on the network),
- All *interacting in parallel*, such that
- The joint measurements of all the computers on the network result in the appearance of single observed, intersubjective reality.

These six features of a P2P simulation are functionally identical to hypotheses (1)-(6). Thus, if hypotheses (1)-(6) are true, our reality is functionally identical to a peer-to-peer simulation. Let us now examine what this new metaphysical model might explain.

Explaining the Physically Unexplained?

Our world has a number of baffling physical features. They include:

1. Quantum superposition
2. Quantum indeterminacy
3. The quantum measurement problem
4. Wave-particle duality
5. Quantum wave-function "collapse", and
6. Quantum entanglement
7. The Planck length
8. The relativity of time to observers (no single, objective "universal clock").

It's worth noting, to begin, that physicists commonly recognize that we/they have "no idea" why our world has features (1)-(8) (see, for example, this video). Contemporary quantum theory--and extant interpretations of quantum mechanics--explain how quantum mechanics works (i.e. what follows from quantum-mechanical equations), but not why our world has these baffling features in the first place.

Interestingly, all eight physical features listed above emerge naturally from the structure of a peer-to-peer network simulation. Here's how:

- A peer-to-peer simulation just is a ***superposition*** of different parallel representations of the simulated environment on different computers on the network (viz. each computer has its own ever-so-slightly different representation of where things in the simulation are, such that the union of the different representations of "reality" is a giant superposition of alternate states),
- "The" location of any object or property in a P2P simulation is therefore also ***indeterminate***, given that each computer on the network has its *own* representation of where "the" object or property is, and there is no dedicated server on the network to represent where the object or property "really" is (any

object or property "really" is represented at many different positions on the network, thanks to slightly different representations on many computers all operating in parallel),

- Any measurement taken by any single measurement device a P2P network also thereby affects the network as a whole (since what one computer measures will affect what other computers on the network are likely to measure at any given instant), giving rise to a massive **measurement problem** (one can only measure an object is on the network by disturbing the entire network, thereby altering where other computers on the network will represent the particle as being),
- Because different machines on the network represent the same object in slightly different positions at any given instant (with some number n of machines representing a given object at position P , some other number n^* of machines representing a given object at position P^* , etc.) a dynamical description of where a given object/property probably is in the environment will have **features of a wave** (viz. an amplitude equivalent to the number of computers representing the object at a given instant, and wavelength equivalent to dynamical change of how many computers represent the object at a given point at the next instant).
- By a similar token, any particular measurement on any particular computer will result in the observation of the object as **located at a specific point** (thus instantiating a **wave-particle duality**), such that
- Any particular measurement on any particular computer will result in the appearance of a **"collapse" of wave-like dynamics** of the simulation into a single, determinate measurement (thus modeling wave-function collapse),
- It is also a natural result of a peer-to-peer network that single objects can "split in two", becoming **entangled** (in a peer-to-peer network multiple computers can, in a manner of speaking, get slightly out of phase, with one or more computers on the network coding for the particle passing through a boundary, while one or more other computers on the network coding for the particle to bounce backwards – in which case, if the coding is right, all of the computers on the network will treat the "two" resulting objects as simply later continuants of what was previously a single object).
- **All time measurements in a P2P simulation are relative to observers.** Each measurement device on a P2P simulation (i.e. game console) has its own internal clock, and there is no universal clock or standard of time that all machines share.
- Because the quantized data comprising the physical information of a P2P simulation will have to be separated/non-continuous much as there are "spaces" between pits of data on a CD/DVD/Blu-Ray disc (see image below), **there must be within any such simulation something akin to the Planck length**, an absolute minimum length below which measurements of space-time cannot be taken in principle (a feature of our world for which, at present, "there is no proven physical significance").

In short, the P2P Simulation Hypothesis promises to provide a unified explanation for many baffling physical features of our reality for which we currently have no such explanation.

Further, there are a couple of tantalizing new lines of evidence regarding the P2P Hypothesis.

First, 2013 National Medal of Science award-winning physicist James Gates has found the existence of *binary-error-correcting block codes embedded in the equations of string theory*: error-correcting codes just like those computer programmers actually use to prevent inconsistencies in online simulations.⁵

Second, the kind of quantum feedback mechanism I predicted on p. 41 of “A New Theory of Free Will”--a mechanism for primitive, libertarian free will to interfere with and alter the normal quantum wave-function (which was also predicted far earlier by the ORCH-OR theory of consciousness)--appears to have been observed (more on this below).⁶

Explaining the Philosophically Unexplained?

Our reality also has a number of philosophically baffling features. Among them are:

1. The mind-body problem
2. The problem of personal identity
3. The problem of time's passage
4. The problem of free will

The P2P model provides what I take to be a unified explanations of these problems as well, while also providing possible new resolutions to them:

1. Explaining the mind-body problem: It's a curious fact that it seems to many of us that no matter how complete a physical explanation might be, such an explanation could never possibly account for phenomenal consciousness (i.e. what red, green, blue, and yellow look like).

The P2P hypothesis promises, I believe, to explain this problem. Observers trapped in a P2P simulation would be convinced--just as many of us are--that there is something about their subjective point-of-view that cannot be captured in the physics of their world. *And they would be right.* The hardware upon which the simulation is running--the processing apparatus (viz. DVD laser apparatus/processor)--would comprise their subjective point-of-view, and be inaccessible to them within the simulation. More generally, the P2P model holds a reality like ours is comprised by two fundamentally different types of things:

⁵ Gates (2010).

⁶ Penrose & Hameroff (2013).

(A) "hardware" (i.e. consciousness/measurement apparatus), and (B) "software" (i.e. physical information) interacting.

2. Explaining the problem of personal identity: Many of us are tempted to say that we--our identity--consist in something over and above any biological or psychological facts about us. On the one hand, the "animalist" theory of personal identity doesn't seem quite right. We can seemingly imagine our consciousness (i.e. ourselves) jumping from one body to another (viz. Locke's prince and cobbler example or "Freaky Friday"). On the other hand, psychological continuity (i.e. "sameness of psychology") seems problematic because we can seemingly imagine perfect psychological duplicates of us who are not us (the "light" of our consciousness would remain with us).

The P2P Simulation Hypothesis promises to explain how this is the case. On the P2P Hypothesis, there *really is* something to us above and beyond the physical or psychological that accounts for our uneasiness: our identities are comprised by phenomena (i.e. hardware/consciousness) in a higher reference-frame inaccessible to us within the simulation except by direct acquaintance with our own first-personal existence.

3. Explaining problems with time: There are broadly two theories of time, the "A-theory" which says that time passes (viz. a "moving spotlight"), and the "B-theory" which says that time is nothing more than an ordered series of events (viz. time just is some events ordered before/after others). Both theories seem to face problems. A-theories seem hopelessly mysterious. B-theories seem to face problems making sense of change (i.e. if an ordered series of events is all that time is, how does time pass?).

The P2P Hypothesis provides a new answer: one that synthesizes both positions via a kind of mechanism/model that we already understand. When I go to play back a CD, the CD is a series of ordered information, and that information is experienced in real-time moving forward only insofar as a distinct observation-mechanism (the CD-player's processor) reads the information. This suggests that in order to make sense of time (i.e. it's being ordered and moving), we need a dualist theory--and the P2P Hypothesis gives us a concrete example of how such a dualist theory works.

4. Explaining (and solving?) the problem of free will: Einstein taught us that the way things appear from one reference-frame may appear the opposite from another reference-frame. Here's a simple example. If you were moving at a uniform speed within an enclosed elevator falling at an extremely fast velocity (say, 100,000 kilometers/hr, you would have no idea you are moving. You would think the elevator was still because the elevator is not accelerating relative to you. A person outside the elevator, however, would see you moving at an immense speed relative to them.

Now consider the problem of free will. From our perspective within our world, all of our actions appear to be determined by the laws of physics. This, obviously, is the issue that gives rise to classical problems of free will (viz. how can we be free if all of our actions are determined by physical law?).

P2P simulations and other online videogames, however, show how (A) the appearance of determinism or causal closure within a simulation can actually be an illusion of sorts generated by (B) causal interaction in a higher-reference frame not determined by any law of physics within the simulation. Allow me to explain.

Anyone who has played an online simulation knows that once one finishes playing a game, one can rewind the game back to the beginning, press the "play" button, and watch the game that just completed inexorably play out just as it did the first time. Accordingly, although the events that played out inside the simulation were the result of inputs from us from the outside, to any observer trapped within the simulation it would have to appear to them that everything in their world is determined by physical law. From their perspective, their laws of physics would appear to be "inexorable." They would think, for instance, that if their world were rewinded back to its beginning, it would have to deterministically evolve just as it did (with each of their actions being determined by its initial state and physical laws!).

Notice what's going on here. Although their world not deterministic vis-a-vis our reference-frame outside of the simulation (our inputs as game-players cause their game to play out the way it does), it *appears* deterministic (indeed, fatalistic!) vis-a-vis their reference frame within the game.

To make a long story short, the model shows how libertarian free will in a higher-reference frame (i.e. free will not determined by any physical law within a simulation) can generate the appearance of determinism in a lower reference-frame. Libertarian free will, in other words, is compatible with determinism--provided we distinguish between reference-frames.

Now, of course, there is one final rub. The simulations that we have created exist within our physical reality--a reality that appears itself to be determined according to physical law. As such, why not think that determinism is true vis-a-vis every reference frame?

This is where things get complicated, but in brief, the reasons I don't think this--the reasons I think we probably have true libertarian free will (the ability to self-cause according to pure thought in a manner determined by no physical or psychophysical law)--is that I think that (A) mind-body dualism is probably true, (B) there are no reasons to think that the actions of non-physical minds are determined by physical or psychophysical laws, and (C) there are reasons to think that non-physical minds are *causa sui*.

I realize that this will sound like preposterous "magic" to many readers, but as I will argue in future posts (and we can talk about this in the comments section here), I think that

physicalism (i.e. the traditional naturalistic worldview) invokes a lot of similar "magic" already in a manner that mostly escapes notice. Since I think there's a whole lot of "magic" in our world no matter what, and since I think the P2P Hypothesis and Libertarian Compatibilism promise to dramatically reduce how much of it there is (viz. quantum features), I think there are grounds for believing we may have libertarian free will.

In short, while Libertarian Compatibilism/the P2P Hypothesis may be crazy, quantum physics and relativity have *already* shown us that our world is crazy--and we might just need libertarian free will to explain some of that craziness. Maybe...it all depends on whether its predictions come out true.

Predictions: The Proof is in the Pudding

This brings us to a final issue, which is that Libertarian Compatibilism and the P2P Hypothesis make unique predictions--predictions that, if verified, would give us more reasons to believe the craziness.

Although I believe further investigation may reveal the P2P Simulation Hypothesis and Libertarian Compatibilism to make more predictions than these, at a minimum the theories make the following unique set of predictions:

- Our universe is a simulation--for which there might be some tell-tale signs.
- Our universe is a hologram--for which there might also be some tell-tale signs.
- We have libertarian free will--for which there might also be some tell-tale signs (i.e. subtle violations of the normal quantum wave-function within brains).

A final note on this last point. As I explain in "A New Theory of Free Will", if such violations of the normal quantum-wave function are observed in brains, those violations may well *appear* to have fully physical explanations within our reference-frame, since, no matter which physical "path" through the multiverse our consciousness choose, there will always be some physical explanation within that path to explain how the quantum violation arose (i.e. even if there are quantum violations, they may appear to us, in our reference frame, to preserve a kind of causal determinism).

References

Arvan, Marcus (2014). "A Unified Explanation of Quantum Phenomena? The Case for the Peer-to-Peer Simulation Hypothesis as an Interdisciplinary Research Program." *The Philosophical Forum*, 45(4): 433-446

----- (2013). "A New Theory of Free Will", *The Philosophical Forum*, 44(1): 1-49.

Beane, S. R., Davoudi, Z., & Savage, M. J. (2012). Constraints on the Universe as a Numerical Simulation. *arXiv preprint arXiv:1210.1847*.

Bostrom, N. (2003). Are we living in a computer simulation?. *The Philosophical Quarterly*, 53(211), 243-255.

Chalmers, D. (2003). The Matrix as metaphysics. *Science Fiction and Philosophy From Time Travel to Superintelligence*, 36.

Gates, James (2010). "Symbols of Power: Adinkras and the Nature of Reality", *Physics World*, reprinted online at <http://www.onbeing.org/program/uncovering-codes-reality/feature/symbols-power-adinkras-and-nature-reality/1460> (accessed on 14 May 2013).

Penrose, R., & Hameroff, S. (2011). Consciousness in the universe: Neuroscience, quantum space-time geometry and Orch OR theory. *Journal of Cosmology*, 14, 1-17.