compos mentis

On "Kant Meets Cyberpunk": An Analysis of Science Fiction and Transcendental Idealism

Vivian (Xiao Wen) Li
University of Toronto

ABSTRACT

In this paper, I argue that it is possible to conceive of non-material systems producing phenomenal spatial experiences, provided that we can factor in the notion of the transcendental ideal nature of time. I will first examine Eric Schwitzgebel's "Kant Meets Cyberpunk," outlining his argument that transcendental idealism may be true if an immaterial Cartesian soul underlying our current reality is creating virtual reality spaces that immerse us within the spatial environment around us. Schwitzgebel also argues that we are unable to access this immaterial soul because it exists on a base level of reality, while we live on a simulated reality plane. Moreover, I will examine David Chalmers and Nick Bostrom's notions of the matrix as well as the notion that we are currently living in a simulation. I will conclude with my argument for the transcendental ideal nature of time, as well as consider possible counterarguments and the implication if we accept my view.

ON SCHWITZGEBEL'S "KANT MEETS CYBERPUNK"

In order to explain the argument I will be proposing about the transcendental nature of time, it is important to obtain background from Eric Schwitzgebel's paper, "Kant Meets Cyberpunk." In the article, he argues that transcendental idealism, the proposition introduced by Kant that the world is made of phenomena (the appearances of objects) and noumena (things which are in themselves), could be accurate if conceived of along with ideas in Cyberpunk and science fiction. Moreover, he argues for conceiving of non-spatial systems organizing and bringing about spatial phenomenal experiences in the world. He claims that Kant's transcendental idealism consists of two theses, one, that spatial properties of objects depend on our minds, and two, that the fundamental nature of how the world works cannot be understood by us. The characteristic of spatiality, then, is something we bring to objects, and properties such as length, breadth, and depth are accessible to us because of the empirical evidence we obtain from our environments (Schwitzgebel 2019, 5). However, this is not to say that we completely create the illusion of spatiality in our minds, rather, Schwitzgebel is stating that there exists an underlying layer of reality which we interact with empirically and gives rise to the spatial properties we attribute to objects.

In both William Gibson's book, *Neuromancer,* and *The Matrix* (1999), a virtual reality (a computer generated space where a spatial environment can both be interacted with and is immersive, making the user feel like they are really in the scene) is created where the people submerged inside these realities can visually act and react to these spaces. In the former, hackers link up into Internet and navigate within the network of various computer programs, while in the latter, people are unaware that they are being held in warehouses and fed inputs from a computer device (Schwitzgebel 2019, 7). Hence, in *The Matrix,* people who have not awakened to the reality of the situation believe that they are drinking coffee or relaxing in Miami when their real bodies (at least in the warehouses) are motionless. Schwitzgebel also states that Cyberpunk is centered on the view that if we are submerged within our immersive spatial environments (something that makes us feel we are really within the scene), even if we can interact with empirical objects around us (objects we can touch/see/taste/move around), we cannot tell the difference between the current spatial environment and the fundamental environment which is controlling these phenomenal experiences. In other words,

it is entirely possible that we are living in a spatial manifold (spatial environment where each spatial part is related to another part) which is entirely based on another spatial manifold, the latter of which we have no control over. For instance, our brains may be strapped to machines in a warehouse, but our phenomenal experiences of the world are that of a dance party and drinking wine (with the empirical objects we interact with being the lights at the party and the wine glass, which exist within that specific spatial manifold). Schwitzgebel insists that it is conceivable we are currently in a world not unlike *The Matrix* and *Neuromancer* because everyday people can conceive of the aforementioned situation. I believe that he does not provide a good reason for why we might believe this is the case, a point I will bring up later in the paper.

Schwitzgebel, after pointing out the property of *Sims* (where an artificial entity is sustained by a computer and living in a shared virtual reality space with other artificial entities) in Nick Bostrom's paper, also argues that there is a difference between the base level of reality and the simulated level of reality. For instance, the base level could be that brains are stored in a warehouse and connected through wires, while the simulated level of reality could be the large and colorful world people live in. In order to create an argument for a non-materialist generator on the base level of reality, he moves to consider the Turing machine. If conceived of in a x-y plane, the Turing machine is an strip of tape that extends infinitely in either x or y directions (Schwitzgebel 2019, 13). It has a tip that can move in both x-y directions, which reads the alphanumeric characters on each square on the tape. According to if-then rules in its system, the machine can then erase the old characters on the tape and write new ones and move one square to its left or right, or stay in place. To complete his move of portraying a non-spatial system as an underlying spatial manifold operating at base level for humans, Schwitzgebel introduces a specific version of the Cartesian soul, which he names Angel. The soul can think, have conscious experiences, has causal abilities, and perfect memory, but does not have spatial positions nor traits such as length, breadth, or depth (Schwitzgebel 2019, 14). Schwitzgebel focuses on the notion that the Turing machine can be conceived of in a non-spatial manner, for instance, by replacing the symbols of the machine with musical pitches (low A or middle C) and the tape with integer numbers. Angel can then use his memory to move from one integer number to another, associating his current mental states (ranging from extreme sadness to happiness) with musical pitches, and rewriting the pitches

associated with the integer numbers in much the same way as a Turing machine would rewrite the alphanumeric character on the tape.

Schwitzgebel also considers whether the substrate on which the base level of reality is operating matters, for instance, whether or not transistors in parallel series could impact the speed a system processes things. However, he proposes a scenario with Kate and Peer, attempting to show that the idea of a non-spatial system creating spatial phenomenal experiences is entirely possible. Kate and Peter are two artificial intelligent beings who are enjoying the day, however, Schwitzgebel argues that a Turing machine could be responsible for their experiences. He also states that since it is possible that Kate and Peter could be run by Angel, thus, if Kate and Peter's existence is conceivable, then the notion that a non-material computer could be controlling their mental states and processes is also possible. If we accept his above points, then it is possible that beneath our experiences, there is a Cartesian soul which is immaterial and has instructions to execute and create a virtual reality space for us to live in, where it is impossible to access the soul because it is on a different level of reality than us. Also, Schwitzgebel states that we may experience spatiality only in terms of tracking our interactions and experiences with this immaterial soul (Schwitzgebel 2019, 20). Thus, if we can conceive of the aforementioned possibility, we can see how transcendental idealism may be true (Schwitzgebel 2019, 21).

ON NICK BOSTROM AND DAVID CHALMERS

Since Schwitzgebel employs a specific version of Kant's view in his paper, examining the historical Kant's notions of transcendental idealism may not help us much in analyzing the former's argument. Meanwhile, the notion of simulations and the existence of Sims is an important move towards a nonmaterial program generator at the base level of reality. Hence, I will now introduce and examine Nick Bostrom's article, "Are We Living in a Computer Simulation." In the paper, Bostrom argues that at least one of the following three statements is true: (1) the human race will extinguish before reaching the posthuman age (a posthuman is a being that exists beyond being a human, for instance, a cyborg with neurons artificially connected to processing units) age, (2) a posthuman is unlikely to run ancestor-simulations (which encompasses all mental belief-states of every human being that ever lived), or (3) we are living in a computer simulation at the moment

(Bostrom 2003, 1). I will focus on (3) as this is the section most pertinent to Schwitzgebel's argument.

In the case of (3), the complete mental history of humans will be captured on a computer and multiple ancestor-simulations would be executed without the posthuman needing to spend many resources (Bostrom 2003, 247-248). If (3) is true, and we are living in a computer simulation at the moment, Bostrom argues that there would be levels of reality, as the computer which is running the simulation may not have the same physical laws as the one we are viewing at the moment (Bostrom 2003, 253). The people living in simulated worlds could also become posthuman and create ancestor-simulations based on their reality. Since someone that is being simulated by a computer could also become the simulator of a new civilization, thus, if we (in our current reality on Earth) become capable of creating ancestor-simulations, we would arrive at the conclusion that we live in a simulated reality (rejecting (1) and (2)). Hence, it is likely that we are also living in simulator-realities created by simulated beings, while they are also living within the simulated environments of another group of posthumans.

While we could be ancestor-simulations living in another simulated world, it seems unlikely to me that there exists such a layering of realities. Bostrom seems to be assuming that every posthuman civilization will behave in the same manner, however, this may not be the case. For instance, civilization A could decide that they are interested in exploring ancestor-simulations, and create such a world. However, civilization B may decide, arbitrarily, that it is not necessary to explore these simulations because they wish to access alternate universes and do not want to use up even a fraction of their resources running simulations. Hence, I do not believe that even accepting (3) would lead to the conclusion that we are currently living in a simulation. Moreover, technology may be completely able to simulate belief-states of humanity, however, this might not lead to a completely individual experience of being in such simulations. In other words, while it is possible to simulate mental processes using expert nanotechnology and neuroscience, even this may not lead to complete individual thinking on the part of the ancestor-simulations. This is a point that Schwitzgebel does not address in his discussion on the incorporation of cyberpunk themes with Kant's notions of transcendental idealism.

Bostrom's simulations are also elaborated on by David Chalmers in "The Matrix as Metaphysics." Chalmers argues that the Matrix Hypothesis is not a skeptical

hypothesis (one where many of our beliefs would be incorrect if it was true, and is also impossible to strike out) (Chalmers 2005, 3). To give some background on the Matrix Hypothesis, it is the notion that we have always been a brain in a vat (or a brain being connected to a giant computer simulation of the world and receiving inputs and outputs from the simulation, thereby creating the illusion of beliefs and sensing) and continue to exist as such (Chalmers 2005, 2). Since we cannot be certain we are not within a matrix, many of our beliefs can also be questioned. We could, for instance, think that we are drinking coffee in Miami but this might not be the case, as we could be in a matrix and thus would not be in Miami (assuming the vat cells are not placed in Miami). In essence, we do not know for certain that we are not in a matrix, even though we may believe we are drinking coffee in Miami. If we are in a matrix, we are most likely not in Miami. But we do not know that we are not in a matrix, and thus we do not know if we are in Miami (Chalmers 2005, 2). This skepticism (that is, we believe certain things, but we are not certain of them) can then be applied to everything we believe about the world.

However, Chalmers rejects the notion that the Matrix Hypothesis is a skeptical hypothesis and argues that it is a metaphysical hypothesis, or one that is concerned with the philosophical elements of underlying reality (Chalmers 2005, 3). As such, the hypothesis that we are simply brains in a vat is a conjecture about the nature of our minds and of the fundamental nature of relaity. Furthermore, if the Matrix Hypothesis is accepted, it does not lead to the conclusion that we are not in Miami drinking coffee, since we are interacting with the bits that are representing Miami as well as the coffee and the coffee mug (Chalmers 2005, 4). Although not explicitly stated, Chalmers' points relate to Schwitzgebel's argument since the latter notes that there are distinct spatial manifolds, as well as different levels of realities. I am not particularly convinced by Chalmers' argument about interaction on different levels, as it does not explain which exact bits of reality and virtual reality correlate in the matrix. This is a problem since it seems that reality and virtual reality can be simply reducible to bits and would copy exactly into another, although the process of how this is done is unclear. For instance, the coffee mug I use in the external world would, in the matrix world, correlate to an aggregate of bits in a coffee mug formation. In response, Chalmers would argue that even in our current world, we are unsure of which quantum particles or waves exist in the coffee mug. The reasoning he would present, however, seems to point at a lack

of knowledge in our current reality rather than answering the question of how it is possible virtual reality and reality can easily correlate into each other.

REBUTTALS AGAINST SCHWITZGEBEL

Delving into the analysis of Schwitzgebel's points, he argues that it is conceivable for our phenomenal experiences to be completely controlled by a base level of reality constituting of machines and systematic mechanisms. In other words, he argues that it is possible we exist in spatial manifolds separate from our biological brains or bodies, and live in a shared virtual reality space. However, the only reason he gives for this point is that people are able to conceive of the possibility of this happening, as seen by the success of cyberpunk narratives (Schwitzgebel 2019, 10). This, to me, is not a good enough reason to believe in his argument, as he employs the use of "conceivable" as almost a means of establishing truth, or at least one possibility that potentially contains truth. However, many people can conceive of similar things without them being true at all (for instance, a unicorn). Hence, the reason he gives at this point in his argument appears to be an appeal to the masses fallacy.

Another issue with his argument is within the Kate and Peter scenario, which he uses as a stepping stone to argue for the conceivability of Angel. To reiterate, he stated that since it is possible that Kate and Peter could be run by Angel, if we can conceive of Kate and Peter existing, thus, it is possible that the system running them is non-material (Schwitzgebel 2019, 19). However, this may be a case of circular reasoning, as he first assumes it is possible that Angel runs the system without justifying Angel's existence. Then, he moves to state that Kate and Peter are conceivable, with the implicit assumption that their phenomenal experiences are being generated by a nonmaterial system. At the end of his argument, he states that since Kate and Peter are conceivable, it is possible their experiences are being generated by Angel. His argument moves full circle, as he is ultimately trying to prove his first assumption (that Angel runs the system).

Schwitzgebel also decides to assign the base level of reality as the one and only accurate level of reality. However, it appears to me that following his logic of a base level of reality generating phenomenal experiences for people in the same spatial manifold, there is also the implication another level could also be generating and controlling the base level on which our current reality operates.

As well, it is difficult to state the meaning of "levels," a problem that manifests more in a materialist setting than an immaterialist (like Angel, for instance). For Kant, it might be difficult to state the physical properties and interactions between phenomena and noumena in terms of atoms, quarks, and physical laws. Meanwhile, for Schwitzgebel, there may be a problem regarding how descriptive the base level of reality needs to be to produce a vibrant simulated reality, or even how the base level of one programming entity can generate so much variety and uniqueness, potentially creating multiverses, and can keep itself from overloading with information.

Although there are a few rebuttals that could undermine Schwitzgebel's argument, his claim is that transcendental idealism only *might* be true if incorporated with Cyberpunk themes. Hence, I would like to continue with this line of thought (thinking that transcendental idealism might be true) to see if it is possible to address my aforementioned rebuttals regarding the base level of reality and a vibrant simulated reality. I will explore the possibility of transcendental idealism with the notion that time could also be transcendental, and conceived of with immaterial programming systems existing on different levels of reality.

THE TRANSCENDENTAL IDEAL NATURE OF TIME

It appears that the transcendental ideal property of time may be necessary when approaching the view of an immaterial system creating phenomenal experiences for human beings on the simulated reality level. The reason for this is because following Schwitzgebel's reasoning and logic, there is no reason to stop at just the base level of reality being the fundamental system that functionally controls and operates the world. In fact, there is also a possibility that the spatial manifold the non-spatial system operates is based on another base level, which in turn is functioning to control this reality. (I am drawing inspiration from Christoper Nolan's "Inception" and the notion of a dream within a dream. Instead of a dream, I am proposing that there exists a system within a system, or a reality within another reality). This would solve the problem of how a single immaterial programming unit could create such variety and colour in multiple universes of simulated reality. The transcendental ideal nature of time (arguing that time is subjective and not a feature of things independent from us) would be necessary to start to understand the different functioning realities within each other, since

they cannot all be experiencing the passing of time at the same rate. For instance, since a system A would be the base level of reality for System B, and System B for C, with some time needing to pass before System B can be influenced by System A, it follows that time would not exist independently at all, but rather be subjected to the different Systems which are running the spatial manifold a subject X is living in.

I now consider a few counterarguments:

(1): One may argue that there must exist a basic operating system (such as in *The Matrix*), as my scenario may be too idealist. For instance, it would be difficult to define why something exists in System Z if we have to map back to System A (following Chalmers, the bits that constitute the coffee mug in System A may become completely warped by System Z). Furthermore, if there is an infinite series of worlds, there has to be an end (or a base reality) to it, or else the idea seems too bizarre.

(2): Another possible counterargument is the question of how these non-spatial systems could interact with our spatial minds, as it seems easier to explain this with *The Matrix*, given the spatiality of the machines. This is harder to understand in the context of Angel, even without considering Angel 2.0, 3.0, or any other iterations of Angel in my scenario (on Systems from A to Z).

I would respond with the following:

With regards to (1): Within a world as complex as *The Matrix*, it is equally as unlikely to map back specific parts of code within the system to the phenomenal experiences of the characters. Moreover, similar to what Schwitzgebel proposed, in that there is a base level of reality of the brain in the warehouse, where there are other empirical objects located within it, it is likely that the world we conceive of existing could also be conceived within another world. Then, the move that Schwitzgebel makes in his paper would be the same move I make, with the only difference being the added proposal that the base level of reality, including the non-spatial notion of it, could be operating within another simulation that both we and Angel (the Cartesian soul that seems to mimic the Turing machine) are unaware of. As for the bizarreness of the infinite series, I argue that this

scenario could be possible, not that it is entirely the case. It is, moreover, equally conceivable to me as Schwitzgebel's conception of the non-spatial systems creating spatial experiences for human beings on a day-to-day basis. We also live within dimensions that are more complex than just three-dimensional (according to Physicists and Mathematicians), dimensions that we cannot always view and visualize. However, just because we have limited mental capabilities, we cannot then determine that it is definitely not the case further dimensions do not exist outside of our comprehension. This is also the case with realities embedded within realities. By realities within realities, I mean that there are immaterial generators that create virtual reality spaces, which in turn create further generators, with each step requiring time to run differently in every system. Of course, if someone asks me to prove my point with scientific and empirical evidence, it would be difficult to do, as Science seems to view the world as reducible to formulas and fundamental waves or particles. However, I do believe that my view is possible within the frame of Schwitzgebel's argument, and is in fact a logical next step according to Schwitzgebel's points.

(2): Moreover, although we could be experiencing spatiality in our current reality, we may really be non-spatial entities living on a virtual reality space. If so, then there does not seem to be a problem with a non-spatial system interacting and simulating another non-spatial system, at least on a theoretical basis.

CONCLUSION AND IMPLICATIONS OF
ACCEPTING MY PROPOSED VIEW

One implication of accepting my aforementioned view is having non-spatial mechanics structure the phenomenal experience of empirical objects. For instance, it may be possible for machines to create, by themselves, non-spatial mechanics which would structure an entirely new being into existence (as seen by the entities chasing Neo in *The Matrix*). In that case however, we would most likely not realize this occurring, as it would happen on a level of reality that is simulated from our reality, and which (at least following the argument of Schwitzgebel) would rely on our reality also being a simulated one that is currently generated by an immaterial and non-spatial machine (perhaps someone named this machine Angel down the road).

Thus, while I agree with Schwitzgebel (that it is possible that a non-spatial system could create spatial phenomenal experiences), I also note that this could only be possible provided that time can also be seen as transcendentally ideal (not existing independently from us) and that there exists multiple base levels underlying each simulated level of reality (or for us, our current reality as we experience it).

REFERENCES

Bostrom, Nick. 2003. "Are We Living in a Computer Simulation?" *The Philosophical Quarterly* 53 (211): 243–255. JSTOR, www.jstor.org/stable/3542867.

Chalmers, David J. 2005. "The matrix as metaphysics." In *Philosophers Explore the Matrix*, edited by Christopher Grau, 132. New York: Oxford University Press.

Schwitzgebel, Eric. 2019. "Kant Meets Cyberpunk." *Disputatio* 11 (published online ahead of print) 11: 1–25.