Précis of *Reality+*

David J. Chalmers

Reality+ is intended as a work of technophilosophy. Inspired by Patricia Churchland's concept of neurophilosophy, technophilosophy is a two-way interaction between philosophy and technology: philosophy sheds light on technology, technology sheds light on philosophy.[1]

One area where technophilosophy is already very familiar is in the philosophy of artificial intelligence. We can ask philosophical questions to shed light on AI: e.g., are artificial minds genuine minds? We can also use AI to shed light on philosophical questions about human minds: what is the nature of the mind, and whicjh systems have minds?

Reality+ applies a technophilosophical treatment to issues about virtual reality (VR). It asks philosophical questions about virtual reality: e.g., are virtual worlds real worlds? It also uses VR to address many traditional questions about reality: e.g., what is the nature of reality, and what can we know about it?

Some terminology: a virtual world is an interactive and computer-generated environment. A virtual reality environment is an immersive, interactive, and computer-generated environment. A typical videogame on a desktop computer involves a virtual world but not virtual reality, because it is not immersive. A typical interactive digital environment experienced using a VR headset is virtual reality, because it is immersive.

The central thesis of Reality+ is virtual reality is genuine reality. This breaks down into three central theses: one metaphysical, one epistemological, and one value-theoretic.

1. Objects and events in virtual reality are real and not illusory.
2. We can't know we're not living in a virtual reality.
3. We can live a meaningful life in virtual reality.

A fourth thesis concerns the philosophy of mind: virtual (and augmented) minds are genuine minds. This roughly involves the thesis that AI systems can be conscious and that extending the mind with augmented reality technology (such as glasses that augment the physical world with digital information) is possible.

A general strand running throughout the book is the use of virtual worlds to respond to the problem of external-world skepticism (a line I first explored in my 2003 article "The Matrix as Metaphysics"). The second thesis above, that we can't know we're not living in VR, may seem to imply external-world skepticism, but the crucial first thesis, that objects in VR are real, blocks the implication. Even if we're in a virtual world like the Matrix, things around us are real, and many things are just as they seem.

Underlying this response to external-world skepticism is a sort of structuralism about the

[1] A version of this précis appears in Uriah Kriegel (ed.), *Oxford Studies in Philosophy of Mind* (2024), published by Oxford University Press.

physical world, akin to the structural realism that has become central in the philosophy of physics. The rough idea is that the truth of our ordinary external-world claims depends primarily on the causal structure of the external world, and that a simulation of a world has roughly the same causal structure as the world it is simulating: so if our external-world claims would be true in the unsimulated world, they will be true in the simulated world.

Chapters 1 and 2 of the book introduce the issues. Chapter 1 distinguishes questions about knowledge, about reality, and about value, both where virtual worlds are concerned and where ordinary reality is concerned, and introduces the three major theses above.

Chapter 2 elucidates the simulation hypothesis, which says that we are living in a lifelong virtual world. Simulations and virtual worlds are nearly coextensive in principle, but "virtual world" is typically used for the (often smaller-scale) worlds we create, while "simulation" is typically used for the (often larger-scale) worlds of the sort that we might perhaps already be in. I distinguishes various sorts of simulation, and make an initial case that we can't know we're not in a simulation.

Chapters 3-5 focus on epistemology. Chapter 3 presents Descartes' central skeptical arguments from the first Meditation through the lens of VR. Chapter 4 explores numerous standard responses to these arguments (from God, idealism, verificationism, dogmatism, simplicity, and more) and argues that none succeeds.

The central argument considered in these chapters is a standard Cartesian argument for external-world skepticism, placed in the key of simulation.

1. You can't know you're not in a simulation.

2. If you can't know you're not in a simulation, you can't know anything about the external world.

---

3. So: You can't know anything about the external world.

In this part of the book, I am especially concerned to defend premise 1, which is a version of the epistemological thesis laid out at the start. I eventually respond to the argument by using the metaphysical thesis at the start to reject premise 2.

Chapter 5 develops a version of the Bostrom-style simulation argument, and uses this argument to make a strong case that we can't know we're not living in a simulation. My preferred version of the simulation argument (refined over the course of the chapter) is as follows:

1. If there are no sim blockers, most humanlike beings are sims.

2. If most humanlike beings are sims, we are probably sims.

---

3. So: If there are no sim blockers, we are probably sims.

Here a "humanlike" being is roughly one with the same sorts of conscious experiences that we have. A "sim blocker" is roughly something that prevents the creation of many more humanlike

beings in simulated worlds than nonsimulated worlds. The conclusion is in effect disjunctive, with disjuncts including sim blockers such as Nonsims will die first, Nonsims will choose not to make sims (Bostrom's two disjuncts), Intelligent sims are impossible, Conscious sims are impossible, Simulators will avoid creating conscious sims, and Sims will require too much computer power. I argue that we can't know that any of these sim blockers obtain. I go on to argue that we should assign a non-negligible probability to the simulation hypothesis, and that consequently we can't know that we are not in a simulation.

Chapters 6-9 focus on metaphysics. Chapter 6 focuses on reality, introducing and motivating virtual realism (virtual worlds are real) and simulation realism (simulations are real), distinguishing different notions of reality along the way. Chapter 7 focuses on issues about god, arguing that a simulator may have some but not all of the properties of a traditional god, and considering the simulation argument. Chapter 8 focuses on information, and the it-from-bit hypothesis where the world is ultimately digital and everything is made from bits. This thesis includes both the pure it-from-bit hypothesis where bits are fundamental, and the it-from-bit-from-it hypothesis where the bits are grounded in something more fundamental.

Chapter 9 puts the pieces together to make an argument for simulation realism. The two key arguments are as follows:

1. If the simulation hypothesis is true, the it-from-bit creation hypothesis is true.

2. If the it-from-bit creation hypothesis is true, most of our ordinary beliefs are true.

---

3. So: If the simulation hypothesis is true, most of our ordinary beliefs are true.

Here, the it-from-bit creation hypothesis is the conjunction of the it-from-bit hypothesis with the hypothesis that our world (including the its and the bits) was created. One could deny premise 1 by holding that if we are in a simulation, there are bits but no "its": photons and other entities we take to exist do not really exist. The case for premise 1 in response goes via the following argument (appropriately generalized).

1. Photons are whatever play the photon role.

2. If we're in a simulation, digital entities play the photon role.

---

3. So: if we're in a simulations, photons are digital entities.

Chapters 10-13 switch the focus from the simulation hypothesis to real virtual reality technology using familiar headsets and the like. (I pass over these chapters relatively quickly, as none of the commentators focuses on these areas.) Chapter 10 argues for a sort of virtual realism in this domain, where virtual objects are real digital objects. Chapter 11 argues that percepion in VR is typically veridical and not illusory or hallucinatory. Chapter 12 takes up augmented reality technology, arguing for a form of realism here too, and for a limited sort of relativism. Chapter 13 considers epistemological issues about deepfakes and whether these might lead to a form of skepticism.

Chapters 14-16 focus on issues about the mind. Chapter 14 considers how issues about mind–body interaction look through the lens of VR. Chapter 15 addresses the question of whether AI systems can be conscious. Chapter 16 introduces the extended mind hypothesis and argues that augmented reality technology can extend the mind.

Chapters 17-19 address questions about value. Chapter 17 considers Nozick's experience machine and uses simulation realism (as well as considerations about autonomy) to argue that whether or not one can live a good life in the experience machine, one can live a good life in a virtual world. One central argument compares life in rich VR (a VR with roughly the complexity of ordinary reality) to terraform reality (life on a new terraformed planet), arguing:

1. Life in rich VR is roughly as valuable as life in a corresponding terraform reality.

2. Life in terraform reality is roughly as valuable as ordinary non-virtual life.

---

3. So: Life in rich VR is roughly as valuable as ordinary non-virtual life.

Chapter 18 addresses questions of moral status and argues that simulated beings can in principle have full moral status (that is, they can matter morally in roughly the way that human beings do). I argue that moral status requires consciousness: philosophical zombies lack moral status. I also argue that contrary to a common view, moral status does not require affective consciousness: philosophical Vulcans (with consciousness but no effect) still have full moral status. Chapter 19 addresses issues in political philosophy about how to build a virtual society.

Chapters 20-24 address some foundational issues underlying the case for simulation realism earlier in the book. Chapter 20 concerns the philosophy of language and the role of externalism, arguing that simulated worlds such as "Sim Earth" should be treated semantically in roughly the way that "Twin Earth" is treated by externalist arguments.

Chapter 21 focuses on the nature of computation, arguing that computation in physical systems (and therefore in simulations) requires a certain sort of causal and counterfactual structure. Chapter 22 focuses on structuralism and elaborates the structuralist arguments for simulation realism (and therefore against skepticism), as follows. Here Nonsim Universe is an unsimulated version of our universe and Sim Universe is a simulation of it.

1. Our physical theories are structural theories

2. If we're in Nonsim Universe, our physical theories are true.

3. Sim Universe has the same structure as NonSim Universe.

---

4. So: If we're in Sim Universe, our physical theories are true.

Chapter 23 concerns the relation between the scientific and manifest images and argues that we have been through a "fall from Eden" in transitioning from the former to the latter. We have moved from a primitivist conception where Color, Space, and so on are primitive Edenic qualities to a functionalist (or structuralist) conception where color and space are picked out by the

roles that they play. Eden can help to diagnose our residual skeptical intuitions. If we are in a simulation, our Edenic beliefs about the external world (say, that objects are laid out in a cerain way in Edenic space) are false; but those Edenic beliefs are already false in the post-Fall world of science. In both a simulation and in the post-Fall world of science, our non-Edenic beliefs about the external world may be largely true.

Chapter 24 considers skeptical hypotheses not yet addressed earlier in the book, from temporary and local simulations to Boltzmann brains, and tries to draw some limited anti-skeptical conclusions.