

Anna Longo*

How the True World Finally Became Virtual Reality

In the chapter “How the ‘true world’ finally became a fable” in *Twilight of the Idols*, Nietzsche enumerates the steps that led from the belief in the accessibility of the true world beyond the illusory world of appearances to the dismissal of this metaphysical myth. However, Nietzsche deems that the elimination of the “true world” brought about the obliteration of the apparent world as well: “The true world is gone: which world is left? The illusory one, perhaps? ... But no! we got rid of the illusory world along with the true one!”¹

In this paper, I will suggest that the philosophical hypothesis that we might live in a simulation² can be considered to be the last and most nihilistic episode in the series of narrations about the true and apparent worlds that Nietzsche sketched. As David J. Chalmers claims, “virtual reality is a sort of genuine reality, virtual objects are real objects, and what goes on in virtual reality is truly real.”³ Accordingly, the illusory world of appearances is considered to be equivalent to any computational simulation of it since the information that is relevant to achieve true knowledge in the physical world can be digitally reproduced. However, this means that the sensible world is completely lost since all the contradictory and chaotic qualities that characterize it as a becoming of appearances are dispensable. While the difference between the metaphysical

¹ Friedrich Nietzsche, *Twilight of the Idols*, in *The Antichrist, Ecce Homo, Twilight of the Idols, and Other Writings*, trans. J. Norman, Cambridge, Cambridge University Press, 2005, p. 171.

² The philosophical hypothesis that we might live in a simulation has been popularized by Nick Bostrom’s paper “Are you living in a computer simulation?”, *Philosophical Quarterly*, 53 (211/2003), pp. 243–255. Many philosophers, scientists, and public figures have been taking the hypothesis seriously, for example: John Archibald Wheeler, “Information, Physics, Quantum: The Search for Links”, in W. H. Zurek (ed.), *Complexity, Entropy, and the Physics of Information*, Boston, Addison-Wesley, 1990, pp. 3–28; Elon Musk (“Joe Rogan & Elon Musk – Are We in a Simulated Reality?”, <https://youtu.be/ocM69oCKArQ>, accessed 25 July 2021), Silas R. Beane, Zohreh Davoudi, and Martin J. Savage, “Constraints on the Universe as a Numerical Simulation”, *The European Physical Journal A*, 50 (148/2014), pp. 1–9.

³ David J. Chalmers, “The Virtual and the Real”, *Disputatio*, 9 (46/2017), p. 309.

* Collège international de philosophie, Paris, France

supersensible realm and the sensible world is overcome, the latter is substituted by a computational model that corresponds to true and objective knowledge of reality. Hence, we are nothing more than brains in a vat whose true knowledge is completely independent of the nature of the beings we interact with: they are essentially patterns of information that are truly known once they can be algorithmically reproduced.

It is important to point out that this conception of knowledge is the condition for present-day machine learning, an automated system for producing predictions based on the recognition of patterns of information. From this perspective, any existing entity can be reduced to the algorithmic procedure that is responsible for its behaviour, for the specific outputs that are calculated in response to specific inputs. This allows one to predict the results of interactions (events that are determined by the encounter of differently behaving entities) in such a way that decisions (to release information) can be made that are suitable for producing the desired effects. For example, water can be known as a pattern of behaviour that is correlated to temperature variations, such that changes can be modelled as depending on a simple rule that allows one to predict the series of phenomenal appearances that unfolds ordinately during the process of heating up or cooling down. Within this conception, intelligent agents operate according to an inductive learning method that allows them to enlarge their knowledge of reality and construct increasingly accurate abstract models. Such models are predictive hypotheses about the functions that describe the regular patterns that identify the unfolding of phenomena (how something changes in relation to the changes of some parameter) or the constant spatial relations that characterize specific objects despite their apparent variations (for example, if one cannot say that all ravens are black, one can recognize all the possible ravens thanks to structural invariants). Like artificial intelligence systems, human agents can improve their knowledge by exploring data concerning the behaviour of systems in order to make ever more accurate predictive hypotheses. The cognitive procedure that allows humans to build algorithmic models of phenomena (to identify systems with rules or functions that determine their structural relations as well as the way in which they affect or are affected during interactions) follows the rules of probabilistic statistical inference. A hypothesis about this complex procedure is used to program learning machines. Moreover, the knowledge (not yet achieved) of this learning rule, or what we can call the logic of induction, can be used to determine the information that the brain considers to be pertinent

in recognizing the pattern of information that corresponds to the model of a phenomenal object (a category or concept): such information can be digitally supplied to obtain the desired brain responses, i.e. the recognition of specific objects that call for specific actions that are meant to gain the desired effects. True knowledge is then the coherent set of predictive hypotheses that ought to be believed in order to successfully interact with objects that are identified with patterns of algorithmically produced information. Hence, true knowledge is an algorithmic rule that allows one to link the most efficient response to the incoming information under the hypothesis that the received data can be interpreted as a part of a regularly unfolding series (a specific object “causing” the brain stimulation): the recognition of a pattern justifies the expectation of future variations and such expectations justify the selection of specific actions. Since any object is but a pattern of information, it does not make a great difference if such information is physically communicated to the brain through the senses or if it is conveyed by electrodes stimulating neurons. In any case, the cognitive system will operate correctly and true inferences can be made in both physical and virtual reality, since both material and digital objects are mere patterns of information. What Nietzsche called the apparent world, or the reality of chaotic and contradictory sense impressions, is then reduced to an ordered set of computable series of information whose truth can easily be determined by checking if the rule is able to generate possible variations that have not yet been observed. Such a world, composed of nothing but algorithmic procedures or computational models (anything that is truly known can be correctly simulated by a computer), is believed to be the product of a correct procedure for the production of knowledge. Metaphysics is then overcome by proving that true knowledge of the empirical world can be obtained independently of knowing the reason why reality is such rather than otherwise. What makes a hypothesis believable, in fact, is that it has been produced according to the rational procedure that ought to be followed in order to produce satisfactory models of perceived phenomena and can be grasped by analysing the predictive theories that are collectively considered to be true. This analysis of the most trustworthy hypotheses is supposed to unveil the cognitive or inductive procedure for constructing efficient models; these hypotheses can be tested by simulating learning procedures in artificially intelligent machines. To put it in another way, metaphysics is overcome by showing that the truth about experienced reality can be stated independently of its nature (its supersensible origin) but under the condition that reasonable beliefs are held as regards the complex inductive rule (or logic) that allows rational

beings to construct models to efficiently cope with streams of data by selecting the information that is pertinent to classifying the different causes of brain stimulation and by responding appropriately.

Chalmers's thesis – according to which virtual reality is as true as the apparent world we used to call “reality” – is inscribed in the conception of knowledge that supports our contemporary automated system of knowledge production and that considers any existent being to be an algorithmically generable pattern of information. Moreover, as we will see in this paper, Chalmers's thesis about the equivalence between the physical and the digital world can be deemed to be the effect of the achievement of the positivist effort for the “elimination of metaphysics”⁴ and an extreme consequence of the effort to establish objective criteria for judging the truth value of propositions about the apparent world. I will argue that Nietzsche's prediction about the obliteration of the apparent world has actually been fulfilled by Chalmers, and I will show why his theory must be considered one of the many fables that humans have been producing in order to organize the world according to their own ends. A fable that is believed to express true knowledge about reality but which is just an effect of the imposition of a non-necessary norm for constructing worlds suitable for the satisfaction of human needs. A fable whose necessary condition is the acknowledgment of the human creative capacity for lying, which is in continuity with the real as true becoming.

The Fable of the True World

That the world we are immediately aware of is an illusion is an old philosophical assumption. Ever since Plato's allegory of the cave, different versions of the myth have been presented so that the history of Western philosophy can be seen as a series of variations on the theme of the relation between the true and apparent worlds. From Nietzsche's perspective, this is the story of the progressive

284

⁴ I refer here to the paper published by Rudolph Carnap in 1932, “Überwindung der Metaphysik durch Logische Analyse der Sprache”, *Erkenntnis*, 2 (1/1932), pp. 219–241, and translated into English by Arthur Pap as “The Elimination of Metaphysics through Logical Analysis of Language”, in A. J. Ayer (ed.), *Logical Positivism*, New York, Free Press, 1959, pp. 60–81. The reference to Carnap is motivated by the fact that Chalmers recognizes the influence of the *Aufbau* on his own research program (see David J. Chalmers, *Constructing the World*, Oxford, Oxford University Press, 2012.)

intensification of nihilism or the imposition of a system of valuation that favours the preservation and growth of the human species but which is blind to the real value of life as creative becoming.

In Plato's famous allegory, we are like prisoners in a cave of sensual perceptions where we take shadows for real things. The knowledge we have about the projected shapes is revealed to be false,⁵ not because the expectations based on the observed regularities are incorrect, but because such theories concern phenomena in the apparent world of the cave rather than the objects of the true world (ideas). It is important to point out that knowledge here does not merely consist in the capacity of addressing the physical objects that pass between the fire and the prisoners' backs and that produce the shadows on the wall (we could say the material but invisible causes of phenomena that our science aims to grasp). In Plato's myth, in fact, true knowledge can only be gained by looking into the sun, by turning away from any material concern and sensible attraction. The sun is a source of truth that is inaccessible to sense organs,⁶ a brightness that our eyes cannot bear: the objects of true knowledge are suprasensible and they can only be grasped in thought. The soul, as the non-material inner light, provides us with access to the true world and with it the possibility of contemplating the real objects of knowledge. The true world is attainable for Plato because, as Nietzsche says, "he lives in it, *he is it*."⁷ Such identification with the light of truth is possible because the philosopher has detached him- or herself from any sensible interference in order to participate exclusively in the true world.

Since Plato, philosophy has been characterized by a progressive distrust in the speculative power of reason (the faculty of ideas) and an increasing confidence in rationality, the faculty that aims to construct the appearances of the world

⁵ As Socrates suggests in this dialogue: "However, what if among the people in the previous dwelling place, the cave, certain honors and commendations were established for whomever most clearly catches sight of what passes by and also best remembers which of them normally is brought by first, which one later, and which ones at the same time? And what if there were honors for whoever could most easily foresee which one might come by next?" Plato, "The Allegory of the Cave", trans. T. Sheehan, <https://web.stanford.edu/class/ihum40/cave.pdf>, accessed 25 July 2021, book VII, 514a–517a.

⁶ "But I think that finally he would be in the condition to look at the sun itself, not just at its reflection whether in water or wherever else it might appear, but at the sun itself, as it is in and of itself and in the place proper to it and to contemplate of what sort it is." *Ibid.*

⁷ Friedrich Nietzsche, *Twilight of the Idols*, p. 171.

according to objective (universal) rules. While doubting the accessibility of the true world beyond the cave (the metaphysical reasons for the world to appear as it does), philosophers started to be interested in the mechanism that produces the observed regularities in the shadows. The goal was to turn the prison into a home that can be fashioned according to the satisfaction of human needs and desires (rather than according to spiritual aspirations). Men of knowledge decided to focus on the apparent world in order to discover the laws that organize the structure of its appearing and that can be exploited in order to increase the efficacy of actions. In the empiricist and positivist version of the story, the disposable world is the ideal or metaphysical one, while useful knowledge is based on experience and relates to the mathematizable constant relations among the observed facts. The platonic light of reason is now seen as the obscure normativity of unprovable principles (metaphysical dogmatism) and the real light is the rational capacity to produce hypotheses whose truth depends upon their practical efficacy. As David Hume suggested, most philosophical speculations should be forgotten since they are useless:

If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask: does it contain any abstract reasoning concerning quantity or number? No. Does it contain any experimental reasoning concerning matter of fact and existence? No. Commit it then to the flames: for it can contain nothing but sophistry and illusion.⁸

However, the abolition of the true world has not been achieved by the empiricists and the positivists, who were merely sceptical about the possibility of obtaining certainty and who thought that absolute truths were just inaccessible for us. The true world is a fable that can still be told even when the representation of the apparent world cannot be considered to be necessarily true: for sceptics like Hume, any theory is, to a certain extent, a fable, even though fables can be evaluated with respect to pragmatic criteria. The abolition of the “true world”, on the other hand, occurred when the fable was prohibited as meaningless, as the product of an uncoherent and unacceptable use of language. The actual obliteration of the true world rests on the belief that logically necessary truths can be stated about the apparent world, which, for this very reason, is the only object of

⁸ David Hume, *Enquiry Concerning Human Understanding*, as quoted by A. J. Ayer in the “Editor’s Introduction” to *Logical Positivism*, p. 10.

true knowledge. That is why I suggest that logical positivism can be deemed to be a further step in the trajectory sketched by Nietzsche: the true world is indeed abolished together with the elimination of metaphysics and any discourse on things whose descriptions are not logically implied by the axiomatic definitions of the most elementary items composing the only accessible and true reality.

The originality of logical positivism is that the impossibility of metaphysics as a science does not depend upon what can be known, but upon what can be said. As Carnap puts it:

A language consists of a vocabulary and a syntax, i.e. a set of words which have meanings and rules of sentence formation. These rules indicate how sentences may be formed out of the various sorts of words. Accordingly, there are two kinds of pseudo-statements: either they contain a word which is erroneously believed to have meaning, or the constituent words are meaningful, yet are put together in a counter-syntactical way, so that they do not yield a meaningful statement. We shall show in terms of examples that pseudo-statements of both kinds occur in metaphysics. Later we shall have to inquire into the reasons that support our contention that metaphysics in its entirety consists of such pseudo-statements.⁹

Accordingly, while the “true world” is abolished as the object of pseudo-statements, the apparent world is revealed to be the referent of statements that can be proven true or false and can therefore be truly known. From this standpoint, knowledge is a process that consists of finding the axiomatic premises that imply the macroscopic and phenomenal descriptions as consequences. This process unfolds by testing the hypotheses that are supposed to be supported by the available evidence, which allows one to verify the degree of logical implication between premises (atomic observational statements) and consequences (predictions): the aim is to judge the truth value of any scientific hypothesis by checking if it can be analytically derived from the definitions of elementary terms. Such definitions concern the behaviour (how they are affected and how they affect each other) of elementary entities, the interactions among which produce more complex objects characterized by specific ways of acting and reacting. Beings should then be defined according to their causal role within a network of beings. Any being can therefore be identified as a stable pattern of

⁹ Carnap, “The elimination of metaphysics through logical analysis of language”, p. 61.

relations (which does not vary under the action of external forces). Thus, philosophy is no longer the science of the supersensible objects of thought, but a critical discipline that analyses propositions in order to assess their truth value according to purely logical and rational criteria.

In order to fully understand why Nietzsche claims that the abolition of the true world also entails the abolition of the apparent world, I propose to consider more recent developments regarding Carnap's insights. In particular, I will take into account Chalmers's thesis that virtual beings are objects of true knowledge in the same way that material entities are.

***The Matrix* as Metaphysics**

I propose to consider *The Matrix* as a version of the story of the true and the apparent world that completes the series sketched by Nietzsche. As is well known, this movie is about the discovery that the world we live in is a computer simulation and that we are actually brains in a vat under the control of intelligent machines that use human bodies as batteries. After realizing the truth, the protagonist decides to fight the malevolent AI and liberate humanity from slavery. At first sight, it might seem that *The Matrix* is a contemporary edition of the allegory of the cave; however, the similarities are only superficial. The essential difference is that most beliefs about the physical organization of reality are the same inside and outside the simulation, so that once we are disconnected from the Matrix, we find no "true" knowledge, but a world where most of the truths we possessed are still valid. To put it another way, the objects of true knowledge are the same inside and outside the cave since such objects are not material beings but regular patterns of brain stimulations. Such patterns can be perfectly instantiated by bits of information, which is why we might not realize that we are brains in a vat rather than wired to a genuine reality. The hypothesis that we live in a simulation without being aware of it can then be considered proof that any perceived world is the effect of brain stimulation and that it does not make a big difference if the causes of phenomenal representations are material or digital objects (streams of data). The "Matrix Hypothesis" is not a sceptical argument meant to cast doubt on the truth of our scientific knowledge of reality, but a thought experiment confirming that our knowledge of the apparent world is true independent of the nature of the entity our brains are wired to. As Chalmers explains:

I think that even if I am in a matrix, my world is perfectly real. A brain in a vat is not massively deluded (at least if it has always been in the vat). Neo does not have massively false beliefs about the external world. Instead, envatted beings have largely correct beliefs about their world. If so, the Matrix Hypothesis is not a skeptical hypothesis, and its possibility does not undercut everything that I think I know.¹⁰

To understand why the hypothesis that we might live in a simulation is meant to prove that our knowledge is true independent of the nature of the objects that stimulate the brain, it is important to clarify Chalmers's epistemological perspective. His main source of inspiration is Carnap's structuralism,¹¹ i.e. the idea that it is possible to understand the most elementary physical relations (how atomic elements are defined by their respective mode of interaction) in a way that enables us to determine all possible macroscopic effects. However, Chalmers does not commit to the assumption that the structure of reality is reflected by logical relations and that the totality of truths can be deduced from the definitions of primitive observational terms. Rather, he thinks that acceptable inferences should not contradict the set of the most fundamental beliefs and that such beliefs constitute a coherent ground of reliable hypotheses that, even though they cannot be taken to be absolutely true, are the premises that allow one to judge the truth value of the proposed consequences. As a cognitivist, Chalmers believes that the very fact that we represent the world as a coherent set of phenomena that can be identified as regular patterns of brain stimulations depends upon the underlying causal structure of reality. Since reality is a network of entities defined by the nomic or causal role they play with respect to each other,¹² brain states are themselves causally determined within the same web of structural relations. In agreement with cognitive philosophers such as Thomas

¹⁰ David J. Chalmers, "The Matrix as Metaphysics", <http://consc.net/papers/matrix.html>, accessed 25 July 2021, p. 2. This paper was originally written for the philosophy section of the official *The Matrix* website (2003) and was subsequently published in C. Grau, (ed.) *Philosophers Explore the Matrix*, Oxford, Oxford University Press, 2005. The same argument appears in Chalmers's book *Constructing the World*, Oxford University Press, 2012, in the excursus "The Structuralist Response to Skepticism" (pp. 431–440). I am quoting the text available on David J. Chalmers's website.

¹¹ See the "Introduction" in Chalmers, *Constructing the World*.

¹² *Ibid.*, p 432.

K. Metzinger and Andy Clark,¹³ Chalmers claims that knowledge can be obtained through the process of making hypotheses about the structures (i.e. real objects) that cause recurrent patterns of impressions. Accordingly, the representation of reality is the effect of the recognition of regular relations among received pieces of information (phenomenal objects are like virtual maps of the possible causes of recurrent series of stimulations). Scientific enquiry thus consists of making increasingly precise hypotheses about the causes of the regularly observed patterns, which are supposed to depend upon the rules of interactions among unobservable particles. In short, brain states are determined by interaction with real objects, while phenomenal objects are hypothetical virtual maps based upon regular patterns of stimulation that are supposed to identify the external causes of perceived changes in the brain, e.g. the brain state that corresponds to the perception of a table depends on matching between the hypothetical pattern of stimuli that is supposed to characterize a table and the pattern of present impressions (information input). Moreover, the pattern of perceptions that identify a table can be explained in terms of interaction among more basic entities. Perceived patterns can be considered to be the effects of the rules that determine interactions at the most basic level, rules being the hypotheses as to the premises from which all possible true brain states are derived as consequences. This means that ordinary concepts – those associated with the representation of macroscopic objects or with recurrent patterns of brain stimulation – ought to be inferred from scientific concepts about the most basic structural patterns. Accordingly, beliefs about the underlying physical structure of the world can be used to clarify the relations among the concepts that are associated with the brain state caused by macroscopic objects. Normative beliefs, then, are rules of inference that allow us to judge the truth value of descriptions of the apparent world without supposing a logical structure of reality but which still assume that a coherent set of beliefs about reality can be adopted. As Chalmers explains:

The inferential role of a concept can be construed as a normative role, constituted by good inferences that the concept might be involved in. On one construal, the inferential role will be an a priori role, involving the a priori justified inferences that the concept is involved in. On another, it will be a sort of analytic role, involv-

290

¹³ See Thomas K. Metzinger “Why Is Virtual Reality Interesting for Philosophers?”, *Frontiers in Robotics and AI*, 5 (101/2018), pp. 1–19; Andy Clark, *Natural Born Cyborgs: Minds, Technology and the Future of Human Intelligence*, Oxford, Oxford University Press, 2003.

ing the trivial or cognitively insignificant inferences that the concept is involved in. Either way, inferentialism will give special weight to *entry inferences*: good inferences from thoughts constituted by primitive concepts alone to thoughts involving the concept in question.¹⁴

While learning consists in confirming the hypothesis about the very basic rules that support further inferences, knowledge is the capacity to produce correct inferences (to assert true consequences from given premises). This capacity allows conscious beings to anticipate future states starting from an observed present state in such a way that they can select the actions whose consequences are the most desirable in the scenario that will probably be actualized.

Now, the reason why Chalmers claims that we can gain true knowledge in virtual reality is that we can actually control our avatars in order to interact in a satisfactory way (for example, in a game scenario) and that such knowledge depends on the same learning process that we apply in material worlds: in any case, it is a matter of making correct inferences about the events that should be expected when some conditions are observed. In the same way that we can grasp the underlying structure that determines the regular patterns of phenomena in a material environment, we can also understand the basic rules of interaction among the digital objects that stimulate our brain as streams of data. Moreover, when a virtual reality is, like in *The Matrix*, a simulation of the material world, we correctly expect the basic structures of the two worlds to correspond to each other, although they are instantiated by different kinds of structures (material and digital). Since knowledge is based on patterns of brain stimulations, the capacity to make good inferences can be acquired independently of the nature of the objects we causally interact with: the structure of digital objects causes brain states that are analogues of those caused by the structure of material objects. As a consequence, most of the beliefs that are true in the material world are also true in the simulated world. As Chalmers explains:

Carnap held that all our hypotheses about the external world are in effect structural hypotheses, concerning the existence of objects satisfying a certain structure. And he held that we can know in principle whether any structural hypothesis obtains. If so, then we can know that the external-world truths obtain. Even if

¹⁴ Chalmers, *Constructing the World*, p. 463.

we are in the evil-genius scenario, or the Matrix scenario, the relevant structure among our experiences obtains.¹⁵

In the Matrix scenario, a computer is simulating the structure of the world and, even though this structure is made of data, the brain states obtained and their relations are exactly the same. For example, the brain state corresponding to the perception of a table will be the same when the table is touched by my organic hands or by my avatar's hands. The material and the virtual are both apparent worlds since they are the phenomenal effects of the interaction between the brain and the structure of the elements that stimulate it. That they are both apparent worlds means that the qualities that we attribute to things, for example colours or smells, are but the effects of the interaction between regular causal patterns and the brain: a material apple is not red; it is the non-coloured structure of the apple that causes the brain state of redness, in the same way a digital apple is a pattern of data that produces the conscious state of redness.

I think it is much better to hold that even after Galileo, ordinary claims such as "The apple is red" are true. The apple is not Edenically red, but it is structurally red: that is, it has the property that plays the structural role associated with redness in causing experiences and the like. Likewise, ordinary claims such as "The apple is round" are true. The apple is not Edenically round, but it is structurally round: that is, it has the property that plays the structural role associated with roundness, in causing experiences and the like.¹⁶

Thus, it does not make a real difference if the redness of the apple is caused by a material object or by a digital object. In both cases, we can obtain the knowledge necessary to predict the future states of the apparent world in order to interact efficiently. Any apparent world can be truly known and the truth of this knowledge rests on the fact that, in order to be a coherent set of phenomena, a world must be a pattern that emerges from the rules that determine the behaviour of its fundamental elements.

The inferences that are correct in the material world are valid in its simulation as well: rather than a sceptical argument casting doubt on the truth of our knowl-

¹⁵ *Ibid.*, p. 431.

¹⁶ *Ibid.*, p. 439.

edge, the Matrix Hypothesis can be considered to be a thought experiment that proves that true knowledge concerns the formal aspects of the underlying structure of the world and that such knowledge is valid independent of the qualitative nature of this structure. This entails that a virtual object can be considered to be a real object rather than a fictional object:

Simulations are typically designed to replicate the abstract causal organization of an original system. A property such as being a calculator depends only on this organization, which is also present in a simulation, so a simulated calculator is a calculator. The same reasoning explains why a virtual calculator is a calculator.¹⁷

However, the question as to the actual nature of the underlying structure, i.e. if it is material or digital, cannot be answered by science, since the latter can only access the fundamental functions (the respective roles played by fundamental elements) that characterize the structure. Such a question is then, according to Chalmers, a legitimate metaphysical question:

Where physics is concerned with the microscopic processes that underlie macroscopic reality, metaphysics is concerned with the fundamental nature of reality. A metaphysical hypothesis might make a claim about the reality that underlies physics itself. Alternatively, it might say something about the nature of our minds, or the creation of our world. [...] In particular, I think the Matrix Hypothesis is equivalent to a version of the following three-part Metaphysical Hypothesis. First, physical processes are fundamentally computational. Second, our cognitive systems are separate from physical processes, but interact with these processes. Third, physical reality was created by beings outside physical space-time.¹⁸

Accordingly, the Matrix Hypothesis is an acceptable metaphysical claim and, even though it is not possible to verify it from within the simulation (it does not produce true knowledge), the proposed speculation does not contradict most of the normative beliefs we are rationally committed to. As Chalmers explains, the Matrix Hypothesis is equivalent to other metaphysical theories about the nature of the structure that is the object of scientific knowledge from the perspective of its functioning (science knows how elements work independent of their qualita-

¹⁷ Chalmers, "The Virtual and the Real", p. 325.

¹⁸ Chalmers, "The Matrix as Metaphysics", p. 3.

tive being: a lever, for example, is defined by its function rather than by the matter it is made of). For instance, the hypothesis that the world has been created by a god standing outside space-time is equivalent to the hypothesis that our reality is a computer simulation that has been programmed by intelligent machines. However, metaphysics is not considered to be a real science that produces true knowledge of the supersensible realm in which the phenomenal appearance of the objects of experience is grounded. Metaphysical speculations are irrelevant with respect to the justification or invalidation of the scientific construction of phenomenal reality. Following the positivist endeavour of liberating scientific enquiry from philosophical justification, metaphysics can be considered to be overcome as it is no longer needed to legitimize the truth of our statements about experienced reality. Moreover, metaphysical discourses are now judged according to the criteria of science as statements whose truth value cannot be assessed. Yet, while Carnap defined them as meaningless, Chalmers claims that they can be considered to be meaningful even though they do not convey any useful information. After all, scientific hypotheses can be judged to be true or false independent of metaphysical assumptions.

If one believes that God created the world, and if one believes that God is outside physical space-time, then one believes the Creation Hypothesis. One needn't believe in God to believe the Creation Hypothesis, though. Perhaps our world was created by a relatively ordinary being in the "next universe up", using the latest world-making technology in that universe. If so, the Creation Hypothesis is true. I don't know whether the Creation Hypothesis is true. But I don't know for certain that it is false. The hypothesis is clearly coherent, and I cannot conclusively rule it out. The Creation Hypothesis is not a skeptical hypothesis. Even if it is true, most of my ordinary beliefs are still true.¹⁹

294

Metaphysical reflections, then, do not provide any access to a true world with respect to which the reality we are immersed in would be a mere appearance. Rather, metaphysical hypotheses are acceptable when they do not contradict the truth of normative beliefs about the perceived world. These normative beliefs refer to the functions of the elements that compose the world as a mechanism that causes phenomenal representations in the brain. Everything is thus considered to be part of a machine whose structure determines illusory qualitative images

¹⁹ *Ibid.*, p. 5.

in the brain, a machine with which the brain can interact to bring about desired effects (rewarding brain states). Knowledge of the causal structure of reality allows conscious beings to engage in practices that depend upon the recognition of the possible functions that a being can assume within a network where other beings express a specific mode of reacting in response to incoming inputs and acting by releasing outputs. Here, truth is a matter of formulating inferences that allow one to efficiently interact with the environment by anticipating the results of the actions that are supposed to bring about the desired reactions. Hence, true knowledge is independent of any metaphysical assumption about the nature of known beings that are known merely as expected series of actions and reactions.

However, even though truth no longer depends upon belief in the true world (the exterior cause of the being of the apparent world) and true knowledge is supposed to concern the world we are immersed in, this makes of the scientific representation of reality nothing more than a fable, a fable about the mathematical structure that characterizes beings and makes their digital simulation possible. As Nietzsche claims in *The Will to Power*, it is because we simplify appearances by reducing them to the repetition of identical cases that we are justified in believing that they depend upon a logical or mathematical structure.

Our psychological perspective is determined by the following:

1. that communication is necessary, and that for there to be communication something has to be firm, simplified, capable of precision (above all in the identical case). For it to be communicable, however, it must be experienced as adapted, as “recognizable”. The material of the senses adapted by the understanding, reduced to rough outlines, made similar, subsumed under related matters. Thus, the fuzziness and chaos of sense impressions are, as it were, logicized.
2. the world of “phenomena” is the adapted world which we feel to be real. The “reality” lies in the continual recurrence of identical, familiar, related things in their logicized character, in the belief that here we are able to reckon and calculate.²⁰

²⁰ Friedrich Nietzsche, *The Will to Power*, trans. W. Kaufmann and R. J. Hollingdale, New York, Vintage Books, 1968, § 569 (Spring-Fall 1887), pp. 306–307.

Accordingly, the idea that the world consists of a series of facts that can be objectively inferred from knowledge of the basic mathematical structure of reality is a consequence of the falsification of sensible events. As Nietzsche puts it:

Our subjective compulsion to believe in logic only reveals that, long before logic itself entered our consciousness, we did nothing but introduce its postulates into events: now we discover them in events – we can no longer do otherwise – and imagine that this compulsion guarantees something connected with “truth”. It is we who created the “thing”, the “identical thing”, subject, attribute, activity, object, substance, form, after we had long pursued the process of making identical, coarse and simple. The world seems logical to us because we have made it logical.²¹

The mathematical reconstruction of reality is certainly useful in order to survive in the reality we are part of, but the usefulness of a hypothesis with respect to human ends does not prove its truth. According to Nietzsche, the abolition of the metaphysical world as the real source of truth merely resulted in making logic the new criteria for establishing what ought to be considered true: “logic would be an imperative, not to know the true, but to posit and arrange a world that shall be called true by us.”²² It is because propositions regarding reality must be evaluated with respect to the posited rational rules (the formula that describes how fundamental entities behave) that any world that can be correctly described is considered real, even though such reality remains nothing but a non-necessary construction. This construction, however, is useful, so its value is relative to human goals: rationality imposes the arbitrary demand of constructing the common world as a norm, but does not provide any true knowledge of the sensible world, whose sensual appearance exceeds the artificial simplification that is imposed on it. Moreover, it is because rationally constructed worlds are environments in which we can actually make successful decisions that we think that the truth of our knowledge about them has been proven. However, as Nietzsche observes:

In valuations are expressed conditions of preservation and growth. All our organs of knowledge and our senses are developed only with regard to conditions of pres-

²¹ *Ibid.*, § 521 (Spring-Fall 1887), pp. 282–283.

²² *Ibid.*, § 516 (Spring-Fall 1887; rev. Spring-Fall 1888), p. 279.

ervation and growth. Trust in reason and its categories, in dialectic, therefore the valuation of logic, proves only their usefulness for life, proved by experience – not that something is true.²³

Since stable beliefs are needed in order to survive and endure as defined identities, we project in the real world the existence of stable beings (which can be recognized by observing the regularities of the events that are supposed to be the effects of their interactions). Nevertheless, this useful construction of the world is a falsification of reality, an apparent world that we ought to believe to be true because we cannot actually live without this kind of a lie. That is why the abolition of the true world as a meaningless fable brought about also the dismissal of the apparent world: with the apparent world being reduced to a rational construction based on idealized elements, we have lost the overwhelming sensual excess of which appearance actually consists.

The Unbearable Truth of Reality or Why We Need Lies

If the true metaphysical world is a fable and the apparent world has been substituted by a useful construction, what is truly real? According to Nietzsche, the truth is that there is no truth because there are no stable beings (about which something true can be said): the world is a pure becoming and becoming is a pure appearance without underlying appearing beings. As he claims, “the antithesis of this phenomenal world is not ‘the true world’, but the formless in-formulable world of the chaos of sensations – another kind of phenomenal world, a kind ‘unknowable’ for us.”²⁴ The world in itself is the same as it is actually perceived: a chaos of sensual properties and qualities that cannot be known without being betrayed and reduced to a collection of simplified entities or constant and quantifiable relations. However, knowledge of the world as pure appearing (rather than as the appearance of some idealized set of elementary and mathematized beings) is not possible, since nothing actually true can be said about it. Hence, “the world with which we are concerned is false, i.e., is not a fact but a fable and approximation on the basis of a meager sum of observations; it is ‘in flux’, as something in a state of becoming, as a falsehood always changing

²³ *Ibid.*, § 507 (Spring-Fall 1887), p. 275.

²⁴ *Ibid.*, § 569 (Spring-Fall 1887), p. 307.

but never getting near the truth: for—there is no truth.”²⁵ As a consequence, the scientific representation of the apparent world is a useful lie. The problem, according to Nietzsche, is that this lie, which allows the human species to preserve itself, also prevents the expression of humankind’s highest values. Even though scientific knowledge as an attempt to dominate nature by reducing it to a set of repeatable simplified beings is an expression of the *will to power*, it is an inadequate expression that limits the creative forces of becoming by imposing forms whose value is relative to the satisfaction of human needs rather than being equal to the absolute value of creative forces. The nihilism of positive knowledge depends on the fact that the world is conceived as a dead mechanism deprived of “life” and “will”, a mechanism without meaning that tends toward a final state of equilibrium or entropic dissolution (the final death of everything rather than eternal self-creation through self-destruction). Moreover, the representation of reality as a set of parts defined by their respective functions includes humans whose performances (practical and cognitive) are judged with respect to the rational norms that are supposed to ensure the dominant role of the species and its growth rather than the expression of the fundamental Dionysian instincts. The rules governing the construction of the scientific world are imperatives that establish the correct or valuable inferences that rational beings ought to make in order to preserve an order of the world that is suitable for the welfare of the population. As Nietzsche claims:

The task is to make man as useful as possible and to approximate him, as far as possible, to an infallible machine: to this end he must be equipped with the virtues of the machine (– he must learn to experience the states in which he works in a mechanically useful way as the supremely valuable states; hence it is necessary to spoil the other states for him as much as possible, as highly dangerous and disreputable).²⁶

298

Such an “economic evaluation” of human beings is, according to Nietzsche, the most extreme danger, the real expression of nihilism as the negation of the will to power, the suppression of the truth that men are not defined by their preferences and social roles but by being equal to the anti-identitarian forces of becoming. Rather than being brains in a vat that are mechanically compelled

²⁵ *Ibid.*, § 616 (1885-1886), p. 330.

²⁶ *Ibid.*, § 888 (Spring-Fall 1887; rev. Spring-Fall 1888), p. 474.

to behave according to the laws of rationality to maximize individual utilities, men ought to aspire, according to Nietzsche, to be artists, to embrace the real sensible chaos of becoming and celebrate being part of the irrational power of the whole.

The extreme point of nihilism is reached when humans are “truly” known or rationally constructed as automatons that perform correct calculations of the actions needed to achieve the given goals in a world reduced to the virtual reality of a game. The extreme point of nihilism corresponds to the reduction of the human to a pattern of information that can be known or modelled in the same way as any other system is known and modelled, i.e. as a series of operations that can be simulated by an algorithm. This stage corresponds, in fact, to the total obliteration of the difference between representation (the constructed phenomenal world) and reality since the representing subject has been finally reduced to a constructible phenomenon whose value is that of a means like any other. In particular, humans – like anything else – are reduced to exploitable resources of information and are thus an object of evaluation according to economic criteria: that is why, according to Nietzsche, a hero or a new species of men is now needed to reverse all values.

The reduction of man to an algorithmic procedure is the condition for creating a perfect virtual reality. In fact, comprehending the functioning of cognition is necessary in order to provide the brain with all the information necessary to obtain the representation of a coherent reality where one can act efficiently while holding true beliefs. The understanding of cognition concerns the elucidation of the rules of rational decision-making, which are based on the correct anticipation of the unfolding events (since the whole process was generated by the algorithmic rules that identify phenomena). Knowledge of these rules can be used in order to predict individuals' behaviour in such a way so as to select the best decisions with respect to expected strategies: this is, for example, what the predictive algorithms we interact with on digital platforms do. However, the reduction of humans to rules on decision-making based on predictive models of phenomena and specific preferences make them not only predictable by machines, but also models for constructing intelligent machines. We are about to reach the point at which humans are no different than their artificial simulations or artificially intelligent systems. However, in order to become models for the efficient behaviour of machines, humans have to submit to the norm of constructing the

kind of reality where utilities can actually be maximized and where intelligent algorithms can be successfully used to improve strategies and to produce useful knowledge. The infinite possible states of such a reality – the market reality is a good example – can be computed by accessing the rules that determine the respective actions and reactions that produce observable events. An apparent world in which anything that can happen can be anticipated (even though it is not always correctly predicted) and therefore simulated, a world where types of agents are modelled as algorithms that compete to satisfy a given set of preferences by learning the preferences of others, is a lie that is so strongly believed that it comes to be the formula of any possible reality. If any possible objectively known world can be reduced to a computer simulation, then any immersive and coherent virtual reality game can be correctly believed to be real since one can learn to interact efficiently by learning its rules. We believe so strongly that the world is truly a network of encounters among beings defined by their behaviour (algorithmic rules) and that such knowledge sets us free to satisfy our preferences that we no longer feel the chains that force us to construct the world as a rational game that can be successfully played. It no longer matters whether we know the origin of the structural laws that organize the world we are immersed in, as far as they are perceived as the necessary conditions of any rational world-making. It seems even more acceptable to speculate that such rules have been made by a foreign AI and that we live in a computer simulation than to admit that the rules have been made by us and that we are the authors of the lie that we believe to be the true structure of reality. Against the type of man that takes the lie of the algorithmic definition of beings to be the only true condition of any apparent world, Nietzsche opposes the artist, who knows that we cannot survive without lies and that the only authenticity consists of being faithful to the creative power of lying by contributing to the ever-changing appearance of a world where no identity appears.

300

Metaphysics, morality, religion, science – in this book these things merit consideration only as various forms of lies: with their help one can have *faith* in life. “Life *ought* to inspire confidence”: the task thus imposed is tremendous. To solve it, man must be a liar by nature, he must be above all an *artist*. And he *is* one: metaphysics, religion, morality, science – all of them only products of his will to art, to lie, to flight from “truth”, to *negation* of “truth”. This ability itself, thanks to which he violates reality by means of lies, this artistic ability of man *par excel-*

lence – he has it in common with everything that is. He himself is after all a piece of reality, truth, nature: how should he not also be a piece of *genius in lying*?²⁷

Nietzsche's conception is the perfect reversal of Chalmers's. For the latter, any fiction is "truly real" under the condition of being logically coherent, of depending on a knowable structure that causes its correct representation in a rational mind, i.e. a brain that performs the calculations that are needed to preserve itself in any environment, while for Nietzsche no constructed world is real, and thus any reality that is believed to be truly known is a fiction. Hence, there is only one reality and nothing true can be said about it. Nevertheless, besides "whereof one cannot speak, thereof one must be silent," there is another option. This option is not available to men of logic but only to the artist, for whom the chaos of becoming appears in the form of an artwork, a form that does not make it knowable as a defined being but does make it sensible. An artwork is a form imposed on becoming as an affirmation of the will to power, a lie that does not pretend to be the truth but is sincerely offered as a lie, as a pure appearing (rather than the appearance of something defined or simplified). The artist is faithful to the truth of the impossibility of truth and he or she offers his or her lies to contribute to the creativity of always changing forms through which the highest power of the will is expressed: "art is worth more than truth."²⁸ In creating forms as true lies, the artist aspires to be identical to the will to power rather than to be an identity acting in agreement with human valuable ends. In an artwork, the unrepresentable becoming of which reality consists is embraced for what it is, without any ambition to provide a true image of it, but with the pride of being part of the flux in which any stable form is, at the same time, a transitory expression of the whole and which will come back eternally as an instant whose value cannot be compared or measured.

301

Conceived as becoming or will to power, the apparent world is finally liberated from the fable of the "true world", from the "metaphysical hypothesis" that somebody or something might have created it from the outside as a determinate being (a structure). As we have seen, this hypothesis cannot be ruled out by Chalmers and by all the philosophers who conceive of reality as an effect of the being of a structure, of a set of rules that cannot explain their own coming

²⁷ *Ibid.*, § 853 ("Art in the 'Birth of Tragedy'"), pp. 451–452.

²⁸ *Ibid.*, p. 453.

into being. For Nietzsche, on the other hand, the metaphysical hypothesis of the eternal recurrence excludes the possibility of believing in a creator of the world. As eternal recurrence, reality is a circular becoming that has neither a beginning nor an end, an infinite flux limited by nothing and constantly producing and destroying anything it can produce by reproducing itself. As Nietzsche poetically describes in the last pages of *The Will to Power*:

This, my Dionysian world of the eternally self-creating, the eternally self-destroying, this mystery world of the twofold voluptuous delight, my “beyond good and evil”, without goal, unless the joy of the circle is itself a goal; without will, unless a ring feels good will toward itself – do you want a name for this world? A solution for all its riddles? A light for you, too, you best-concealed, strongest, most intrepid, most midnightly men? – This world is the will to power – and nothing besides! And you yourselves are also this will to power – and nothing besides.²⁹

While the man of knowledge evaluates the states of the world with respect to utility and himself with respect to the efficacy of his inferences, the artist-philosopher evaluates the states of the world according to their incomparable value of necessary manifestations of becoming and himself as becoming equal to the will to power, thus contributing to the creation of sensible appearances.

References

- Ayer, A. J., “Editor’s Introduction”, in A. J. Ayer (ed.), *Logical Positivism*, New York, Free Press, 1959, pp. 3–28.
- Beane, Silas R., Zohreh Davoudi, and Martin J. Savage, “Constraints on the Universe as a Numerical Simulation”, *The European Physical Journal A*, 50 (148/2014), pp. 1–9.
- Bostrom, Nick, “Are you living in a computer simulation?”, *Philosophical Quarterly*, 53 (211/2003), pp. 243–255.
- Carnap, Rudolph, “Überwindung der Metaphysik durch Logische Analyse der Sprache”, *Erkenntnis*, 2 (1/1932), pp. 219–241.
- Carnap, Rudolph, “The Elimination of Metaphysics through Logical Analysis of Language”, trans. A. Pap, in A. J. Ayer (ed.), *Logical Positivism*, New York, Free Press, 1959, pp. 60–81.
- Chalmers, David J., “The Matrix as Metaphysics”, <http://consc.net/papers/matrix.html>, accessed 25 July 2021.

²⁹ *Ibid.*, § 1067 (1885), p. 550.

- Chalmers, David J., *Constructing the World*, Oxford, Oxford University Press, 2012.
- Chalmers, David J., “The Virtual and the Real”, *Disputatio*, 9 (46/2017), pp. 309–352.
- Clark, Andy, *Natural Born Cyborgs: Minds, Technology and the Future of Human Intelligence*, Oxford, Oxford University Press, 2003.
- Metzinger, Thomas K., “Why Is Virtual Reality Interesting for Philosophers?”, *Frontiers in Robotics and AI*, 5 (101/2018), doi: 10.3389/frobt.2018.00101, pp. 1–19.
- Nietzsche, Friedrich, *The Will to Power*, trans. W. Kaufmann and R. J. Hollingdale, New York, Vintage Books, 1968.
- Nietzsche, Friedrich, *Twilight of the Idols*, in *The Antichrist, Ecce Homo, Twilight of the Idols, and Other Writings*, trans. J. Norman, Cambridge, Cambridge University Press, 2005.
- Plato, “The Allegory of the Cave”, trans. T. Sheehan, <https://web.stanford.edu/class/ihum40/cave.pdf>, accessed 25 July 2021.
- Wheeler, John Archibald, “Information, Physics, Quantum: The Search for Links”, in W. H. Zurek (ed.), *Complexity, Entropy, and the Physics of Information*, Boston, Addison-Wesley, 1990, pp. 3–28.