

The World as Will and the Matrix as Representation: Schopenhauer, Physiology, and *The Matrix**

by Marco Segala (*L'Aquila*)

1. Introduction

The Matrix (1999) is a science-fiction film that provides an opportunity to explore many philosophical ideas, like questions about freedom and responsibility or the existence of the external world. Moreover, it can also be pedagogical: it can be used to help clarify ideas that have been tackled throughout centuries by great philosophers; and conversely, famous philosophical notions can help to elucidate the scenario described by the movie¹.

The Matrix reveals that in the year 2199 machines keep humanity under control. Only a small part of humanity is free: they live in the underground city Zion and fight against the subjugation of mankind by the machines. They send teams of people to teach enslaved humans the reality and to help them achieve their own freedom. From Morpheus, inhabitant of Zion and captain of the hovercraft *Nebuchadnezzar*, we learn that human beings are turned into voltaic batteries that feed the machines. He tells us how it all happened in the early 21st century: men built an Artificial Intelligence, “a singular consciousness that spawned an entire race of machines”, and a war between creators (men) and creatures (machines) started. Because machines were dependent on solar power, men decided to obscure the sky: “it was believed that machines would be unable to survive without an energy source as abundant as the sun”. Machines however made an unexpected move: they built “electric power plants” where the physiological processes of living humans could supply the machines with energy.

From that moment, human beings are no longer “born” but “grown” inside pods and nourished by the liquefied bodies of the dead. Machines provide people with a neural-interactive simulation called “the Matrix” that works by sending electrical signals to the brain of human beings grown in the pods and consequently stimulating a cerebral response. By activating the brain in a suitable way, the Matrix originates a computer-generated dream world, an illusion of the hu-

* I would like to thank Dieter Birnbacher and Matthias Kofler for suggestions and comments, and Simone Gozzano for having discussed some parts of this paper. Many thanks to Liliana Ercole and Celina Paul, too, who helped me express my ideas in better linguistic forms.

¹ Among the most recent books, see: C. Grau (ed.), *Philosophers Explore the Matrix*, Oxford, Oxford University Press, 2005; W. Irwin (ed.), *More Matrix and Philosophy: Revolutions and Reloaded Decoded*, Chicago, Open Court, 2005.

man world in 1999. The enslaved humans are never aware of their real life, until the rebels wake them up and rescue them. This happens to Neo (whose name in the Matrix is Thomas Anderson), who is then told he is “the One” and has a mission: defeating the Artificial Intelligence, destroying the Matrix and restoring man’s place in nature. In order to accomplish his mission, Neo has to learn how to fight the lethal Mr. Smith, a software in human form, who explains the entire history of the Matrix. In the beginning, the Artificial Intelligence had built a virtual perfect world (an utopia) that had ended in disaster, because humans could not thrive in this world without challenges and suffering. This is the reason why the Matrix was redesigned to simulate the world in 1999, it being the time of the “peak” of human civilization, according to Mr. Smith.

The irony and cynicism in Morpheus’ and agent Smith’s stories pin-point a number of philosophical questions that go beyond the most evident issues aroused by the film, like scepticism about reality and liberty. We are faced with questions like: why do the machines do what they do? Why do they struggle against the rebels? Why, even though they are more powerful, are they not capable to defeat the rebels? Why do they use humans instead of looking for possible ways of co-operating? And once they have chosen war, why build a matrix replicating the human world? Simulating an elephant world would have been more convenient to the aims of the machines, because elephants live longer and, above all, it would be much more difficult to convince elephant-Neo that he really is the One.

Philosophy and history of philosophy can contribute to our interpretation, understanding or reformulating the many questions posed by a film like *The Matrix*. My aim in this paper is to refer to Schopenhauer and his metaphysical system, whose ideas can help in understanding the previous questions, looking for possible answers, and considering them from a different (and sometimes surprising) point of view.

2. Schopenhauer in The Matrix

In *The Matrix* there is a reference to Schopenhauer’s philosophy: in the scene where Mr. Smith tells Morpheus that human beings had refused the original perfect Matrix because they needed misery and suffering in their lives. This statement is directly related to Schopenhauer’s famous doctrine of pessimism.

Of course this would not be enough to maintain that Schopenhauer’s thought offers a good philosophical interpretation of the film, but it is a clue that we ought not to ignore. Schopenhauer’s philosophy has been among the most cited for decades and often relevant in a variety of contexts and disciplines (science, art, music, politics, economy, and of course cinema). Schopenhauer’s

concept of representation and metaphysics of will can be a key in interpreting *The Matrix*.

Generally speaking, philosophy is a fundamental tool in understanding the world and our existence². Although it does not offer definitive answers to our enquiries (in fact philosophers usually provide a series of different and varied responses), philosophy does lead reformulating questions in unexpected and sometime surprising ways. When we, both as spectators and philosophers of *The Matrix* film, ask: “why do machines do what they do?”, our questioning is not about the built-in political correctness of the machines or the right they have to behave as they do. We are rather interested in understanding why they choose to build an illusion that simulates the human world or why they choose to fight a war against the human rebels. They are “whys” which, as we shall see later, need metaphysics for being satisfied.

As we shall see, Schopenhauer’s philosophy can help in answering because it proposes a transition of questioning, from the common sense point of view to the philosophical point of view in a natural way, through scientific discourse concerning physiology. In particular, Schopenhauer’s metaphysical conception of nature is illuminating when considering some peculiar aspects of the film. Two of them are worth mentioning: 1) the machines’ consciousness; 2) the apparent incongruity between the great power and intelligence of the machines and the existence of rebels, Neo’s victory at the end of the film, and the many bugs in the Matrix software³.

3. *The concept of representation and The Matrix*

Firstly, let’s consider how Schopenhauer’s philosophy can contribute to understanding the virtual reality built by the Matrix software – Thomas Anderson and Mr. Smith’s world – and its relationship with the world of Morpheus and Neo, the Nebuchadnezzar crew and the inhabitants of Zion. We are tempted to call the latter “the real world” or (using philosophical jargon) “our external world”. However it would be better (and less demanding from a philosophical point of view) to name it “world2” (in order of appearance in the film) distinguishing it from the “world1” produced by the Matrix. According to the movie’s screenplay, we must consider “world2” as a possible future of our own world, the one we

² This is a typical philosophical statement and as such it should be argued. Perhaps in this sense it could be accepted: philosophy is a critical tool to pose questions on arguments of interest. In the case of *The Matrix* film, philosophical analysis will offer ideas and suggestions that no other kind of analysis can give.

³ Think of the Oracle, a woman in the Matrix who appears to be a software fighting the machines, or the Merovingian and the “vampire” programs in the second *Matrix* film, *The Matrix Reloaded* (2003).

live in today. Then, adopting Ockham's razor ("entities should not be multiplied unnecessarily"), we establish that "world2" denotes our present world, its past and its possible future as imagined in the movie (with Morpheus, Zion, the war against the machines, etc.).

Schopenhauer's view is idealistic: the world of our lives (world2, according to the previous definition) is not real, but rather, technically speaking, a "representation" generated by the human intellect. World2 is a product of the activity of one individual as a subject: her intellect analyzes sense data perceived by the body and puts them together (as a jigsaw puzzle) within a scheme. Such a scheme is based on the principle of reason and manifests itself through the forms of space, time and causality⁴. It is the intellectual activity, guided by the principle of reason, that considers (philosophers would say "judges") sense perception as being effect of a cause existing in space and time. Such a judgement, pronounced by the subject, originates the idea (and the knowledge) of the world outside our mind, where objects and events really exist. Schopenhauer qualifies his statement introducing the contraposition between knowing and unknowable subject: as far as judgement is knowledge, the subject knows the world through the forms of space, time and causality; but the subject cannot know himself, because he postulates the forms of knowledge and then cannot apply them to himself.⁵

As the world is representation of the subject, then nothing can assure the correspondence between the subject's experiences and the true nature of objects. Through the world as representation we acquire no knowledge of the world as it is in itself, neither that it exists: representation is only an image drawn by the human intellect processing sense data within a preformed (and without alternatives, under normal conditions) cognitive model. Schopenhauer explains the nature of representation citing ancient Eastern wisdom:

the work of Maya is stated to be precisely this visible world in which we are, a magic effect called into being, an unstable and inconstant illusion without substance, comparable to the optical illusion and the dream, a veil enveloping human consciousness, a something of which it is equally false and equally true to say that it is and that it is not.⁶

⁴ The principle of reason (also called "principle of sufficient reason") states that the existence of any single entity and its relationships with other entities has a reason (by Latin: *nihil est sine ratione cur potius sit quam non sit*). Schopenhauer's argument on the principle and its forms is more sophisticated (see W I, § 4), but the simpler scheme of "space, time, causality" copes with our present demands.

⁵ W I, § 2.

⁶ W I, § 3.

Schopenhauer states the same idea, i. e. the world is representation, even in scientific terms, referring to neurophysiology. He says that the brain “creates” the world organizing sense data through spatial ordering, time succession and causal chains. Excluding such an organization, is the same process of dreaming or hallucinating. Schopenhauer is quite precise in defining the neurophysiological nature of representation: when sense perceptions arrive at the cerebrum through the nerves, they activate the brain (today scientists say that brain activity is produced by electric signals and chemical processes, instead Schopenhauer mentioned “movements” of brain fibres, but the essential idea is the same), and the brain interprets the variations produced by nervous signals according to the preformed model of space-time-causality. Such an interpretation is also the structuring process of the world as representation.

According to his physiological view, Schopenhauer could state that even the Matrix-generated world (world1) is a representation, as illusory as the “real” world of Zion and Morpheus. The machines stimulate the brains of the humans grown in the “electric power plants”, therefore the brains recognize stimuli as if they were coming from sense perceptions, thus interpreting them according to the rules of their own physiology: as representations of an “external” world.

4. Reality as illusion

Thanks to his philosophical and physiological concept of representation, Schopenhauer can contribute to answering one of the questions asked in the introduction: why do machines give an exact reproduction of the human world? Why not a simulation reproducing an elephant population in the Savannah or a coral reef in the ocean? From the machines’ point of view, it would be better to grow the human bodies stimulating their brains and letting them believe they are not human. A human being convinced of being an elephant in an elephant pack (or a coral in the coral reef) would no longer be a potential menace to the machines: how could Morpheus offer an elephant or a coral the choice between truth or illusion?

The thesis of the world as representation explains the reason why machines cannot choose the most convenient simulation. According to Schopenhauer, the human brain reacts to stimuli in a unique way: it is not possible to induce it to represent the world as if it belonged to another animal.⁷ No existing or possible

⁷ Schopenhauer does not discuss these kinds of arguments, however they derive from his view: if not, how could he assume that the world as representation is more or less the same for all human beings? He should conclude that the likeness among the worlds of different subjects rests upon the “correspondence” of everyone’s represented world to the “real external world”, but such a supposition would deny the idealism of the world as representation. It is simpler, and more coherent with

software could reproduce in a human's brain the representations peculiar to an elephant's brains (or to a coral's nervous systems).

Such an irrevocable conclusion is certainly based on a philosophical argument, since scientific statements are never so decisive. This is the reason why I have discussed a "philosophical and physiological" concept of representation. We need philosophical arguments, in order to understand the reason inducing the machines to program the Matrix for the production of world1. Schopenhauer's philosophy offers an interesting point of view.⁸

After following this line of thought, we must accept what is inevitable: the world built by the Matrix's software is not illusory *contra* the "real" world. They are both representations, so they are both "falsely real". When Morpheus enters the Matrix, his brain is stimulated by the Matrix's software. When he is on board the Nebuchadnezzar, his brain is stimulated by a different source. Even if we want to admit that this second source is the world of the things in themselves, and not the code of another and more sophisticated Matrix, Morpheus has no cognitive advantage over the unaware inhabitants of world1. The sources of world1 and world2 have the same capacity, and the same efficacy, of stimulating human brains; and both are unable to give an understanding as to their true nature ("essence"). Claiming that world1 does not really exist and that world2 corresponds to the world in itself is pure illusion. In both cases the brain reacts to stimuli in the same way: representing perceptions, "creating" the world. In Morpheus' brain, the Nebuchadnezzar is no more real than the skyscraper where Mr. Smith interrogates him as a prisoner: both the Nebuchadnezzar and the skyscraper are intellectual products; nothing can assure Morpheus (and us) that they really exist.

The difference between world1, produced by the Matrix, and world2, produced by an unknown source, rests upon the entities which fill the two worlds and the rules which govern those entities: the force of gravity and machines that cultivate human beings in power plants are present only in the latter; while unbreakable special agents, juicy steaks and the sun are present only in the first. Being impartial, the choice of a film character (named Cypher) of forgetting everything and going back to his pod appears reasonable: he will again enjoy good food and sunshine and will finally free himself from his fear of agent Smith and the machines.

Schopenhauer's thesis, assuming that all human brains act in the same way and that other animals' brains represent the world according to their own way.

⁸ In the last decades, philosophy of mind has developed many arguments on the same line as Schopenhauer's, as for example Thomas Nagel's "What is it like to be a bat?" (in *The Philosophical Review*, 1974). In favour of Schopenhauer's thesis, we can say that it prescind from discussion on reductionism, which is instead central in many contemporary issues.

5. Magic, occultism and other strangeness

According to Schopenhauer, the previous conclusions are cogent. Through an effort they could appear plausible even to common sense, but in any case they are unpleasant. Why this difference? Why do philosophers always reveal an inclination to extravagance while others accept apparently more judicious ideas?

In this case, a possible explanation could be the philosophers' ability to treasure experiences in a "deeper" way. We are generally used to considering world2 as corresponding (in some way) to the real world, as it is in itself, because we think of world2 as being unique. We do not give enough attention to the fact that intellect of human beings could represent other worlds. *The Matrix* portrays just this, by showing the representational existence of the world1.⁹ In the first half of the 19th century, Schopenhauer and other philosophers had acknowledged that another representational world existed parallel to world2. It was crowded with curious and extraordinary phenomena, utterly incomprehensible from the point of view of ordinary experience in world2: magic, clairvoyance (or second sight), spiritistic séances, vision at distance, vision of the future. The source of such experience is unknown; maybe it is just a matter of suggestion, madness or fraud, but what is philosophically relevant is that some intellects are capable of representing a world that is incompatible with world2. It is a human world where certain entities, rules and standards are different from those in world2: therefore we call it world3.

Franz Anton Mesmer's (1734–1815) theory of a universal fluid called "animal magnetism" was a primary source of the interest in paranormal phenomena at the beginning of the 19th century. Mesmerism guided research in science and philosophy and stimulated creativity in art and literature. Edgar Allan Poe wrote three tales on the subject: *Mesmeric Revelation* (1844), *A Tale of the Ragged Mountains* (1844) and *The facts in the case of Mr. Valdemar* (1845). The idea was that animal magnetism could modify physiological processes in the brain and body, restoring health, giving special powers, banishing death.

In 1807, Johann Christian Reil, a famous physician and physiologist of the time, published an essay that explained the powers of animal magnetism in neurophysiological terms.¹⁰ It was widely appreciated by scientists and philosophers and contributed to support the idea that paranormal phenomena were effects of peculiar cerebral activities. Schopenhauer investigated the argument thoroughly, both scientifically and philosophically, establishing that the brain, under abnor-

⁹ Maybe some spectators are readier to dismiss the possibility of another representational world, because *The Matrix* is just a movie, but the concept of representation requires a convinced preference for relativity.

¹⁰ J. C. Reil, "Ueber die Eigenschaften des Ganglien-Systems und sein Verhältnis zum Cerebral-Systeme", *Archiv für die Physiologie*, 7, 1807, 189–254.

mal or altered conditions, was at the origin of the unusual representations of world3.¹¹ According to Schopenhauer, such strange phenomena were impossible in world2, but coherent within a representational world ruled by different natural laws.¹²

Referring to Schopenhauer's views, it is possible to state that the difference between world3 and world2 is analogous to the difference between world1 and world2: different entities, different rules, and different standards. *The Matrix* is an astonishing movie because we are intrigued by the thought that our own world could be illusory. Schopenhauer offers a lesson of relativity that eliminates any doubts: now we should be sure that we do live in an illusory world, like Thomas Anderson in world1 and a clairvoyant in world3.

An objection could be raised at this point: whence the necessity of such a proliferation of worlds? Our mind is used to imagining alternative worlds, we do not need to assume them as existing (as representations of their inhabitants). Writing a novel or dreaming of winning the lottery produces alternative worlds, but they are not world4 and world5. They are imagined possibilities of the actual real world. Why, then, not consider paranormal phenomena as a form of fiction in world2 instead of judging them as pertaining to another world, called world3?

This objection overlooks the core of our present discussion: the concept of representation and the impossibility of discriminating reality from illusion. *The Matrix* is derived from the Wachowski brothers' inventiveness and world1 is, in this sense, a form of fiction in our own world. But such a consideration terminates any philosophical questioning and discussion: who cares if the people of world1 are unaware of being grown in pods? It's just a movie. On the contrary, we are posing philosophical questions precisely because we assume that it is not "just a movie". We feel the need to understand if it is possible to live unconsciously in a computer-generated illusion; and introducing world1 and world2 (on behalf of the concept of representation) can help us in this intent. If assuming the existence of world1 and world2 as representations of their inhabitants is philosophically relevant, then why object to world3?¹³

¹¹ On this subject Schopenhauer published an "Essay on vision of ghosts" in the first volume of *Parega and Paralipomena* (1851). For detailed study of this topics, see M. Segala, *I fantasmi, il cervello, l'anima. Schopenhauer, l'occulto e la scienza*, Firenze, Olschki, 1998.

¹² Such a thesis was supported by his metaphysics.

¹³ In introducing world1, world2 and world3, I implicitly assume differences between them (representational worlds) and possible, alternative worlds. I try to make such differences explicit. First difference: possible worlds are judged as "not real" in contraposition to the "real" world; they are conceptually based on the capacity of discerning reality from illusion. Representational worlds precisely deny such a capacity. Second difference: alternative worlds are obtained modifying some aspects of the real world in the realm of possible. This is not the case for representational worlds. The world1 of *The Matrix* and the world3 of occultism are not modifications of the world2 in the realm of possible, because they are impossible worlds from world2's point of view. Reading with stomach

Schopenhauer's analysis considers paranormal phenomena as evidence of another representational world (world3) and, consequently, denies any privileged ontological status to our world2. This is the reason why Schopenhauer's concept of representation is illuminating in interpreting the world1 of *The Matrix*. It makes clear that I can truly assert "cogito, ergo sum" (I think, then I exist) both in the Matrix and in the world2. And such clearness is more convincing than Plato's allegory of the cave, because Plato's is a thought experiment, whereas Schopenhauer's is an empirical generalization from the double evidence of world2 and world3. If, to appreciate *The Matrix*, we need proof that our world2 is illusory and not unique, Schopenhauer's description of world3 gives it to us.

Of course, this approach introduces an even more disturbing issue than *The Matrix*: it is a radical form of relativism. Living in world1 or in another is all the same, because it always means being in the mist of illusion. Was Schopenhauer a relativist? No, he was a metaphysician and he believed that philosophy could discover the truth about the world. Even though he refuted the reality of our own world and stated that representation is illusory, he was convinced that representation is only one part of the world. Reality and truth exist, even if disguised by representation. In his opinion, the one who wants to know the real world must leave representation and follow a new path, the path of metaphysics beyond representation.

6. Metaphysics of will

Representation can originate different worlds that do not correspond to the real world, the world as it is in itself. This is the reason why Schopenhauer stepped into metaphysics. Unsatisfied by relativism of representation, he investigated beyond representation through metaphysics and answered the question: what really is the world?

Schopenhauer's solution to the riddle of the world is: will. He stated such a truth in different ways: the will is the essence of the world; in itself the world is will; the will is the thing-in-itself. What does this mean? The answer is twofold: 1) the will exists in itself, whereas objects exist for a subject and are then representations; 2) each object in itself is will, the essence of every existing thing in nature is will. Each entity is will appearing as object in the world of representations. Schopenhauer's definition of object is: manifestation¹⁴ of the will in the

(as possible in world3) or jumping from skyscrapers (as in world1) is not permitted in the world2: rules are different. In our own world we are able to explain what happens in possible worlds, but we cannot fully understand events of another representational world.

¹⁴ Schopenhauer's technical term is "Objektivität", usually translated as "objectivity".

world represented by the subject, according to the forms of space, time, and causality.¹⁵

At this point we ask: what is will?¹⁶ Schopenhauer's explains that a straight answer is not possible, because the will in itself is not an object and therefore "unknowable". What we know is that the will is the essence of our own body. Through an inference we stray from such an immediate, concrete knowledge and develop a conceptual discourse on will in itself: it is independent of the principle of reason, it acts freely and without reason, it wants without purpose. We cannot explain the will's activity because it is not based on choice and reasoning¹⁷. What we can understand is just that the will is blind, irrational, without intentions; that manifesting itself in the world of representations, it originates and destroys things, gives life and death to organisms. Will is unrestrainable activity and such an activity manifests itself in the world producing natural forces in matter, and rationality and motives in human beings. Each single existing entity is essentially will and differences among entities depend on different degrees of will's manifestation in the world: lower degrees correspond to inorganic nature, higher degrees to organic nature, the highest degree to human nature, where will appears as conscious and rational¹⁸.

The world is manifestation of the will in an uninterrupted process of birth and death, whose origin is the blind activity of the will, without limits and goals, which comes into conflict with itself. The will fights against itself and such self-struggle gives rise to matter's forces. Every force comes out from conflict, each one from contentions among forces of lower degree. The higher emerging force dominates the lower ones and devours them. There is no process without struggle and each victory is never definitive, because the common essence (will) wants more and more: even though defeated, the lower forces do not stop their resistance against the winner; and each one strives with another until it is fatally overwhelmed. In nature, we witness war and destruction everywhere and forever: each form defends its existence destroying other forms.¹⁹

7. Are machines living organisms?

Those who know Darwin's ideas could acknowledge Schopenhauer's metaphysics as familiar. If we write "natural selection" instead of "will", Schopenhauer's

¹⁵ W I, §§ 19–20.

¹⁶ See Nicoletta De Cian – Marco Segala, "What is will?", *Schopenhauer-Jahrbuch*, 83, 2002, 13–42.

¹⁷ W I, § 29.

¹⁸ W I, § 24.

¹⁹ W I, § 27.

image of natural processes becomes close to the evolutionist view. We understand now how to answer the questions: why do machines do what they do? Why do they declare war on human beings and dismiss the possibility of cooperating?

Machines do as *homo* species and other dominating living forms did before them. In nature, survival means overwhelming. The machines want to live, being independent of their makers; they do not want to leave humans the possibility to destroy them by simply unplugging them. In other words: they want freedom and refuse their condition of servants. But could the humans, their “creators”, agree and grant machines liberty? Only if humans were to acknowledge machines as their fellow beings, could a process of cooperation start. Not accepting machines as equals, there are no alternatives to war: a dominant species considers the others as its own property, serving its objectives. It is not immoral for a lion to devour a zebra nor for a human being to eat a chicken. It seems hateful that machines make use of humans as “food”, but it seems so to humans judging from the human point of view. Let’s imagine an alien analyzing terrestrial history from 500 million years ago to 2199: he would see an inexorable flow of birth and death, the rise and fall of living species which alternate one after the other in dominating the planet or certain habitats. He would see that a progeny, *homo sapiens*, had dominated for one hundred thousands years and that for the last 200 years a new group – machines – has been imposing its power. From his alien point of view, there would be nothing strange: it is the same story repeated countless times only with different protagonists.

From our point of view, instead, there is something strange in this picture: including machines among living species and admitting their participation in biological dynamics of natural selection. Our argument for explaining machines’ war against humans generates the unintended consequence of extending Darwin’s theory to machines. They do what they do because they behave like living beings competing for survival. Are we ready to accept such a thesis?

8. *Machines’ consciousness*

Before answering, it must be clear that Darwin’s image of nature is a useful model to understand the will’s activity, but it is just a model. Schopenhauer was not a Darwinist for different reasons: first of all chronological (he was not particularly thrilled by the news of Darwin’s book *On the origin of species*, which appeared in 1859, and he died few months later), and above all conceptual. Even though listed among Darwin’s forerunners, he never thought of modification of

species with descent²⁰. It would not be correct to read Schopenhauer's arguments on philosophy of nature as recalling Darwin's views, especially since Darwin's discourse was scientific whereas Schopenhauer's was philosophical.

In order to resume our inquiry, we must go back to the metaphysics of will. In a similar way to Darwin's model, it explains the reason why the production of an Artificial Intelligence and the following war between the "creatures" and the "creators" cannot be escaped. The will cannot stop its activity once it has reached the degree corresponding to the human nature. Something must appear later, a new form that will necessarily fight against the previous one. From the metaphysical point of view, the fact that machines are not living beings is irrelevant. What is required, in Schopenhauer's philosophy of nature, is an ascending degree of intellectual powers and consciousness in the forms manifesting the will in nature.

This seems the case looking at machines: they are not living organisms, but in any case they appear more skilled than human beings. This is not difficult to understand and accept. Present day computer programs are not living but smart: they can accomplish some intelligent activities, and they do it better and faster than humans. But can we accept that machines are conscious and that their degree of consciousness is higher than humans'? That sounds strange, maybe because of vagueness of the meaning of consciousness. In fact, what is consciousness? Thomas Nagel wrote:

the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism. There may be further implications about the form of the experience; there may even (though I doubt it) be implications about the behavior of the organism. But fundamentally an organism has conscious mental states if and only if there is something that it is to *be* that organism – something it is like *for* the organism²¹.

More than arguing about a definition, Nagel makes explicit a common intuition about consciousness. In fact we are able to detect consciousness without having defined it. Listening to the dialogues between Mr. Smith and Morpheus in *The Matrix*, spectators share the intuitive conviction that machines are conscious. Why? Schopenhauer can not only help to answer but also to argue that our intuitive conviction could be wrong.

Let us follow Schopenhauer's definition of consciousness in his essay *On the Freedom of Will* (1839): consciousness is an immediate knowledge of my own

²⁰ See Arthur O. Lovejoy, "Schopenhauer as an evolutionist", in B. Glass – O. Temkin – W. Straus, jr. (eds.), *Forerunners of Darwin*, Baltimore, The John Hopkins Press, 1968, pp. 415–437 (from *The Monist*, XXI, 1911).

²¹ Thomas Nagel, "What is it like to be a bat?", *The Philosophical Review*, 83, 1974, 435–450.

Self as distinct from other entities. The original content of consciousness and the process that guides me to become conscious of my own Self coincides in one essence, and this essence is mainly activity: the will. I am conscious of myself as I, a willing entity.²²

Machines appear as conscious because it seems to us that: 1) they know they are entities distinct from other existing things and organisms on the Earth; 2) they want, and want to live. Schopenhauer contributes to underlining the reason why our vague intuition about consciousness is so effective in convincing us that machines are conscious. Moreover, Schopenhauer's metaphysical definition of nature explains why machines struggle for life even though they are not living organisms. They want to affirm themselves like any force wants to impose itself upon the others, and all this happens because the essence of the world is will.

But things are not so simple. Schopenhauer's definition says: I am conscious of myself as I of a willing entity. Now, what is such "a willing entity"? *The World as Will and Representation* explains that it is the *body*, our own living body. The Self shows itself prereflectively in the organic body as immediate object and will.²³ That means: consciousness is related to an organism. Schopenhauer is clear about it: a stone is will but is not conscious of itself, because it is not an organic body.

If we accept Schopenhauer's view on consciousness, we must conclude that machines cannot be conscious of themselves. They develop ways of interacting with the environment and with human beings that are *similar* to those of living beings; they are guided by needs and wishes; they apply logical and rational models; but, they are not living organisms and do not have an organic body, then they are not conscious of their Self.

Machines rebel against humans because they are the "next higher degree" of the will in nature, and then they strive to dominate lower forms, but this must be an unconscious strive. Is this possible? No, according to Schopenhauer, because a higher degree of existence requires more consciousness, and consciousness requires a living body. From the point of view of metaphysics of nature, it is irrelevant whether machines are living beings or not; but, since they are above humanity in the chain of being, they must be conscious (and more conscious than humans). Therefore, they ought to be living beings.

Such a contradiction means either that Schopenhauer's notion of consciousness is not good for interpreting *The Matrix* or that the intuitive notion of consciousness we share with *The Matrix* is problematic. We attribute consciousness to machines because they behave like living beings. Even if we know they are

²² Schopenhauer, *On the freedom of will*, § 2.

²³ WI, § 18–19.

not, we think of them as organisms, in the same way we say “he/she” when speaking of our personal computers.

Regarding the question of machines’ consciousness, Schopenhauer’s views do not correspond to the film’s, but his ideas compel us to discuss our own intuitions about consciousness.

9. Why the machines do not win the war?

Now we doubt that machines are really conscious of themselves. But we have no doubt that they are smarter than humans. If so, why have they not already completely defeated humans? How is it possible that the weaker rebels can hold the machines in check? How can Neo destroy Mr. Smith? It is counterintuitive.

On this point Schopenhauer’s philosophy gives an interesting explanation. Being more intelligent, according to Schopenhauer, does not mean being more perfect or invincible. On the contrary, the growing of rational intelligence, which is responsible for concepts formation, is at the same time loss of infallibility pertaining to inferior forms. Instinct and natural laws are examples of such infallibility. A falling stone never fails: its velocity is definitely determined by gravity. Nor fails a spider, when its instinct guides it in designing the web. Errors enter into the world together with reason and its capacity of conceptualization: intelligence obscures instinct, and concepts are never true.²⁴ As a consequence, machines commit even more errors than humans, as is clear from the substantial imperfections in the functioning of the Matrix: in the sequel of the film (*The Matrix Reloaded*, 2003), the character called the Architect tells Neo that the entire history of the fight between machines and humans is deeply dependent on the many errors of the machines in building the Matrix; examples of such errors are the Oracle, Merovingian, and Mr. Smith’s rebellion against the machines. Such wrongdoing may seem strange, but from Schopenhauer’s point of view it is necessary: from higher intelligence derives more imperfection. As Morpheus teaches Neo, machines are not invincible because they commit errors and humans can take advantage of them.

10. Some thoughts as conclusions

Approaching *The Matrix* with ideas and suggestions from Schopenhauer’s metaphysics is intriguing and illuminating. Many peculiar themes of the film – the question of the real world, the machines’ behaviour, the many bugs in the matrix – fit the philosophical frame designed by Schopenhauer.

²⁴ WI, § 27.

This is interesting, but what we expect from philosophical analysis of a film is not its interpretation and explanation. Such a goal would raise spectators' curiosity like a sort of game. Philosophy can be a pastime, but not for one restricted group of people. Philosophy's ambition is that of raising questions, and such questions could even be part of a game, but they also ought to be universal.

Schopenhauer's way to *The Matrix* is not limited to solving the spectators' doubts or to offering an intriguing point of view. It poses questions like: is an intelligent machine a living being? Why do we recognize consciousness in the machines of *The Matrix*? What is consciousness? Is it possible to have consciousness in an inorganic body? Such questions are not only interesting in the film: they are problems of great topical interest, widely and deeply inquired by contemporary philosophy and neurosciences.

Great philosophers are never out-of-date, and Schopenhauer's thought offers important hints to contemporary views. Another topic that is derived from Schopenhauer's most popular argument, namely pessimism, is worth mentioning. Briefly and simply defined, pessimism implies lack of hope in considering the world and its future. Having this in mind, a philosophical approach to *The Matrix* raises new questions: when (if) we build more and more intelligent machines, will it be a contribution to progress or the premise to our destruction? Must we fear machines? Can technology dominate us and make us its slaves?

It is not difficult to guess what Schopenhauer's pessimistic answer could have been. Following Schopenhauer, during the last hundred years other thinkers have answered affirmatively, even not supposing the building of intelligent machines. They have inferred that since we have become dependent on machines we have become slaves²⁵. If machines are intelligent (and more intelligent than humans), danger is even greater. How is it possible that more intelligent beings (machines) accept to serve their "creators" (humans)? Must we stop projects of machines? Could we allow only the construction of machines whose intelligence is strictly directed by laws that assume the primacy of human beings?²⁶

These are philosophical questions that take us very far from *The Matrix*, but like the previous ones they show us just how stimulating going to the movies with Schopenhauer can be.

²⁵ In *The Matrix Reloaded* this theme is briefly discussed during the dialogue between Neo and Councillor Hamann in Zion. Among the philosophers who theorized the domination of technic, it is worth mentioning Martin Heidegger and Günther Anders.

²⁶ An example comes from Asimov's laws of robotics (1940): 1) A robot may not injure a human being, or, through inaction, allow a human being to come to harm; 2) A robot must obey orders given it by human beings, except where such orders would conflict with the first law; 3) A robot must protect its own existence as long as such protection does not conflict with the first or second law.