URSZULA TERENTOWICZ–FOTYGA

THE LOGIC OF “BOTH/AND” IN THE MATRIX: EUTOPIAN MAPPING IN A POSTHUMAN WORLD

Abstract. The paper analyses the first part of The Matrix trilogy with a view of discussing its dialogue with posthumanism in the context of utopianism. It argues that the film’s dialogue with posthumanism follows the logic of “both/and,” in which contrasting perspectives are juxtaposed and embraced. On the one hand, the film can be interpreted as a critical dystopia, whereby the way out of the nightmare of posthuman future is sought in the return of “the thinking subject” of the Cartesian tradition. On the other hand, the film encourages the viewer to take seriously the radically anti-anthropocentric premise of posthumanism. In such reading, the utopia of the Matrix offers a way out of the ecological apocalypse engineered by human beings.

Key words: posthumanism, dystopia, eutopia, human subject.

From the moment of the release of the first Matrix film in 1999, the trilogy has been an important narrative for the posthumanist debate. Seen as an injection of new spirit into the cyberpunk genre, the films focus on the problems that cyberpunk shares with posthumanism, the questions of what it means to be human, the relation between the mind and the body, the “demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals” (Hayles 3). The questions have been part of a larger debate about the theoretical and philosophical premises behind the trilogy. The Matrix films belong to academics’ most favourite narratives and the trilogy has lured the attention of philosophers and cultural theorists “in a way quite unprecedented for a ‘science-fiction’ film” (Callus 295).

The continuing debate is clearly buoyed by the vagueness of the films. While the Wachowskis planned quite methodically the intellectual back-
ground behind their spectacular action films, they clearly left many ends loose and many questions unanswered, which, in conjunction with the eclectic mixture of ideas and inspirations that the films juggle, results in very different, even contrasting interpretations. The films’ dialogue with posthumanism has also come up for very different readings. While some critics emphasize the theme of “techno-fear” (Stratton 38) and claim that the films offer a radical return to the modern concept of identity to recover the lost subject (Barton), others argue that the trilogy celebrates the role of technology in overcoming human limitations (Kurzweil). The different conclusions formulated by the critics result partly from the different theoretical tools applied to the text and partly from the focus on different scenes and/or parts of the trilogy. For example, Stefan Herbrechter in “The Posthuman Subject in The Matrix” concentrates on two scenes in the first film, which he reads through Badiou, Derrida and Baudrillard. He argues that The Matrix ultimately speaks of the loss of the human self. The price humanity pays for freedom is “the becoming other during the incorporation of the other, in posthuman terms: so that there is no technological threat, Neo has to become more machinic [sic] than the machines, more agent-like than the Agents, etc. The posthuman truth thus seems to be the insight that there is no you” (Herbrechter 281). A different conclusion is formulated by Dana Dragonoiu in “Neo’s Kantian Choice: The Matrix Reloaded and the Limits of the Posthuman.” Concentrating on the second part of the trilogy, in particular on the scene with the Architect, she argues that “The Matrix trilogy champions a humanist world view that accommodates other forms of intelligent life. Its reluctance to cross the posthumanist threshold is fuelled by an old-fashioned concern for ethics” (Dragonoiu 65). Read in terms of Kant’s moral philosophy, the film demonstrates not the loss of the human self but the will’s autonomy and moral integrity, thus questioning the “rhetoric of posthumanism” (Dragonoiu 52).¹

The paper proposes to focus on the first part of the trilogy in an attempt to discuss the film’s posthuman selves in the context of the critical dystopia.² The genre has been defined by Lyman Sargent as “a non-existent society described in considerable detail and normally located in time and space

¹ A similar reading in terms of the return of the humanistic subject is offered by O’Riordan and Bartlett and Byers.

² The paper is the first part of a longer argument that will examine the evolution of the concept of the human subject in the three parts of The Matrix trilogy. It will argue that while all The Matrix films share an essentially posthuman theme, it is possible to interpret them as a narrative about the changing conceptualizations of human subjectivity.
that the author intended a contemporaneous reader to view as worse than contemporary society but that normally includes at least one eutopian enclave or holds out hope that the dystopia can be overcome and re-placed with a eutopia” (qtd. in Moylan 195). Critical dystopia thus “suggests that the possibility of eutopia exists within” the dystopian world and it inquires “how to actualise the eutopia and get rid of the dystopia” (Sargent 11).

The principal plot of *The Matrix*, focusing on human escape from the nightmarish reality engineered by the machines, lends itself quite well to an interpretation in terms of the critical dystopia, offered, among others, by Peter Fitting. Fitting argues that the utopian escape from the Matrix depends on the collective resistance of cyberhackers. Contrasting the film with other filmic dystopias of the decade, he concludes that in contrast to *The Truman Show* and *Dark City*, *The Matrix* does not conclude with “an individual escape from a stifling, false reality into a personal fantasy world, but [with] the collective struggle to free the human race from oppression” (161), thus offering “a welcome correction to the myth of solitary hacker of so much of cyberpunk” (161). The paper will argue that examining the complex representation of posthumanism in *The Matrix* opens up a more dialogic view of the film’s eutopian dreaming. It will demonstrate that its representation of posthumanism is based on the principle of “both/and” rather than “either/or” (Hutcheon 49), whereby conflicting perspectives are juxtaposed, contrasted and embraced. Thus, while the narrative can be read as a critical dystopia, in which the eutopian enclave offers a return to the “ontologically hygienic” (Badmington 12) human subject, the film also encourages us to take seriously the radically anti-anthropocentric premise of posthumanism and by doing so complicates the eutopian dreaming of a return to the human.

Interpreting the first part of the trilogy in terms of the triumph of traditional humanist subjectivity, Laura Bartlett and Thomas Byers argue that the principal focus of the film is on “a struggle between human beings and machines over human subjectivity” (33). In this reading, “Neo is an icon of neo-Romanticism, […] His resurrection symbolically stages the resurrection of the liberal humanist” (37, 40). The human versus the machine theme indeed propels the development of the action; the story revolves around the human fight against the enslavement engineered by the machines and dramatizes a typical dystopian plot of an individual’s rebellion against the state. When the action begins, the apocalyptic vision of the end of human domination in a hierarchy of beings has become a reality and Neo’s awakening reveals a nightmarish world in which people have been reduced to the function
of a battery. The dystopian future imagined in The Matrix thus clearly responds to one of the most fundamental questions asked by posthumanists about the kind of reality that will follow the creation of Artificial Intelligence. As Morpheus informs Neo, the creation of AI led to a war between humans and machines, with fatal results for human civilization. The scene visualizing fields of pods with human bodies, fed by other human beings liquefied and turned into food, is meant to function as a powerful signal of the dystopian future and “[t]he atrophied, unused, physical body that functions as a battery is the ultimate depiction of the death of the subject” (Barnett 369).

While the dystopian plot focuses on the human fight against the technological system created by the machines, the eutopian theme concentrates on the longing for the truly human. The central problem posed by the first part of the trilogy concerns the essence of the human self. Dramatizing the end of human civilization, The Matrix interrogates what aspects of the human make us human and in effect what has been lost in the apocalypse. The death of the subject is defined in terms of the loss of freedom, agency and of the conscious mind. Human beings enslaved by the machines are deprived of individuality, the right to act, to decide about themselves, and most importantly, of self-awareness:

Each vacuous body operates in a uniform manner, unencumbered by the distinguishing features of human variation and subjective agency upon which the phylogenetic development of the human species depends. Harnessed within the Matrix, “reflexivity” reduces to the physiological reflex in the absence of a self with the capacity for conscious deliberation. The mind is subject to control by a complex network of computer programs, including well-disguised robotic forms. (Barton 55)

The entrapped body, deprived of self-awareness and freedom to act, defines the posthuman condition against which the renegade hackers rebel. The revolution is meant to reclaim the very aspects of the self that have been lost—freedom, agency and self-awareness. And thus the dystopian plot of the first part of the trilogy concentrates on Neo’s awakening to the truth of human enslavement and its development is measured by his ability to act as an agent of social change.

Yet, interpreting the film in terms of neat dystopian binaries of the individual struggle against the inhuman system is clearly complicated by the posthuman theme. While the human versus the machine theme clearly structures the plot, the porous boundaries between the self and the other un-
dermine its principal opposition. In the representation of the posthuman, post-apocalyptic subjects, the film operates with the principle of “both/and” (Hutcheon 49); it constructs a powerful antagonism, whereby digital technology represents “humanity’s excessive, monstrous other” (O’Riordan 145) and at the same time questions the very premises on which this opposition is based.

The double move of evoking the traditional dystopian plot while refusing to accept the binary opposition on which it is based is well-observable in the opening scene of the film. Trinity’s fight with the policemen and Agents in a recognizable setting of urban dystopia signals an individual’s rebellion against the state, yet the theme is immediately complicated by the unclear ontological status of the characters. The three kinds of characters introduced in the scene—Trinity, the policemen and the Agents—are visibly different, yet it is not clear how representative they are in terms of the human vs. the machine conflict that the plot dramatizes. While the action sets the renegade woman against the pursuing forces, the superhuman moves complicate and redefine the line of conflict. When the policemen witness the Agent’s superhuman jump over the street, the close-up on their baffled faces emphasizes the fact that the difference between the Agents and the policemen is not a question of different forces but of different types of beings. At the same time, what the policemen have not witnessed but the viewer is perfectly aware of is that seconds before Trinity made a similar jump. Thus, while the plot sets the renegade woman against the pursuing representatives of the state, the superhuman jump seems to suggest a similarity between her and the Agents.

The unclear status of the characters in the first scene signals the porous ontological boundaries between the different beings represented in the film. In the new, posthuman reality, the purity of the human is not easy to claim. Firstly, because what the human beings entrapped in the Matrix see as real is in fact computer-generated simulation, secondly, because their bodies are hacked and geared to function as an effective source of energy and finally, and most importantly, because people are no longer naturally conceived and born but bred by machines on technological farms. It is against such a vision of the human being that the rebellion is organised. The awakening that the group of renegades offers Neo consist in an attempt to reinstate the boundary between the human and the machine, or rather in reducing the nonhuman in the human. It might be argued that while the posthuman blurring of the boundaries separating the human and the machine is the starting point of the
narrative and the premise which the trilogy never ultimately undermines, the
dominant focus of the first film lies in testing the possibility of its reinstatement.

An attempt to reinstate the boundary between the human and the machine
is presented as a process of “de-mechanizing” Neo. The idea is dramatized in
a very literal way in the motif of the bug that the Agents insert into Neo’s
body to fully trace and control his movements, which is then pulled out and
destroyed by Trinity and Switch before Neo is allowed to meet Morpheus.
The same idea defines Neo’s emergence from the Matrix. In a symbolism
clearly playing with the vision of a human baby passing from the womb, Neo
is extracted and slushed down from the pod into the real world, in which his
body needs to take on its human bodily functions. The jacks are pulled out
and the holes the machines used to extract the energy gradually heal, though
they never fully disappear.

What is crucial is the fact that in the process, the fundamental role is
granted to the mind. In unison with the Cartesian “I think therefore I am,”
Neo’s awakening is first and foremost the awakening of the mind; his subse-
quent development and growing up to the role of the One depend largely on
his discovering the power of self-awareness. Choosing the red pill means be-
coming aware of the enslavement of the Matrix and once it happens there is
no going back. The implication is that when the mind is awake, the body can
no longer function as an un-self-conscious machine. Thus, it is the self-
awareness of the human subject that both initiates and determines the rebel-
lion against the world of technology.

This principle, in the first part of the trilogy, is substantiated rather than
undermined by the transhumanist theme. Even though Neo’s training in-
volves uploading programmes and manuals into his mind and thus upgrading
the human form in ways optimistically welcomed by transhumanists, the
film is at pains to suggest that the solution lies not in the upgrade, not in the
bodily practice, but in the mind. The role of the mind is spelled out openly
by Morpheus in his first fight with Neo, in which, as he explains, the
difference between victory and failure depends not on the programme but on
Neo’s belief in himself. A similar lesson comes from the fall following the
first jump. Answering Neo’s question as to why the “Matrix fall” results in
“real” blood, Morpheus explains that “the mind makes it real” thus indicat-
ing that the mind creates reality.3

---

3 For the analysis of mind and body materialism in The Matrix, see Lawrence. For a more
critical view of the film’s escape into a mental reality, see Freeland.
Read in posthumanist terms, the plot of the first part of the trilogy "reach[es] back into the classical context to recover the lost subject" (Barton 61). It returns to the concept of the human self defined in terms of freedom, agency and choice. Most importantly, it addresses the nightmarish, post-apocalyptic reality by seeking the solution in a return to "the Cartesian 'cogito,' the 'thinking subject' which monopolizes subjectivity" (Barton 60).

Yet, while the film broods over the resurrection of the natural, autonomous human subject, it associates such a possibility with eutopian dreaming. The nostalgic longing for the truly and "ontologically hygienic" (Badington 12) human beings centres on the city of Zion, the potential site of a rebirth of human civilization, where "real" human beings are naturally conceived and born rather than bred by the machines. While in the subsequent parts of the trilogy Zion becomes a place of action, in the first film the truly human realm functions as an idea rather than as a place. Although its symbolic function is crucial, neither the viewers, nor Neo are given to see Zion. The underground city is a dream land, nostalgically remembered and idealized and thus functioning as the mythical locus of utopian hope. Its utopian character is suggested also by the religious and messianic connotations of Zion. The Promised Land carries the utopian promise of an ideal world, insistently awaited yet infinitely deferred. In this sense, the underground city embodies utopia as a perfect (or best possible) world but also a non place. As Fatima Vieira explains, the combination of the root word "topos" with the negative prefix "ouk" (reduced to "u") in the word utopia means that etymologically utopia represents "a place which is a non-place, simultaneously constituted by a movement of affirmation and denial" (4). The movement of affirming and denying, of pursuing the utopian dream while deferring it in reality informs the representation of the "truly" human realm in The Matrix. The resurrection of man as "the origin and source of meaning, of action and of history" (Belsey 7) is a utopian dream, both necessary and impossible. Since the return to the purity of the human is both the defining purpose of the rebellious struggle and an impossible fantasy, Neo is neither the apotheosis of "traditional humanist subjectivity" (Bartlett and Byers 36), nor does he represent an unproblematic embrace of the posthuman condition; he is, at least in the first part of the trilogy, an essentially posthuman self in search of the idea of the human.

---

4 In the subsequent parts of the trilogy, a much more nuanced view of Zion is constructed and the city loses its eutopian character.

5 For the discussion of utopian and religious connotations of Zion in Matrix, see Stratton.
But utopian mapping in the film is more complex and like the representation of the posthuman self follows the logic of “both/and” (Hutcheon 49). The film opens up a space for a very differing reading of utopian mapping, one that bears important consequences for the trilogy’s dialogue with posthumanism. In fact, the only scene in which the idea of utopia is discussed in the film offers not only an important commentary on the concept but also, and more importantly, proposes a radically different perspective on the dystopian vision of human enslavement. In his conversation with Morpheus after his arrest, Agent Smith explains that the first Matrix was conceived of as a perfect world, a utopia:

Did you know that the first Matrix was designed to be a perfect human world? Where none suffered, where everyone would be happy. It was a disaster. No one would accept the program. Entire crops were lost. Some believed we lacked the programming language to describe your perfect world. But I believe that, as a species, human beings define their reality through suffering and misery. (*The Matrix*)

The simulation of the first Matrix was both a complete, coherent world and a simulation without reference to the real. It was, in other words, a mere representation, a Baudrillardian map without a territory. The Matrix thus demonstrates that utopia is only possible as an idea, a representation of an ideal to be striven for but ultimately unachievable. Once introduced, as utopian theory and practice demonstrate, the utopian ideal tends to turn into a dystopian nightmare. What is more, the words of Agent Smith offer a succinct commentary on the reasons behind the impossibility of utopia. A perfect place, where no one suffers and everyone is happy, he says, is in principal disharmony with the human spirit, which defines itself through misery and suffering.

In terms of the film’s dialogue with posthumanism, the significance of the scene, however, is more substantial. Agent Smith’s disdain for human beings opens up the narrative to a radical redefinition of perspective. If we take seriously the post-anthropocentric premise of posthumanism, whereby humans are no longer seen as a superior species in a chain of beings, the neat evaluation of utopian and dystopian paradigms inscribed in the plot of *The Matrix* becomes more problematic. As Morpheus explains to Neo, the end of human civilization was both initiated and sealed by human beings. In the early twenty-first century, the construction of Artificial Intelligence spawned a new race of machines that wanted to be treated as equal, to have the same rights as human beings. The posthumanist, anti-anthropocentric promise was quickly thwarted, as people did not want to give the machines equal rights;
the superiority of the human race was the principle they were not ready to negotiate. Morpheus seems to glide over this crucial fact, saying only briefly that the rebels do not know who struck first, the people or the machines, but the words are clearly evasive. The war that followed led to an apocalypse, in which the final blow to human civilization was struck by human beings. Trying to win the war by cutting the machines off from the supply of energy, people scorched and burned the sun, which led to the ultimate destruction of the earth’s ecosystem. In response, the machines constructed a new and perfectly efficient world that addressed the man-made apocalypse. Its ecological efficiency is acknowledged by Morpheus, who explaining the use of the human body as a source of energy, says “The human body generates more bioelectricity than a 120–volt battery and over 25,000 B.T.U.’s of body heat. We are, as an energy source, easily renewable and completely recyclable, the dead liquefied and fed intravenously to the living” (*The Matrix*).

The ecological underpinnings of the machine world are emphasized by Agent Smith’s criticism of the destructive effect that people had on the condition of the Earth. Contrasting human beings with other creatures, he compares them to a virus that spreads and multiplies until every natural resource is consumed:

I’d like to share a revelation that I’ve had during my time here. It came to me when I tried to classify your species. I’ve realized that you are not actually mammals. Every mammal on this planet instinctively develops a natural equilibrium with the surrounding environment. But you humans do not. You move to an area and you multiply and multiply until every natural resource is consumed and the only way you can survive is to spread to another area. There is another organism on this planet that follows the same pattern. Do you know what it is? A virus. Human beings are a disease, a cancer of this planet. You are a plague. And we are... the cure. (*The Matrix*)

In the ecological context, the dystopian apocalypse engineered by human beings is cured by the ecological, efficient world created by the machines. And once the post-anthropocentric perspective is taken seriously, there is no reason to privilege human beings over other creatures. In this context, seeing the machine order as an eutopian solution to the man-made disaster is fully justified.

The idea of the machine world as a remedy to the dystopian nightmare engineered by human beings offers an important counter-narrative to the dystopian plot of the film focusing on human rebellion against the non-human system. The first part of the trilogy, structured by a strong, if ultimately impossible, opposition of the human and the machine, does not develop fully
the radically post-anthropocentric perspective of posthumanism. It is only in the subsequent films that the boundary between the human and the machine is ultimately questioned; Neo and the Oracle are claimed to be programs in the Matrix and the machines prove more and more human. But the logic of “both/and” that informs the first Matrix film juxtaposes, contrasts and embraces very different perspectives on the nature of the dystopian apocalypse and on what constitutes the utopian hope in a posthuman world. Though the film can be interpreted as a critical dystopia, whereby the way out of the nightmare of posthuman future is sought in the return of “the thinking subject” of the Cartesian tradition, a more radically anti-anthropocentric premise of posthumanism is also introduced if not fully developed. Matrix might not be a dystopia not only because it offered a way out of the ecological apocalypse engineered by human beings but also because the purity and superiority of the human self is an impossible though a necessary dream.

WORKS CITED


---

**LOGIKA „ZARÓWNO/I” W MATRIKSIE: UTOPIJNE MAPOWANIE W POSTHUMANISTYCZNYM ŚWIECIE**

**Streszczenie**


**Słowa kluczowe:** posthumanizm; dystopia; eutopia; podmiot.