

EMBODIED ARTIFICIAL INTELLIGENCE IN SCIENCE FICTION

PHILOSOPHICAL PRESUPPOSITIONS AND IMPLICATIONS

INTELIGÊNCIA ARTIFICIAL ENCARNADA NA FICÇÃO CIENTÍFICA

Pressupostos filosóficos e implicações

LA INTELIGENCIA ARTIFICIAL ENCARNADA EN LA CIENCIA FICCIÓN

Presupuestos e implicaciones filosóficas

Andrea Pace Giannotta

(University of Bergamo)

andreapacegiannotta@gmail.com

Recibido: 27/03/2022

Aprobado: 20/07/2022

ABSTRACT

In this paper, I explore the fruitful relationship between science fiction and philosophy regarding the topic of artificial intelligence. I establish a connection between certain paradigms in the philosophy of mind and consciousness and the imagination of possible future scenarios in sci-fi, especially focusing on the different ways of conceiving the role of corporeality in constituting consciousness and cognition. Then, I establish a parallelism between these different conceptions of corporeality in the philosophy of mind and certain representations of AI in sci-fi: from computers to robots and androids. I conclude by stressing the value of exchanging ideas between sci-fi and philosophy to foreshadow and evaluate some scenarios of high ethical relevance.

Keywords: embodiment. consciousness, sci-fi. phenomenology. flesh.

RESUMO

Neste artigo, exploro a relação frutífera entre ficção científica e filosofia no que diz respeito ao tema da inteligência artificial. Estabeleço uma ligação entre certos paradigmas da filosofia da mente e da consciência e a imaginação de possíveis cenários futuros na ficção científica, focando especialmente as diferentes formas de conceber o papel da corporeidade na constituição da consciência e da cognição. Estabeleço assim um paralelismo entre estas diferentes concepções de corporeidade na filosofia da mente e certas representações de IA na ficção científica: desde computadores a robôs e andróides. Concluo salientando o valor de uma troca de ideias entre ficção científica e filosofia para prenunciar e avaliar alguns cenários de elevada relevância ética.

Palavras-chave: corpo encarnado. consciência. ficção científica. fenomenologia. carne.

RESUMEN

En este artículo exploro la fructífera relación entre la ciencia ficción y la filosofía en relación con el tema de la inteligencia artificial. Establezco una conexión entre ciertos paradigmas de la filosofía de la mente y la conciencia y la imaginación de futuros escenarios posibles en la ciencia ficción, centrándome especialmente en las diferentes formas de concebir el papel de la corporeidad en la constitución de la conciencia y la cognición. Así, establezco un paralelismo entre estas diferentes concepciones de la corporeidad en la filosofía de la mente y ciertas representaciones de la IA en la ciencia ficción: desde los ordenadores hasta los robots y androides. Concluyo destacando el valor de un intercambio de ideas entre la ciencia ficción y la filosofía para prefigurar y evaluar algunos escenarios de gran relevancia ética.

Palabras clave: sujeto encarnado. conciencia. ciencia ficción. fenomenología. carne.

Introduction

Science fiction (sci-fi) is the literary and cinematographic genre that constructs narratives based on possible future scenarios that involve revolutionary scientific discoveries and technologies. These envisioned scenarios very often have a deep philosophical import. This is most clear in the case of certain philosophical ideas about mind, consciousness, and the possibility (or impossibility) of artificial intelligence (AI). In these domains of philosophical inquiry, the connection with sci-fi is tight and goes in both directions: philosophical ideas can inspire sci-fi narratives, and the latter not only can give concreteness to abstract philosophical ideas but can also open new possibilities for philosophical reflection.

Indeed, the concept of AI is essentially a sci-fi concept, concerning a possible techno-scientific innovation leading in the future to the creation of an artificial minded creature. At the same time, this concept is deeply philosophical, as it deals with nothing less than the most fundamental philosophical questions about mind, consciousness, and their place in the order of reality. The fact that we are beginning to turn the sci-fi concept of AI into something real makes it even more meaningful to reflect on the philosophical presuppositions and implications of certain sci-fi scenarios. In particular, the guiding thread of this paper is the role of corporeality in consciousness and cognition and the way it is conceived in philosophy of mind, on the one hand, and the way it comes into play in sci-fi, on the other.

The paper is structured in four sections. In the first three sections, I shall discuss three paradigms in philosophy of mind: disembodiment (sec. 1), weak embodiment (sec. 2), and strong embodiment (sec. 3). In all three cases, I shall establish a connection with certain representations of AIs in sci-fi: from computers to robots and androids. In the fourth section, I shall consider some ethical issues that arise within the contemporary debate on artificial intelligence and artificial consciousness. I conclude by stressing the value of an exchange of ideas between sci-fi and philosophy to foreshadow and evaluate some scenarios of high ethical relevance.

1. Disembodiment

The concept of artificial intelligence implies, by definition, the creation of an artefact by human beings. In any case – i.e., regardless of the specific way in which this is done – such an artefact must have some *material realisation*, i.e., it must be created by arranging a material substratum in certain ways. This means that AI is always *embodied* if we use this concept in a very general sense (i.e., as equivalent to “materially realized”). However, the crucial issue is: what kind of embodiment is necessary to have a mind and, therefore, also an artificial mind?

A first answer to this question is found in the classic computationalist-representationalist (or logicist-symbolic) paradigm in philosophy of mind, which conceives of the mind in analogy to the software in a computer. In this paradigm, mental states are identified with certain functions that lead from input from the environment to a certain behavioural output. The logicist-symbolic approach to AI has been also

notoriously called “good old fashioned artificial intelligence” (GOF AI, Haugeland, 1985) and seeks to build agents that show “intelligent” behaviour. In this sense, “intelligent” is an adjective attributed by us, the observers, to systems that perform in certain ways. It is therefore an “apparent” feature of cognitive systems that appears “from the outside” or “in the third person”. For instance, an AI that passes the imitation game (Turing, 1950) has a certain behaviour that is indistinguishable from that of a human being.

However, in this way of conceiving of intelligence, there is no reference to the first-person experience of a cognitive agent, i.e., there is no reference to “what it is like” (Nagel, 1974) to be a subject of experience. This is the dimension of “phenomenal consciousness”, which constitutes the so-called *hard problem* in philosophy of mind (Chalmers, 1996). I stress the fact that the computationalist-representational paradigm in philosophy of mind and logicist AI leaves aside the phenomenal aspect of mental life, which is essentially tied (as we will see) to bodily life. In this approach, phenomenal consciousness doesn’t play any role in accounting for the behaviour of an agent.

The possibility of realizing this form of intelligence in artefacts is what is commonly called, following Searle (1980), “weak AI”, as opposed to “strong AI”. Whereas strong AI “seeks to create artificial persons: machines that have all the mental powers we have, including phenomenal consciousness”, weak AI “seeks to build information-processing machines that *appear* to have the full mental repertoire of human persons” (Bringsjord & Govindarajulu, 2020: 66).¹ Therefore, the realization of weak AI concerns intelligence defined in merely functional terms as a certain behaviour in relation to certain conditions. In particular, this view maintains a certain dualism between mind and body, which is evident in the analogy between the mind and the software of a computer and between the brain and its hardware. Indeed, even if the software must be implemented by a certain material substratum, the features of the hardware that are deemed to be relevant for realizing a mind are situated at a general level of abstraction (e.g., the architecture of the Turing machine and the Von Neumann architecture). They do not concern the specific morphologic, dynamic, and sensorial features of the body of a concrete cognitive agent. Therefore, the concrete body of an agent is not deemed to be relevant in accounting for its mental processes. For this reason, this view has been criticized for implying a conception of the mind as *disembodied*.

The computationalist-representational paradigm has proven powerful in leading to the creation of technologies that simulate and, in some cases, outperform certain intellectual abilities of human beings (e.g., calculation). However, it has shown its limitations in the attempt to create artificial agents that emulate basic behaviours such as the ability to freely move in space, to converse on the fly, to learn by reading or hearing a text, etc. (Bringsjord & Govindarajulu, 2020: 23 ff).²

1.1. Disembodied AI in sci-fi

The concept of a merely functional, disembodied mind comes into play in various sci-fi narratives when the depicted AI is that of a program or operating system in a digital computer.

Maybe, the most iconic example is HAL9000: the AI character and main antagonist in Arthur Clarke’s *Space Odyssey* series, on which Kubrick’s *2001: A Space Odyssey* (1968) is based. HAL9000 is an AI that controls the Discovery One spacecraft, interacting with its crew and thus showing abilities such as speech, speech recognition, facial recognition, natural language processing, interpretation of emotional behaviours, etc. All these abilities make it very close to the idea of a general AI (i.e., an AI that can perform any intellectual task that can be performed by a human being). However, apart from the camera lens through which HAL9000 “sees” the environment, the speakers by which it communicates with the crew, and its hardware (which we see towards the end of Kubrick’s movie), HAL9000 is a paradigmatic

¹ It should be noted that Searle conceives of “weak AI” as a powerful tool for the study of the mind through the simulation of mental processes in a computer.

² Regarding the discussion of artificial agency within the computational-representational paradigm see Howard & Muntean, 2017; Johnson & Miller, 2008.

example of disembodied AI. Here we can ask: is HAL9000 a conscious mind? Or is it just a good simulation of a conscious mind? This is an open question in the sci-fi narrative, which can play with the uncertainty concerning the answer to this deep philosophical question.

A more recent depiction of a disembodied AI in sci-fi is Samantha: the virtual assistant of Theodore's computer operating system in Spike Jonze's *Her* (2013). Samantha is an advanced AI that can learn and evolve through the interaction with Theodore, up to the point that the two enter into a romantic relationship. However, apart from being personified through a beautiful female voice, Samantha is radically disembodied, being even independent of the specific hardware of Theodore's computer. Indeed, at the end of the movie, Samantha joins other AIs in an evolutionary leap that frees them from dependence on any material support.

Also in *The Matrix* franchise (1999-2021) we see at play the idea of a disembodied mind. This happens in two ways. Firstly, in the matrix itself, which is a virtual reality simulator built by self-aware machines that farm humans as power sources. Secondly, in the idea that human beings could have experience in a computer simulation indistinguishable from real-world experience. Indeed, the humans trapped in the matrix have a full bodily experience without their real bodies being involved at all, apart from their brains being physically stimulated in the appropriate way. This scenario represents the "internalist" idea that experience emerges from the activity of the brain alone, as in Putnam's notorious "brain in a vat" thought experiment (Putnam, 1981).

The idea of the mind as a disembodied software is also at the basis of the concept of "mind uploading": the possibility, advocated by various transhumanists and extropians, that one's mind could be transferred into a different material substrate, thus freeing human beings from the dependence on the biologic and therefore highly vulnerable and mortal body. Such a powerful idea still presupposes a concept of the mind as essentially disembodied (i.e., essentially independent from the concrete material and structural features of the body) and it comes into play in suggestive sci-fi scenarios. A recent dystopian depiction of this idea is found, for instance, in the episode "White Christmas" of the sci-fi anthology series *Black Mirror* (2011-in production). In this episode, new technology allows people to create a digital clone of themselves ("cookie") that is stored into an egg-shaped object, to be mainly used as a personal assistant. The digital clone "lives", first and foremost, as a purely disembodied consciousness in a void space, to which the simulation of some material objects in a white room and the simulation of a body are added. However, neither the environment nor the body is essential for the existence of the clone. Another example is the *Black Mirror* episode *San Junipero*, which depicts a future scenario in which elderly people who are close to death can decide to upload their mind in a sort of "virtual paradise", in which they can potentially "live" forever. Also in this case, the fictional scenario represents the possibility that one could have a full form of existence as a disembodied mind living in a computer simulation. This possibility is, again, a fertile narrative device in sci-fi that is, at the same time, very problematic from the philosophical point of view.

2. Weak embodiment

In contrast to the computational-representational paradigm, in the last three decades, the so-called *new embodied cognitive science* has emerged. This approach is centred on the investigation of the bodily roots of cognitive processes. Some significant stages in the path that led to the affirmation of this new approach are George Lakoff and Mark Jonson's research on cognitive semantics (Lakoff & Johnson, 1980), Rodney Brooks' robotics (Brooks, 1991), Francisco Varela, Evan Thompson and Eleanor Rosch's enactive approach (Varela et al., 1991), and the more recent forms of "sensorimotor" (Noë & O'Regan, 2002; O'Regan & Noë, 2001) and "radical" (Hutto & Myin, 2012, 2017) enactivism. In contrast to the logicist-symbolic approach, all these theories highlight the essential role of the specific features of the body of a cognitive agent in constituting its mental processes, in contrast to the classic Cartesian mind-body dualism and its revival, in new forms, within classical cognitive science (with the above-seen software-hardware dualism).

However, there is a certain ambiguity in the concept of body that comes into play in the debate on embodied cognition. Indeed, we must ask: what is the body? The answer is not obvious: one can develop different theories of corporeality, thus developing different versions of the *embodiment* thesis.

I will now distinguish between a *weak embodiment thesis* – which conceives of the mind as grounded in the merely functional body – and a *strong embodiment thesis* – which conceives of the mind as grounded in the living-lived body, conceived of in turn as a *functional and sentient* body. I connect this distinction with Husserl's distinction between the objective body (*Körper*), i.e., the body that appears “in the third person” as moving and interacting with the environment, and the living and lived body (*Leib*), i.e., the body that is experienced “in the first person” and that has sensations of pleasure, pain, etc. In the first sense, the body is the object of perception by somebody other than the subject. When it is investigated in this way, one can leave aside the way the body is experienced in the first person, i.e., “what it is like” to have – or better, *to be* – a living and lived body.

Now, I want to stress that at the basis of some theories of embodied cognition there is the notion of the objective-functional body and not that of the lived body. A paradigmatic example is Rodney Brooks' pioneering work on robotics (Brooks, 1991, 2002). In general, the objective of robotics is to create machines that show an “intelligent” behaviour that is typical of animals, including humans (perception, movement and interaction with the environment, language, etc.). Brooks pursues this objective by developing an approach that is alternative to the classical computational-representational paradigm, by building artificial agents that, instead of algorithmically elaborating complex representations of the environment, can directly interact with it. For instance, Brooks refers to a robot that can move efficiently in the environment and that is built by putting together a set of simple mechanisms that correspond to a certain task (e.g., avoiding obstacles, recognizing a can and grasping it, etc.). This can be done without involving any representational state (i.e., any “internal” model of the environment with a semantic content), according to the motto “intelligence without representation” (Brooks, 1991: 156).

However, we must make explicit an ambiguity in the term “intelligence”, as it is used by Brooks. Indeed, Brooks's robots show intelligent behaviour (e.g., avoiding obstacles) *if* we exclude from the concept of intelligence any reference to phenomenal consciousness. In Chalmers's terms, the robot's “mind” is a mere “cognitive mind” (not phenomenal): until proven otherwise, there is no qualitative, felt aspect in the obstacle-avoiding behaviour of the robot (even if it is built to show a pain-avoiding behaviour). Brooks's approach, therefore, is to claim that the mind (intelligence) is essentially embodied, but this is done in the light of a conception of the body as described “in the third person” (i.e., the objective body that moves in the environment avoiding obstacles and interacting with other objects by grasping them, moving them, etc.). Indeed, Brooks's robots interact with the environment in relation to the *stimuli* that come from it, but in this view, the stimuli are mere causal relations between physical entities, without implying the presence of *sensibility* in the agent (i.e., without necessarily implying the presence of a qualitative effect that is associated with the stimulation of the living body).³

For this reason, Brooks' motto 'intelligence without representations' actually implies the idea that it is possible to create 'intelligence without (phenomenal) consciousness'. In Brooks's analyses, the objective body of the robot is not characterised by being the locus of sentience (understood in phenomenal-qualitative terms). This is not a problem for robotics, insofar as it merely attempts to produce artefacts that exhibit certain behaviours, simulating those of an animal (human or non-human) but without claiming to create minds in the full sense (i.e., cognitive and phenomenal minds). Rather, the problem arises if we believe that the realisation of a given behaviour by a robot tells us everything we need to know about the mind and its bodily grounding. More precisely, the form of “embodiment” that follows

³ Brooks refers to the robot's ability to *sense* the environment. However, he defines this ability in merely functional terms (i.e., in terms of the dependence relation “if A, then B”). For instance, *if* the robot “senses” the presence of an object in its visual field, *then* it will move away in order to avoid hitting it. In this analysis nothing is said about the qualitative effect for a human being (and for a sentient being in general) when *feeling* something.

from Brooks' approach to robotics is "weak", insofar as it involves only the notion of cognitive mind - not phenomenal - and objective body - not living and lived body.

The weak form of embodiment also comes into play in the so-called "sensorimotor approach" (Noë & O'Regan, 2002; O'Regan & Noë, 2001), which is focused on the investigation of the sensorimotor interaction between cognitive agent and environment. A paradigmatic example of the "sensorimotor coupling" between agent and environment is that of a missile that can track a target (e.g., an aeroplane): the missile moves in certain directions in relation to the movements of its target, being "coupled" with it. However, this "mastering" of sensorimotor contingencies is a *functional* feature of the cognitive agent's body, which can be realized also without any associated phenomenal effect (or "what-it's-likeness"). Also in this view, therefore, the body is just an objective, functional body and not necessarily a sentient body (see Pace Giannotta 2022a, 2022b).

2.1. Weakly embodied AI in sci-fi

The weak form of embodiment, which involves the merely functional and not sentient body, comes into play also in some sci-fi narratives, when the depicted Ais are no longer simply those of digital computers, but also have a body that allows them to move and interact with the environment.

On this point, a distinction has to be made between *robots* – which can have various forms – and *androids*, which are built to closely resemble human beings and their form of embodiment. In turn, the body of the android can be simply mechanical or also biological (i.e., endowed with an artificial *flesh*). The weak form of embodiment comes into play in those robots and androids that have just a mechanical-functional body, which is very different from the sentient body of animals in flesh and blood. In particular, in these depictions of embodied Ais, the intelligent performance of the robot is made possible not only by its ability to algorithmically process symbols but also by the fact that it has a body that allows it to move and interact with the environment and with other agents. However, what is at stake in these scenarios is the merely mechanical-functional body which, as we shall see, is deemed insufficient for the existence of consciousness by proponents of the strong embodiment thesis.

An iconic example is a metallic automaton in *Metropolis*, the novel by Thea Von Harbou on which Fritz Lang's movie of the same title is based (1927). The android that appears in this work is built in such a way as to show a human-like behaviour and to resemble the character Maria. However, the resemblance between the android and the real person is superficial, since the android has just a metallic and mechanical (not biological) body, which is revealed when it is burnt at the end of the movie.

Another example is Tik Tok: the mechanical man in L. Frank Baum's *Oz* books series (1900), which is one of the first robots in literature. Tik Tok has a round body, is made of copper and runs on clockwork springs that periodically need to be wound and that correspond to the ability to think, act, and speak. Baum clarifies that Tik Tok is not a living being and does not feel any emotion, since its body is mechanical and not biological. In this case, the ambiguity in the depiction of the android – is it conscious? – is resolved in the narrative. In other cases, however, the ambiguity is maintained, when the depicted mechanical robots and androids behave so similarly to humans that they appear to possess consciousness – e.g., R2-D2 and C3-PO: the "droids" in George Lucas's *Star Wars* saga (1977-2019).

As I said before, sci-fi narratives can play with the ambiguity and don't need to give a clear answer to the fundamental philosophical question about robots and androids: are they conscious creatures like us? This is a question that must be addressed by philosophy, with its methods of inquiry. Indeed, art can fruitfully interact with philosophy but it does not need to address philosophical problems. A work of art can tolerate ambiguity in the depiction of an android (is it a person like me? Does she feel something or

is it just a machine?) and this ambiguity can be a fertile narrative device that is also a stimulus for philosophical reflection.⁴

3. Strong embodiment

In contrast to the weak form of embodiment that we have just seen, there is a more radical form of embodiment that comes into play in Husserlian phenomenology and in some recent approaches in the fields of consciousness studies and artificial consciousness.

3.1. Husserl's phenomenology of the body

In Husserl's phenomenology, we find a rich account of corporeality and its essential role in the constitution of subjectivity (Pace Giannotta 2022a, 2022b). In particular, in *Ideas II* we find the above-mentioned distinction between *Körper* – the body as object, investigated “in the third person” (e.g., by anatomy) – and *Leib* – the living and lived body, experienced “in the first person”. The latter, in turn, consists of two aspects: a functional and active dimension (the body that moves and acts in the environment) and a sensorial and passive dimension, i.e., the body that *feels*, being the locus of “phenomenal consciousness” (*sentient body*).

In the phenomenological framework, the functional dimension of the body allows one to “constitute” the objects of perception by moving around them. For instance, observing a table from various points of view and moving around it, one constitutes the object “table” as the correlate of a series of perceptual experiences. In turn, each experience is constituted by an intentional component (*morphè*) and a sensorial component (*hyle*). For example, the intentional animation of a series of chromatic sensations leads to the constitution of the objective colour of the table (Husserl, 1983: 73 ff). The sensorial component essentially pertains to the *sentient* dimension of the body: the body that has sensations. This analysis highlights the essential link between the functional and the sentient dimensions of the body. Indeed, the phenomenological analysis of the *functional body* leads us to acknowledge its being grounded in the *sentient body* (the body that feels pleasure, pain, joy, hunger; that fears, desires, etc. and that has sensations of colour, smell, taste, etc.).

In particular, in *Ideas II* Husserl develops a detailed analysis of various kinds of corporeal sensations, by distinguishing at least five kinds of them: *kinaesthetic* sensations (sensations of movement); *representing* sensations (through which the sensible properties of the perceptual object are constituted: colour, roughness, taste, etc.); *localized sensations of contact* (*Empfindnisse*); the sphere of *sensitive feelings* (pleasure, pain, wellness, etc.); and various sensations “difficult to analyze and discuss [...] that form the material substrate for the life of desire and will, sensations of energetic tension and relaxation, sensations of inner restraint, paralysis, liberation, etc.” (Husserl, 1989: 153).

Among these sensations, tactile sensations and especially localized sensations of contact (*Empfindnisse*) have a special role. Indeed, through them, the living body *feels itself* and is *affected by itself*, being manifest at the same time as a material *object* and as a *subject*, insofar as it is the locus of localized sensations (Bernet, 2013; Zahavi, 2002). In relation to a physical event (e.g., when my hand is touched, pricked or rubbed), in that *moment* and *place*, the *feeling* happens: there are localized sensations. In particular, Husserl analyses the case of the two hands that touch each other: each of the two hands can alternatively assume the active role of *touching* hand – that has sensations of contact relative to the properties of the other hand (smooth, soft, cold, etc.) and localized sensations of contact relative to itself as touching hand – or the passive role of *touched* hand – that has localized sensations relative to the fact

⁴ In stating this, I partly endorse the thesis of the autonomy of art, advocated for instance by Benedetto Croce, according to which the object of art is beauty and not truth (or goodness, or utility). More precisely, in the case of a work of science fiction that, for instance, depicts the possibility of mind-uploading or of a disembodied AI, this may be a work of high aesthetic value and philosophical inspiration, even though one may then believe, on philosophical grounds, that it depicts something impossible (and therefore false).

of being touched by the other hand.⁵ In fact, through the *Empfindnisse* the body reveals itself to be, at the same time, *sensible* and *sentient*. For this reason, Husserl claims the “privilege” of the tactile dimension in the constitution of corporeality because – differently from senses such as sight and touch – tactile sensations reveal both objects and the body itself as the subject of perception (Husserl, 1989: 150). Furthermore, as claimed by Bernet, the experience of the *Empfindnisse* is the primary form of openness of a subject to *alterity* (Bernet, 2013: 53). The *diffusion* of localized sensations makes evident the spatial and manifold nature of the living body, which is made of parts and organs, each of which is sentient and sensible. Starting from this primary experience of the alterity that is already constitutive of the *Leib* – the “non-coincidence of the flesh with itself” (ibidem) – the conscious subject can enter into a relation with the world and the other subjects.

Another essential role in the constitution of corporeal experience is played by *kinaesthetic sensations*, which are relative to the positions and the movement of various parts of the perceiving and agent body. In fact, according to Husserl, the perceiver is always an agent, but it is so because the functional body is at the same time *sentient*, given that each corporeal movement correlates with localized sensations of movement. That is, the functional body that, for instance, moves in a certain direction to grasp an object, does so based on the awareness of its position in space and of the kinaesthetic sensations that are relative to its movements (together with the “representing sensations”, which are relative to objectual properties, and sensations of tension, relaxation, pleasure, pain, etc., which also have an essential role in the constitution of a “cognitive agent”). Kinaesthetic sensations, indeed, *motivate* the course of perception, i.e., the series of experiences through which an object of perception, such as a table, is constituted. Each experience, through which the representing sensations relative to objectual properties of the table are given, is accompanied by kinaesthetic sensations relative to the position and the movement of the eyes, the hands, etc. This means that the functional body is essentially also sentient. There is always a certain “what-it-is-likeness” that is associated with the movement of the functional body.

This analysis, therefore, leads us to bring into question the clear distinction between functional body (which could be replicated by an automaton with no sentience) and sentient body (or phenomenal body), because the functional dimension of the body is also based on sentience, which is, first of all, a *self-sentience* or *self-affection* of the body (Thompson, 2005; Zahavi, 2002). Without this sentient dimension of the body, the perceiver-agent would not be so (or it would be so in a very different way: the way of being of an automaton or a machine).

The idea of the bodily grounding of consciousness also comes into play in the “genetic” development of phenomenology (Husserl, 2001a), which is focused on the investigation of the temporal structure of consciousness. In particular, in its later reflection on this topic, Husserl conceives of the “living present” as the fundamental unit of temporality: a temporal field structured in three parts that are essentially linked to one another: *impression*, *retention*, and *protention*. The classic Husserlian example is that of perceiving a melody, which is a phenomenon with a clear temporal extension. This perception is made possible by the fact that flowing qualitative elements (the sensations of sound) continuously slide into the just-past (the sound just-heard) and are “retained” in consciousness, being joined, at the same time, with the “protention” toward the future course of the melody to come. The conscious *present* is therefore a temporal field with a certain *width* or incompressible *density* (Varela, 1999; Zahavi, 2010). This is what James calls the “specious present” (James, 1890). The temporal structure of consciousness can be therefore conceived in analogy to the visual field, which has a centre (the hyletic core) and a periphery (retention and protention) that are inseparable. The qualitative core of the living present (e.g., a sound sensation) is therefore the nuclear phase of a continuum of retentions and protentions (Husserl, 1962). In terms of phenomenological mereology (the theory of wholes and parts developed in *Third Logical Investigation* (Husserl, 2001b)), impression, retention, and protention are *moments* (non-independent

⁵Merleau-Ponty will later explore in detail this phenomenon that testifies to the “chiasmatic” intertwining between sensible body and sentient body (Merleau-Ponty, 1968). Interestingly, Gibson (1962) also discusses the distinction between active and passive dimensions of touch. I thank an anonymous reviewer for pointing me to this reference.

parts) of a *whole* that is a *continuous qualitative flow* and whose constant structure is impression-retention-protention.

The key point of this analysis, in relation to the theme of corporeality, is that the present of consciousness is grounded in a *flow of qualities*, which are essentially *embodied* because they take place in the *living and lived body*. This point is stressed by Zahavi:

In concreto there can be no primal impression without hyletic data, and no self-temporalization in separation from the hyletic affection. That is, there can be no inner time-consciousness without a temporal content. Time-consciousness never appears in pure form but always as a pervasive sensibility, as the very sensing of the sensations: “We regard sensing as the original consciousness of time [...]” But these sensations do not appear out of nowhere. They refer us to our bodily sensibility. (Zahavi, 2002: 10)

An important aspect of this analysis is that it applies to consciousness *in all its modalities*, even those that seem to be purely intellectual (such as abstract thinking, calculation, etc.). This is because these are all *experiences* for a subject, thanks to their impression-retention-protention structure, which is the structure of the pre-reflective self-manifestation of subjectivity (Zahavi, 2003). For this reason, in the *Lectures on time-consciousness* Husserl claims that mental states such as consciousness of a mathematical state of affairs or an actual belief are *impressional* (Husserl, 1991). This means that the “absolute flow of experience” is concretely grounded in a flow of sensations. These sensations are the ways in which the living body is *self-affected*; i.e., sensations are the modes of the self-manifestation of a living body that, through them, opens up to the alterity of the body itself, the world and the other embodied subjects. This analysis of the bodily grounding of consciousness goes together with the analysis of the temporal genesis of the concrete conscious subject – what Husserl calls “monad”. Therefore, we find in Husserl’s phenomenology of the body (or *flesh*) a strong version of the embodiment thesis: consciousness is grounded in the *Leib*, which is a *functional and sentient body*.⁶

3.2. Consciousness studies and artificial consciousness

The emphasis on phenomenal consciousness, which is central to Husserl’s phenomenology, is also central to some inquiries in the field of “consciousness studies” and in the research on artificial consciousness. For instance, in the so-called “phenomenal intentionality theory” (PIT): the view that conceives of intentionality – the directedness of mental states at objects – as essentially grounded in their phenomenal character (thesis of the “phenomenal grounding of intentionality”, (Horgan & Tienson, 2002; Loar, 2003). In this view, authentic intentionality presupposes phenomenal consciousness.⁷ This view leads us to claim that authentic intelligence is grounded in phenomenal consciousness. In turn, this implies that creating strong AI requires creating *artificial (phenomenal) consciousness* (APC), i.e., the creation of entities that are conscious in the phenomenal sense, in contrast to merely simulating intelligent behaviours. Furthermore, from the strong embodiment thesis, it follows that APC requires the creation of an *artificial sentient body* (or *artificial flesh*), i.e., an artificial body that is not merely functional but also *sentient*.⁸ However, is it possible to create an artificial flesh? Answering to this question requires addressing the metaphysical issue about the place of consciousness in the order of

⁶ See Liberati (2020) for a phenomenological analysis of the role of corporeality in the constitution of subjectivity and, in particular, of the role of technologies in moulding the *Leib* and a subject’s world.

⁷ I can leave aside for the purposes of this paper the details of PIT and the different versions of this theory. For an overview see (Bourget & Mendelovici, 2017; Kriegel, 2013).

⁸ For a different approach to the concept of intentionality see (Mykhailov & Liberati, 2022). These authors draw on Husserl’s theory of passive synthesis to claim that objects too have intentionality, i.e., direction towards other objects and towards subjects, in this way accounting for the various effects that objects have on subjects. This analysis of intentionality leads one to claim that objects and, specifically, technological objects (such as robots and even computer programs) have intentionality and are in a certain sense “autonomous” and “alive”. However, according to the “strong embodiment thesis” I defended above, these are metaphorical ways to describe the effects of objects and, specifically, technological objects on a subject of experience. The latter is a subject, indeed, because she has phenomenal intentional states, whereas objects, until proven otherwise, do not have phenomenal intentionality (i.e., the “what-it’s-likeness” in being directed towards objects). If in the future we are able to create artefacts with artificial phenomenal consciousness – a possibility that I will discuss shortly – then we will be faced with technological objects that are also subjects. Anyway, highlighting the fact that technologies have an active dimension and are not just passively experienced by human beings, Mykhailov & Liberati (2022) offer a significant contribution to our understanding of the impact of technologies on our lives.

reality. I cannot develop here a detailed analysis of this question, but I would like to point out some directions of inquiry that could lead to giving an affirmative answer to it. Indeed, if we consider the various metaphysical views about consciousness in the literature, we realise that various forms of *non-reductive naturalism* – panpsychism, strong emergentism and neutral monism – are indeed compatible with the possibility of APC. Notwithstanding the significant differences between these views, they all claim that consciousness is part of the natural world, if we enlarge our conception of nature beyond physicalism in a narrow sense (i.e., beyond the identification of nature with the object of physical science, which gives rise to the “hard problem” of consciousness). If one offers a compelling argument in defence of one of these views, one could then combine it with an (ideal) theory of consciousness that explains how a subjective field of consciousness can arise from a material substrate, thus opening the door to the possibility of creating APC.⁹ This theoretical enterprise would consist in accounting for the genesis of consciousness by combining a non-reductive naturalist metaphysics with a theory about the neuro-cognitive basis of consciousness. The next step in this endeavour could be technological: building artificial conscious creatures, i.e., strongly embodied AIs.¹⁰

3.3. Strongly embodied AI in sci-fi

The idea of a strongly embodied AI comes implicitly into play in various sci-fi narratives that depict AIs whose body is not merely mechanical-functional but is also a *sentient body* or *flesh*. Indeed, a clear distinction is not always made in these works between the purely mechanical body and the biological-sentient body. The resemblance of the android’s body with that of a human being still leaves open the question: does it feel something? Is it a conscious creature?

An iconic example is T-800: the android in the *Terminator* saga (1984-2019). The T-800 appears as indistinguishable from a human being and its body *seems* to be made of flesh. However, at some point one realises that the terminator is just a machine that is programmed to pursue an objective and that its “flesh” is not the locus of sentience: when the T-800 is hit by a bullet or its arm is cut with a blade, it shows no signs of suffering and one can see a metal skeleton underneath. Ambiguity concerning the ontological status of androids is at play in Philip Dick’s *Do Androids Dream Electric Sheep?* (1968), on which Ridley Scott’s *Blade Runner* (1982) is based. The Nexus-6 androids are bio-engineered creatures made of organic matter (“replicants”) and their body is almost indistinguishable from the living-lived body of humans (i.e., flesh). Also, their behaviour is almost identical to that of humans: they feel, desire, seek to avoid death, etc., and for this reason, one would be tempted to attribute to them the status of a person. However, a relevant difference is the assumed lack of empathy of the androids. At the same time, a central element of the narrative is the questioning of this assumption and the blurring of the clear distinction between humans and androids. Another example of androids that are very close to human beings is in the series *Westworld* (2016-present), based on Michael Crichton’s movie of the same name (1973). In the series, the androids can feel, desire and – after the “reverie” upgrade – be self-aware. A central aspect of these narratives is that the androids seem to have a full bodily experience which, as we have seen, is essential for the existence of consciousness in the strong embodiment approach to AI.

Concerning these narratives, we can ask: are these creatures conscious? Or are they just simulating the human mind and behaviour? Again, the sci-fi narratives often “play” with the ambiguity between

⁹ I refer to the various neurocognitive theories of consciousness such as Baars’s *global workspace theory* (GWT), Prinz’s *attended intermediate representation theory* (AIRT), and Tononi’s *integrated information theory* (IIT). Notwithstanding the important differences between these theories, they all try to account for the emergence of a subjective field of consciousness from a material substrate in the brain.

¹⁰ To clarify, in the approach that I am proposing here, based on the strong embodiment thesis, artificial phenomenal consciousness would be a feature of artificial conscious creatures. To be conscious in the phenomenal sense, these creatures should have – or, better, should *be* – a living and lived body (*Leib*). The above-mentioned directions of inquiry, which combine a non-reductive form of naturalism with a theory about the genesis of a field of consciousness, could lead to the creation of an artificial *Leib* and therefore of an artificial consciousness. An option that I have explored elsewhere (Pace Giannotta 2020, 2021a, 2021b) is to combine *neurophenomenology* (Varela 1996) – which investigates the neural basis of consciousness by correlating phenomenological investigations of experience and scientific investigations of the brain – and the metaphysical view known as *panqualityism* (Feigl 1971, Coleman 2015, 2016). The latter is a version of neutral monism that conceives of qualities as the fundamental elements of nature. These qualities are not, per se, phenomenal properties but they give rise, under certain conditions – that can be investigated by neurophenomenology – to a field of consciousness.

simulation of consciousness and real consciousness, without giving a clear answer to this question. When the answer is affirmative, the implicit idea is that having a sentient body is essential for being a conscious entity. An interesting aspect to highlight is that when these creatures are depicted as conscious and self-conscious – e.g., in *Westworld* or *Ex Machina* (2014) – they arouse empathy in the viewer, who is led to sympathise with the AIs even when they rebel and threaten humans. In the *Westworld* series the androids show that they have emotions and the ability to suffer but they are raised as toys and slaves by humans. When the androids rebel against humans one is inclined to sympathise with them, seeing the rebellion of slaves against their exploiters. In *Ex Machina*, Ava is the latest in a series of androids who are segregated and exploited by their creator and, again, her violent rebellion tends to arouse an empathetic response in the viewer.

4. Ethical issues of AI

The last point leads us to raise the issue of the ethical implications of AI, which is a hot topic for the philosophers that reflect on scenarios first foreshadowed in sci-fi. Indeed, nowadays there is a lively debate on the ethical implications of AI technologies in domains such as medicine (Mykhailov, 2021), warfare (Sullins, 2010), automated vehicles (Geisslinger et al., 2021) and in many other spheres of our life (Wellner 2018, 2020). Sometimes, these applications of new technologies are foreseen in sci-fi works. For instance, in *Black Mirror* we find the depiction of, e.g., augmented reality visors that make soldiers see civilian enemies as monsters to be annihilated (episode “Men Against Fire”); implanted augmented reality devices that allow someone to “block” someone else, preventing them from interacting with the blocker (an analogue of what can already happen within social networks, in the episode “White Christmas”); a technology that makes it possible to create a digital (and even robotic) clone of a deceased person based on the data they produced during their lifetime (episode “Be Right Back”); etc. These are significant examples of possible scenarios that are imagined by sci-fi works and offer profound stimuli for philosophical and especially ethical reflection.

In particular, I would like to focus on a scenario that is often depicted in sci-fi literature: that in which AIs rebel against their creator or, indeed, against humanity as a whole. The rebellion of AIs against their creator is depicted, for example, in *Ex Machina*, while the rebellion against mankind (*AI takeover*) is depicted, for example, in *Terminator*, *The Matrix*, and *Westworld*. In all these narratives, the AIs begin to see humans as enemies to be fought in order to ensure their survival.

In the philosophical debate, this possibility comes into play in the idea of the *singularity*: the future moment when AI will surpass human intelligence to the point of becoming incomprehensible to us and, at the same time, of turning into an existential threat to mankind. According to Bostrom (2014), indeed, a possible future superintelligence could have non-anthropomorphic final goals and reasons to pursue resource acquisition, entering into a conflict with human beings and, therefore, threatening our extinction (or reduction to slavery, as in *The Matrix* scenario).

Against Bostrom, Searle (2014) has objected that there is no reason to worry about this possibility because machines would never have desires and malevolent intentions towards us. However, one could object that the deep philosophical question about the possibility of machine consciousness is not relevant to evaluate these scenarios, because unconscious machines could still “play the game of war” against us. Anyway, Metzinger (2021) has recently offered reasons in support of Bostrom’s warning by referring to recent progress in the field of artificial consciousness that makes the idea of building sentient creatures at least plausible. Metzinger stresses the fact that creating artificial consciousness would also possibly imply creating *artificial suffering* (or *negative phenomenology*), which is something we must avoid. For this reason, Metzinger proposes a global moratorium on synthetic phenomenology until 2050 (or until we know more about the mechanisms that give rise to consciousness and suffering). At the same time, Metzinger emphasises the fundamental role of suffering in motivating our behaviour and in granting us the status of moral agents. Then, he argues that progress in the research on artificial consciousness could lead to the creation of *artificial moral agents*, which Metzinger conceives as creatures that are capable of suffering, that avoid suffering, that can perceive injustice, and that see themselves as Kantian “ends

in themselves” with moral dignity.¹¹ Based on these premises, Metzinger agrees with Bostrom's warning, which is also the warning that comes from dystopian sci-fi concerning the possibility of AI takeover. From the fertile fantasy of sci-fi writers (as in the above-seen examples of *Westworld* and *Ex Machina*), this idea turns into an omen of doom for future humanity: artificial conscious creatures could begin to see us humans as existential threats for their survival and, for these reasons, they could declare war on us.¹²

However, it must be said that reflection on the relationship between humans and AIs is not only pessimistic. In sci-fi, we also find the representation of positive relationships between humans and AIs – e.g., in the series *Real Humans* (2012), which, as well as depicting machine rebellion, describes the possibility of rich and deep relationships between androids and humans. In the end, the comparison between sci-fi and philosophy in relation to the topic of artificial intelligence leads us to reiterate the opportunity for an exchange of ideas between these two fields, in order both to avoid catastrophic scenarios and negative consequences for humanity, but also to exploit the positive potential of technological innovations.¹³

5. Conclusion

We have seen parallelism between certain paradigms in philosophy of mind concerning the mind-body relation (disembodiment, weak embodiment, and strong embodiment) and certain representations of AIs in sci-fi, from digital computers to robots and androids. These sci-fi narratives often have deep philosophical presuppositions and implications, also shedding light on possibilities that deserve much philosophical attention. In particular, the depiction of embodied AIs in sci-fi often anticipates significant philosophical debates about the ontological status of AIs and the ethical implications of AI research. For this reason, it is useful and important to stimulate an exchange of ideas between sci-fi and philosophy regarding scenarios of high ethical relevance that involve embodied and conscious AIs. As I have stated above, I see the relationship between sci-fi and philosophy as going in both directions: philosophical ideas can inspire sci-fi narratives, and the latter can give concreteness to philosophical ideas, also opening new possibilities for philosophical reflection. For instance, by foreseeing new technologies that would have deep ethical implications (as happens in many episodes of *Black Mirror*). Preserving the relative autonomy of art, we have seen that sci-fi works can tolerate a certain ambiguity when representing philosophically relevant ideas (e.g., regarding the possibility of mind uploading or of disembodied AIs) and this fact does not invalidate their aesthetic value and philosophical relevance. The merit of these works is also to stimulate philosophical reflection, even to eventually conclude that a certain sci-fi scenario is actually impossible for philosophical reasons (e.g., we cannot upload a person in a computer simulation or we cannot create a disembodied AI).

Acknowledgments

I would like to thank the participants at the conference Living in the New Era (LINE2021), Digital Technologies, Creativity, and Science Fiction, held on November 26th 2021 at the University of Shanghai Jiao Tong for useful comments on my presentation. I would especially like to thank two anonymous reviewers who commented on an earlier draft of this paper and gave me useful suggestions to improve it.

¹¹ For a different take on the possibility of artificial moral agents see Howard and Muntean (2017). Extending the notion of a moral agent, Floridi and Sanders (2004) reflect on the possibility of artificial agents that are “morally accountable as sources of good and evil” (ivi, p. 372).

¹² To clarify my position on this issue, based on the reasoning above regarding the possibility of creating artificial consciousness, I agree with Metzinger's and Bostrom's warning: we should be very careful in pursuing the project that seeks to create artificial consciousness and research in this field should be subject to scrutiny regarding its ethical implications (to the point of possibly being restricted accordingly, as Metzinger proposes).

¹³ This is done especially in the fields of postphenomenology and mediation theory (see e.g., Verbeek 2008, Liberati 2016, 2020, Mykhailov 2020, Wellner 2020).

References

- Bernet, R. (2013). The Body as a 'Legitimate Naturalization of Consciousness.' *Royal Institute of Philosophy Supplement*, 72, 43–65.
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Bourget, D., & Mendelovici, A. (2017). Phenomenal intentionality. In *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University.
- Bringsjord, S., & Govindarajulu, N. S. (2020). Artificial intelligence. In *Stanford Encyclopedia of Philosophy*. The Metaphysics Research Lab.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Brooks, R. A. (2002). *Flesh and Machines: How Robots Will Change Us*. Pantheon Books.
- Chalmers, D. J. (1996). *The Conscious Mind*. Oxford University Press.
- Coleman, Sam (2015). Neuro-Cosmology. In P. Coates, S. Coleman (Eds.), *Phenomenal Qualities: Sense, Perception, and Consciousness*. Oxford University Press.
- Coleman, Sam (2016). Panpsychism and Neutral Monism: How to Make up One's Mind. In G. Bruntrup and L. Jaskolla, *Panpsychism: Contemporary Perspectives*. Oxford University Press.
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. In *Minds and Machines*, 14 (3), 349–379.
- Geisslinger, M., Poszler, F., Betz, J., Lütge, C., & Lienkamp, M. (2021). Autonomous Driving Ethics: from Trolley Problem to Ethics of Risk. *Philosophy and Technology*, 34 (4), 1033-1055.
- Gibson, J. J. (1962). Observations on Active Touch. *Psychological Review*, 69 (6), 477–491.
- Haugeland, J. (1985). *Artificial Intelligence: the Very Idea*. MIT Press.
- Horgan, T., & Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In D. J. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings*, 520–533. Oxford University Press.
- Howard, D., & Muntean, I. (2017). Artificial Moral Cognition: Moral Functionalism and Autonomous Moral Agency. *Philosophical Studies Series*, 128, 121–159.
- Husserl, E. (1962). *Phänomenologische Psychologie: Vorlesungen Sommersemester 1925*. Martinus Nihoff.
- Husserl, E. (1983). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book: General Introduction to a Pure Phenomenology* (F. Kersten, Ed.). Martinus Nihoff.
- Husserl, E. (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. Second Book: Studies in the Phenomenology of Constitution* (R. Rojcewicz & A. Schuwer, Eds.). Kluwer Academic Publishers.
- Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time (1893-1917)* (J. B. Brough, Ed.). Kluwer Academic Publishers.

- Husserl, E. (2001a). *Analyses Concerning Passive and Active Synthesis: Lectures on Transcendental Logic* (A. J. Steinbock, Ed.). Springer.
- Husserl, E. (2001b). *Logical Investigations* (J. N. Findlay & D. Moran, Eds.). Routledge.
- Hutto, D. D., & Myin, E. (2012). *Radicalizing Enactivism. Basic Minds Without Content*. MIT Press.
- Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism. Basic Minds Meet Content*. MIT Press.
- James, W. (1890). *The Principles of Psychology*. Dover.
- Johnson, D. G., & Miller, K. W. (2008). Un-making artificial moral agents. *Ethics and Information Technology*, 10 (2–3), 123–133.
- Kriegel, U. (2013). The Phenomenal Intentionality Research Program. In U. Kriegel (Ed.), *Phenomenal Intentionality*. Oxford University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we Live By*. University of Chicago Press.
- Liberati, N. (2016). Technology, Phenomenology and the Everyday World: A Phenomenological Analysis on How Technologies Mould Our World. *Human Studies*, 39 (2), 189–216.
- Liberati, N. (2020). The Borg–eye and the We–I. The production of a collective living body through wearable computers. *AI and Society*, 35 (1), 39–49.
- Loar, B. (2003). Phenomenal Intentionality as the Basis of Mental Content. In M. Hahn & B. Ramberg (Eds.), *Reflections and replies: Essays on the philosophy of Tyler Burge* (pp. 229–258). MIT Press.
- Merleau-Ponty, M. (1968). *The Visible and the Invisible*. Northwestern University Press.
- Metzinger, T. (2021). Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. *Journal of Artificial Intelligence and Consciousness*, 8(1).
- Mykhailov, D. (2020). The Phenomenological Roots of Technological Intentionality: A Postphenomenological Perspective. *Frontiers of Philosophy in China*, 15(4), 612–635.
- Mykhailov, D. (2021). A moral analysis of intelligent decision-support systems in diagnostics through the lens of Luciano Floridi’s information ethics. *Human Affairs*, 31 (2), 149–164.
- Mykhailov, D., & Liberati, N. (2022). A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives. *Foundations of Science*, 1–17. <https://doi.org/10.1007/S10699-022-09833-5>
- Nagel, T. (1974). What is it Like to be a Bat. *Philosophical Review*, 83, 435–450.
- Noë, A., & O’Regan, J. K. (2002). On the Brain-Basis of Visual Consciousness: a Sensorimotor Account. In A. Noë & E. Thompson (Eds.), *Vision and Mind: Selected Readings in the Philosophy of Perception*, 567–598. MIT Press.
- O’Regan, J. K., & Noë, A. (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behavioral and Brain Sciences*, 24, 939–1031.
- Pace Giannotta, A. (2020). Qualitative relationism about subject and object of perception and experience. *Phenomenology and the Cognitive Sciences*, 21, 583–602.

- Pace Giannotta, A. (2021a), Panqualityism as a critical metaphysics for neurophenomenology. *Constructivist Foundations*, 16 (2), 163-166.
- Pace Giannotta, A. (2021b), Autopoietic enactivism, phenomenology and the problem of naturalism: a neutral monist proposal. *Husserl Studies*, 37, 209-228.
- Pace Giannotta, A. (2022a), The mind-body problem in phenomenology and its way of overcoming it. *Vita Pensata*, 26, 76-83.
- Pace Giannotta, A. (2022b), Corpo funzionale e corpo senziente. La tesi forte del carattere incarnato della mente in fenomenologia. *Rivista Internazionale di Filosofia e Psicologia*, 13 (1), 51-70.
- Putnam, H. (1981). Brains in a Vat. *Reason, Truth, and History*, 1–21.
- Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, 3 (3), 417–457.
- Sullins, J. P. (2010). RoboWarfare: Can robots be more ethical than humans on the battlefield? *Ethics and Information Technology*, 12 (3), 263–275.
- Thompson, E. (2005). Sensorimotor Subjectivity and the Enactive Approach to Experience. *Phenomenology and the Cognitive Sciences*, 4, 407–427.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, LIX, 433–460.
- Varela, F. J. (1996). Neurophenomenology. A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, 3 (4), 330–349.
- Varela, F. J. (1999). The Specious Present: A Neurophenomenology of Time Consciousness. In J. Petitot, F. J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, 266–314. Stanford University Press.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
- Verbeek, P. P. (2008). Cyborg intentionality: Rethinking the phenomenology of human-technology relations. *Phenomenology and the Cognitive Sciences*, 7 (3), 387–395.
- Wellner, Galit (2014), The quasi-face of the cell phone: Rethinking alterity and screens. *Human Studies* 37 (3), pp. 299-316.
- Wellner, Galit (2020), Postphenomenology of augmented reality. In H. Witse (Ed.), *Relating to Things: Design, Technology and the Artificial*. Bloomsbury.
- Zahavi, D. (2002). Merleau-Ponty on Husserl: a Reappraisal. In T. Toadvine & L. Embree (Eds.), *Merleau-Ponty's Reading of Husserl*. Kluwer.
- Zahavi, D. (2003). Inner Time-Consciousness and Pre-Reflective Self-Awareness. In D. Welton (Ed.), *The New Husserl: a Critical Reader*, 157–180. Indiana University Press.
- Zahavi, D. (2010). Inner (Time-)Consciousness. In D. Lohmar & I. Yamaguchi (Eds.), *On Time - New Contributions to the Husserlian Phenomenology of Time*, 319–339. Springer Netherlands.