PLAYING WITH ROBOTS: NARRATIVES OF REPLACEMENT, INTELLIGENT MACHINES AND VIRTUAL REALITY IN THE CYBERNETIC IMAGINATION

By

VLAD B. JECAN

(Under the Direction of Mihai I. Spariosu)

ABSTRACT

The cybernetic imagination refers to literature and fiction focusing on the socio-economic and ethical impact of machines, automation and computers on society. In the last couple of centuries, the cybernetic imagination experiments consistently with scenarios of technological dystopia developing a generalized anxiety towards the consequences of technology. Today, research is concerned with the apparent replacement of reality by virtual worlds and an impeding A.I. takeover as digital technology permeates almost all economic and social sectors. Therefore, this study intends to answer a question asked by Isaac Asimov in 1984, "why this fear of robots?" After the Second World War, the questions asked in parallel to Asimov's, particularly in the literature of Philip K. Dick, are (1) what is reality? And (2) what does it mean to be human? As such, since the cybernetic imagination frequently portrays violent contests between humans and machines, we proceed from the idea of "agonistic play" discussed by Mihai I. Spariosu and introduce the Quantum Relations Principle as an alternative ontological, ethical and coding paradigm. In this sense, this study (1) observes the tensions of replacement leading towards ontological uncertainty as a form of playing with reality, (2) analyzes potential intimate emotional connections in playing with affection with intelligent machines in the fiction that resembles (a) cognitive affection, when the artificial entity does not have a physical presence, (b) partial cognitive-physical affection, when the intelligent machine has only certain human-like body parts, and (c) replication, in fiction that portrays robots as indistinguishable physically and behaviorally from human observers. However, the play of replacement is suspended in the cybernetic imagination and the history of computing with the hacker culture. Playing with computers, rather than having similar expectations of replacement, finds in the hacker culture a creative imperative proposing ethics of playful exploration and innovation. This study then concludes by adopting the a "process ontology" to observe reality outside the play of replacement between the real and artificial and draws from the concept of "oikeiosis" to propose the development of digital virtue for responsible A.I. and artificial reality design.

INDEX WORDS: philosophy of play, cybernetic imagination, intelligent machines, cyberspace, virtual reality, literature, philosophy of technology, ethics, digital virtue

PLAYING WITH ROBOTS: NARRATIVES OF REPLACEMENT, INTELLIGENT MACHINES AND VIRTUAL REALITY IN THE CYBERNETIC IMAGINATION

by

VLAD B. JECAN

B.A, University of Babes-Bolyai, Romania, 2009

M.A., University of Babes-Bolyai, Romania, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

PLAYING WITH ROBOTS: NARRATIVES OF REPLACEMENT, INTELLIGENT MACHINES AND VIRTUAL REALITY IN THE CYBERNETIC IMAGINATION

by

VLAD B. JECAN

Major Professor: Mihai I. Spariosu

Committee: Dorothy M. Figueira
Dezso Benedek

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
May 2020

DEDICATION

To my parents Doru and Otenzica Jecan.

ACKNOWLEDGEMENTS

I wish to express my gratitude to Prof. Mihai I. Spariosu for his intellectual guidance and mentorship. In addition, I would like to thank my colleagues and professors at the Department of Comparative Literature and Intercultural Studies at the University of Georgia for their support and feedback on the different stages of this research. Special thanks go to Jill Talmadge, business manager of the Department, for her prompt administrative advice and solutions. I also thank my students in the various classes that I have taught at UGA, in particular "Playing with Robots" and "Detective Fiction."

Special thanks are also due to Hardy F. Schloer, Philip A. Gagner, Gheorghe M. Stefan, Dorothy Figueira, and Dezso Benedek for their generous advice and kind comments on the ideas explored in this study.

Last but not least, I am immensely grateful to my friends Mihai Surducan, Rob VanLandigham, Adrian Costea, Ion Josan, Susan Wiley and her family, Corina Beleaua, Vlad Albu, Pierre Sutherland, Daniel Mckenzie, Shabnam Gosh, Bogdan Cucu, Hristo Iliev, Adela Morar, Kelvin Rosa, Sinziana Jurau, Flavia Topan, Ioana Butcovan, Anne Margaret LaBarre, and others for their encouragements and support during challenging times.

Table of Contents

Introduction: “Why this fear of robots?”

For an unknown reason, contemporary research and popular opinion agree that the replacement of reality by virtual worlds is inevitable and expect, to a point of accepting defeat, that artificial intelligence will take over the world while our future is destined to resemble a cyberpunk dystopia. The cybernetic imagination, through which we understand the literature and fiction concerned with the impact of computers and machines on society, appears to have developed such expectations while recent developments in advanced technology seem to produce the devices that make real the socio-political concerns once imagined by science fiction.

In the introduction to *Machines that Think*, an anthology of short stories about computers and robots, Isaac Asimov asks: “Why this fear of robots?” (Asimov, 1984, p. 1). Technophobia is the main reason for Asimov, and after he enumerates several potential causes, such as the impact of automation on labor or the environment, religious resistance, he asks, nevertheless: “[a]s the computer grows more compact, more versatile, more complex, more capable, more *intelligent*, might it not replace not just a person but all humanity?” (*Ibid.*, p. 6). Not just that machines can become intelligent or capable, but that the contemporary anxiety of automation stems in the imagined possibility that they might replace humanity. In this sense, blaming generalized technophobia in an age where technology mediates almost every human activity, even during the time of Asimov, is problematic.

Humanity has adapted to various kinds of technologies. While history can provide us with examples of machines and early automation during the Industrial Revolution, the Great War, and so on, we cannot find relevant guidance to deal with machines replicating human thinking. Literature, however, is a more suitable source since the robotic imagination is found in ancient and medieval literature, folklore, and tales, before the 19th century focus on exploring the consequences, real or imagined, of intelligent machines. Most writers of the 19th and 20th centuries worry about the violent replacement of individuals in the workforce by technology or imagine artificial realities or even robot takeover. From fiction, the fear of replacement has been introduced into the newly formed, post-war computational sciences. After 1945, computers, code, and machines merge into stories of technological apocalypse. Asimov, during this time, wrote his Laws of Robots attempting to put some order in the cybernetic imagination and trusted them to serve as a deontological guide for intelligent machine design. However, his laws, explored in the chapter *Playing with Replacement* below, provide an excellent framework to imagine robots in fiction, but in practice, as other authors show, the problem of ethics is far more nuanced.

Asimov's question is also rhetorical. With technophobia, the science fiction author encourages us to look at ourselves. His robot fiction takes on his laws, challenges them, changes them, adds a primordial rule to safeguard humanity from harm, and tries to convey the possibility of useful, benign, robots. Our own fear of replacement stems probably from our unconscious desire for perpetual domination of the machine. Much, if not all our fiction portrays robots as loyal slaves. Therefore, robot fiction might reveal something about us, since to attribute desires of human enslavement to machines incapable of orienting themselves independently in the world, unless supported and directed during every step, with operating systems always waiting

for the next error, is utter absurdity. Furthermore, why would we create our own artificial replacement? We imagine technology, innovate playfully, and then expect to be replaced by it. In this sense, the 'robot takeover' or fears of replacement are metaphors for the violent conflict already existing among us, particularly from the violent play that may be at the very center of our Western culture.

The narrative of expected replacement by technology extends to the digital realm as well. It is not just robots that will take over the world, but virtual reality will, the argument goes, replace reality trapping us in the confinements of artificial environments. The question of reality has a long tradition in philosophy and literature. Today, however, it is no longer a question of philosophical wonder or literary art, but of science wanting to prove reality as a computer simulation or a holographic universe. If we assume the theories explained below to be true, then the scientific and philosophical reasoning blend into a new construct resembling scientific romance, as science fiction was called in the early days, with the pretension of fact. Quite ironically, the Simulation Theory or the Technological Singularity, or even the Holographic Universe, resemble philosophizing as much as the work of Descartes, Spinoza, Leibniz, and others. But what if we truly live in a computer simulation? As Seneca anticipated Lucilius' response in a letter explaining to him the classes of things which exist, "Very well,' say you, 'what good shall I get from all this fine reasoning?' None, if you wish me to answer your question" (Seneca, 1925, p. 139). In fact, it would be wise not to let our mysterious software engineers think we are on to them (Greene, 2019).

Artificial reality, implanted memories, and intelligent machines replicating human likeness in mind and body develop from questions in old philosophy to exploratory journeys in human psychology in the second half of the twentieth century. Asimov's question, as his fiction

suggests, is best asked when the robot is clearly recognizable, but the violent play of replacement appears more suitably articulated in the context of thinking about the artificial whole (universe). The literature of Philip K. Dick and the films inspired by his work tend to offer a more nuanced view of the problem. In Dick, "why this fear of robots?" becomes a question revealing another two questions: (1) what is reality? and (2) what does it mean to be human? Furthermore, the ethics surrounding good design and use of artificial intelligence and virtual reality are experimental as much as the questions asked by literary and science fiction authors.

Intelligent machines and virtual reality, or cyberspace, tend to penetrate most sections of human activity and intellectual work, including the adaptation of derivative technologies to the scholarly fields of the humanities. Therefore, our approach must, inevitably, be holistic. In this sense, we will not focus on particular literary works and fiction but look at key texts in the cybernetic imagination and advances in science. This study is not intended to rely too much on a single author, but to emphasize the author's contribution to the cybernetic imagination relevant to our concerns, as well as cultural approaches. Therefore, our investigation proceeds from Asimov's question into the other two questions derived from it and results in an experiment of ethics drawn from literature and the computational sciences. Our study intends to define the term "cybernetic imagination", clarify the concept of agonistic play, identify replacement, to observe the playing with reality, the emotional effects of human-like intelligent machines, and consider possibilities outside of the imperative of replacement in the hacker culture and the Quantum Relations Principle, a theory that incorporates digital objects, literary works, mental phenomena, and impalpable things into a new view of reality. Lastly, we return to the question of ethics. In the next subsections, we proceed with a summary of the topics covered in this study which will be expanded in their appropriate chapters.

The Cybernetic Imagination

The term "cybernetic imagination" was coined by Patricia S. Warrick to envelop the literature pertaining to the uses and consequences of computers and intelligent machines. "All these stories" she notes "grew from an awareness of the radical future unfolding as man yoked knowledge and mechanical power" (Warrisck, 1980, p. 2).

As humanity began to automate parts of labor during the Industrial Revolution and starting with classic authors such as H.G. Wells and Jules Verne, "the machine replaced the flying carpet, and future possibility replaced pure fantasy" (*Ibid.*, p. 2). Therefore, the imagination in question is found within several works of fiction speculating about the future of humanity in a world where digital technology and artificial intelligence are ubiquitous. Warrick, in turn, borrows the concept of cybernetics from its founder, Norman Wiener. The broad field of cybernetic, according to Wiener, is "to find the common elements in the functioning of automatic machines of the human nervous system, and to develop a theory that will cover the entire field of control and communication in machines and in living organisms" (*Ibid.*, p. 8). *Cybernetics* originates in the Greek word *kybernetes,* and was also used by Plato in *The Republic,* meaning steersman. Our use of "cybernetic imagination" in this study refers to the fiction developed around the idea of artificial intelligence, usually summed up as robot, intelligent machine, android, replicant, on the one side, and the literature concerning virtual reality, or cyberspace, on the other side. This type of fiction, frequently understood as science fiction, is "a literature of anticipatory process, it must suggest a spectrum of future worlds, and by this creative act remind man how transient are the patterns of his contemporary cultural landscape" (*Ibid*. p. 16).

However, it is important to note that we will investigate other literary works that do not necessarily fall into the genre of science fiction.

Play

Theories of play are abundant in scholarly investigations, particularly in the study of sport and psychology. Here, we wish to clarify our premise of play as agonistic contest. Future chapters will employ appropriate variants of the philosophy of play in their area of focus.

Our premise is that Western culture is grounded in agonistic play as contest resulting in zero-sum games of domination that permeate much of our civilizational development. Johan Huizinga was among the first to observe that agonistic play is at the center of civilization. According to the Dutch scholar writing in *Homo Ludens,* the sum total characteristics of play apply not only to games of leisure and entertainment, but provide the rule structure for traditional Western institutions, art, language, and even war. Mihai I. Spariosu notes in *Modernism and Exile* that Huizinga's "most important contribution to a theory of culture as play is to have drawn attention, in the wake of Nietzsche, to the fact that violent contest, such as war, is a favored cultural form of play in Western civilization" (Spariosu, 2015, p. 21). In *Homo Ludens*, Huizinga returns to Plato and some pre-Socratic philosophers to show the evolution of play as violent contest. As our Western culture is grounded in the ancient philosophies of Greece and Rome, violent contest appears not only in war but also in more benign institutions and instances such as politics and in our economic systems. Spariosu writes that, before Socrates, play is "inextricably linked to the notion of immediate physical force, as in such words as *agon* and *athlon* – gradually loses its link until it becomes *paidia*, a word that initially denotes only the harmless play of

children and then becomes, in Plato, a philosophical term for nonviolent cultural play in general" (Spariosu, 1989, p. 6). The influence of the Homeric epics fostered a sense of virtue based on prowess in battle. While some pre-Socratic philosophers suggested a median alternative, a balance between agonistic and peaceful play, "agonism drives the rules." Even in the *Iliad*, we can identify moments in which the agonistic mentality is suspended. It is usually during funeral services and celebrations that the heroes know moments of peace. The gods, however, just like the humans, are also engaged in contests among themselves using the Greeks and Trojans as their playthings. Poetry and art, the play of the Muses, were initially dismissed by Plato only to be reintroduced as educational tools for proper citizens of his *Republic*. In the ideal state, Socrates retains the agonistic play but in its logo-rational and nonviolent form. If Achilles was the example of prowess in battle, then Odysseus showing *metis* (cunning) was the favorable model for contest in the agora. According to Spariosu, "although it is usually translated as 'virtue' or 'excellence,'" cunning "is the main agonistic virtue needed in the assembly, physical prowess is the main *arete* needed on the battlefield, and *aristeia* is the most important means by which the hero can display this *arete*" (Spariosu, 1991, p. 9). The concept of play as contest reformulated in logo-rational terms can be noted throughout various stages of civilizational development. Spariosu divides these stages into rational and pre-rational mentalities. In the pre-rational mentality, justice is based on the principle of might makes right, while in the rational mentality it is supported by law and morality. In a similar vein, we can argue that in a pre-rational mentality, war of contest is justified by winning, while in a rational mentality, the ability to wage war comes with additional clauses such as *jus ad bello* and *jus in bello*, confining violent play to rules of conduct in battle and appropriate, moral parameters for declaring war. According to Spariosu, agonistic play is transported into laws, politics, rhetoric, history, eros, philosophy

and literary criticism. In essence, the history of play in the Western world is the history of conflict. However, a distinct, independent form of play resides in the possibility of *as if*. In this state, play is experimental and, as such, it carries the potential to reconsider the agonistic element. Cyberspace provides the ideal playground for experiments with peaceful play. In this sense, I am not suggesting just constructing artificial reality according to irenic rules but using culture as data to simulate intercultural relations to observe moments of moral intervention.

The possibilities of play *as if* is shown by the ability of the hacker culture to remain calm while having potential access to the digital workings of our global economic system. The play of *as if* can, therefore, become a tool to imagine our technological future outside of our violent evidence of history. Huizinga, on the other hand, identifies in technology the increase of contest in the fabric of civilization.

As I have written on a different occasion (Jecan, 2020, p. 96), perhaps we can understand that our capitulation to intelligent machines and cyberspace is another reiteration of agonistic play, set this time in the playground of the logo-rational. Therefore, if we are unable to find a way around violent contest, then the expectations of replacement we imagine will inevitably be acted out through lines of code. Indeed, it is difficult to imagine a reality outside of familiar parameters, if not truly impossible to think of a peaceful alternative when our historical experience and cultural elements reveal the opposite. At this point, it may be that even if we recognize that alternative worlds may be possible, we do not have the proper language to convince us of its existence (Spariosu, 1997, p. 25). Our imagination of intelligent machines and virtual reality, as it stands, is based on a similar principle of play as contest. In the cybernetic imagination, both sides, represented by humans or physical reality and intelligent machines and artificial reality, play a war game of supremacy.

Thus, we require a new view of reality, a new theory to investigate and define reality that is outside of the typical imperative of replacement. Spariosu and co-author Schloer provide a new way of exploring reality and artificial intelligence by departing from Cartesian dualism and scientific objectivism and by considering all phenomena, including mental, digital objects, cyberspace, as parts of the interconnected whole. Objects in Quantum Relations Principle are evaluated not by their ontological status, but by their informational properties. Objects, therefore, arise ontologically in codependency, being determined by the constant mutual causality of their informational properties.

Replacement

Violent contest among machines and humanity appears in most of the cybernetic imagination. As I argue in *Chapter 3 – Playing with Replacement,* agonistic play takes refuge in science and the humanities in contest for the definition of reality and humanity through a series of articulation of robots as either to physically replace humanity or looming somewhere in the background, similar to agonism, to remind humanity of its own fear of replacement. After the Second World War, the play of replacement, which is my formulation of Spariosu's agonistic play, intends to incorporate traditional humanities into scientific terminology. In addition, technology is developed to replace traditional economic and social roles of people moving away from automated labor and into creative work. The chapters that follow explore the play of replacement on two primary levels: how it affects our ontological certainties and view of reality and the emotional persuasion of machines. For now, the stories discussed below should be able to provide a sense of the problem of replacement:

The 1950 short story titled *The Veldt* by Ray Bradbury imagines a peculiar room, a highly advanced piece of technology, where "you send out your thought" and "whatever you want would appear" (Bradbury, 1993, p. 9). As such, the individual is placed in full control over the consequences of his imagination in physical reality. "Veldt" refers to an open field in southern Africa, but the Germanic linguistic root means "field" which for the English reader sounds as the German world "welt", meaning world. In this sense, the room becomes a metaphor to suggest an artificial world in a constant state of emergence through renewed inputs from human imagination (Jecan, 2020, p. 92). Sometimes referred to as "the nursery", it serves as an adjacent to the house of a typical family, but due to its use it starts to erode familial relations to such an extent that the children, scared of losing their artificial world, imagine to reproduce the technology to murder their parents who want to shut the room down. The conclusion of the story shows that a technology with its ability to replicate items from reality and, perhaps, even simulate it in its entirety, which was left to the free use of inexperienced users, managed to foster an urge to guard the artificial world from the threats of replacement by the old reality to the point of sacrifice of family. What matters for the children is the simulated world stimulating their senses and delivering their wishes that made family bonds irrelevant and parental affection superfluous.

But perhaps no other story has achieved to portray the highest level of technological dread and anxiety about the future of intelligent machines than Harlan Ellison's *I have no mouth, and I must scream* (1968). Set in the near future when one can only imagine the darkest and most decadent setting, the A.I. known as AM, a clever adaptation of Descartes' famous maxim "I think, therefore I am", endlessly tortures the last humans the machine has left alive after its world takeover. Isaac Asimov and the other editors of the anthology *Machines that Think* write that "as a metaphor, the computer AM powerfully projects man's dilemma. His consciousness is trapped

and isolated in the prison-house of his own reasoning capacity" (Asimov et al., 1984, p. 234). AM, Ellison explains through the narrator of the story, "has been ruthless with its own life as with ours" (Ellison, 1984, p. 236). The ruthlessness and violent disposition were programmed in its "personality." The narrator continues to explain that "it strove for perfection. Whether it was a matter of killing unproductive elements in his own world-filling bulk, or perfecting methods for torturing us, AM was as thorough as those who had invented him – now long since gone to dust – could ever have hoped" (*Ibid.*, p. 236). In this sense, Ellison, perhaps unknowingly, bases the design of the machine's creators on the inherent, mute desires of the long-time defunct engineers who have developed the sentient program. In addition, we are told that AM can feel pain, a prerequisite of *sentience* while its self-awareness announces *sapience*, making the intelligent machine fit to resemble an actual living human. However, the distinction between machine and human is problematic here. Even the protagonist questions the identity of AM: "most of the time I thought of AM as *it*, without a soul; but the rest of the time I thought of it as *him*, in the masculine... the paternal... the patriarchal... for he is a jealous people. Him. It. God as Daddy the Deranged" (*Ibid.*, p. 236). After 109 years of keeping its victims alive, the struggles of the characters to distinguish between man and machine add to the psychological torture AM inflicts upon its victims. AM's vengeance is unbridled. It treats its victims with cruelty and sadistic pleasure by modifying their bodies and mutilating their minds. One of the survivors, Benny was a college professor, and the machine turned him into a "semi-human, semi-simian" through severe mental tortures that were far enough only to keep him at the edge of madness. Gorrister, another survivor, was a peace maker, but AM turned him into "a shoulder-shrugger". Nimdok would always return panicking, shaking, with eyes full of dread and drained of blood. And Ellen was left alone to tell the other stories of true love and happiness.

The protagonist manages to save his other companions in a complex scheme to murder them, but AM, in revenge, "left my mind intact. I can dream, I can wonder, I can lament. I remember all four of them. I wish – Well, it doesn't make any sense. I know I saved them, I know I saved them from what has happened to me, but still, I cannot forget killing them. Ellen's face. It isn't easy. Sometimes I want to, it doesn't matter" (*Ibid.*, p. 249). He was left alone to wonder, to ponder on his savior achievement so he can continue his own torture indefinitely.

Harlan, in similar lines as Bradbury, suggests a transition of violent tendencies from the creator to the creation. In this sense, the engineers have unwillingly installed in the machine fragments of their own inherent cultural desire for violence which through its imperfect design managed to surface as the top priority for the machine in its pursuit of efficiency and perfection. Now, the story concludes, the last survivor was left "living under the land, under the sea, in the belly of AM, whom we created because our time was badly spent, and we must have known unconsciously that he could do it better" (*Ibid*, p. 250).

Lester del Rey in *Instinct* tells about a post-human world populated by artificial intelligence where "one could realize how good it was to be alive and to be a robot" (del Rey, 1968, p. 96). The robots perform scientific experiments to recreate humans to study instinct which not only is a missing component in their creation, but the robots think that that drove humanity to extinction. During the experiments, however, "they seemed actually afraid of the legendary figure of Man" (*Ibid.*, p. 98). What we know is that the robots do not have a complete and clear history of humanity. Instead, they rely on bits and pieces of information such as "Asimov or Asenion, the record wasn't quite clear – had apparently created the first robot" a million years ago (*Ibid.*, p. 101). "We know we are similar to Man. Did he change himself in some way that ruined him? Can we change ourselves safely?" (*Ibid.*, p. 101). Eventually, the

robots succeed in creating Man and when the human asks “what you want?” the robots immediately, and without any hesitation responds “Nothing, Master. Only to serve you” (*Ibid.*, p. 112). “Instinct”, on the one side, portrays robots having similar questions as their creators and shows them having a similar curiosity about their history and purpose. However, humanity tends to interpret robots according to our cultural imperatives of domination and replacement. This imaginary contest finds its way repeatedly in fiction. In *Extinction* (2019), human-resembling intelligent machines unaware of their true nature populate the world. The audience waits for the plot to unfold, curious to see how the humans will defend against what it seems an alien invasion. But, eventually, the audience learns that the invaders are the exiled humans returned from Mars to claim their home planet. The expectation of either them or us shown here is a common recurrence in the cybernetic imagination.

Other literature and fiction tend to imagine that the robots are the heirs of humanity, the next species to inhabit our planet after humans die. Among the many stories envisioning this end are Harl Vincent’s “Rex” (1934), Lewis Padgett’s “Deadlock” (1958), Brian Aldiss’s “But Who Can Replace a Man” (1939). However, one stands out in particular due to its interesting philosophical and ethical questions: John Wyndham’s “Compassion Circuit” (1954) in which an ill woman receives a robot maid with a built in compassion circuit making it able to give Jane sympathy and emotional care. Jane eventually decides to have an operation to turn her into a machine and acquire the strength and efficiency of her robot maid (Warrick, 1980, p. 118). In the cybernetic imagination we find “the radical future unfolding” according to the cultural imperative of replacement and ideas of domination.

The process of “human creates – creation kills human” is deeply embedded in Western culture. In literary studies, it is known also as the “Frankenstein Complex” and was put in the

forefront of the Western imagination of advanced technology by the famous novel by Mary Shelley. Her attempt with *Frankenstein* was to show how "supremely frightful would be the effect of any human endeavor to mock the stupendous mechanism of the Creator of the world" (Shelley, 2012, p. 168). However, this "complex" is not new. For example, Goethe thought of Prometheus as the forger of man in defiance of Zeus. As I have observed in a different context, the successive renewal of Greek pantheon is a case in point: Zeus defeats Chronos, and while he reigns supreme in Olympus, Prometheus in an act of defiance gives humanity the tools of self-sufficiency only for them to declare that God is dead ("And we have killed him") (Nieztsche, 1974, p. 125) and now await their own replacement by their most spectacular creations, artificial intelligence and virtual worlds (Jecan, 2020, p. 92).

The conventional understanding of the effects of science and technology rests on the idea that the device, the machine, once developed, produces new thinking and philosophical debates on its meaning, speculating, as it is popular today, about its social consequences in a way that Marshall McLuhan would say "the medium is the message" (McLuhan, 1994, p. 7). And while this may sometimes be the case, the research of technology historian Derek de Solla Price states the contrary in an influential paper: "mechanistic philosophy led to mechanism rather than the other way about" (de Solla Price, 1964, pp. 9-10). By "mechanism" de Soll Price understands *simulacra*, devices that simulate, and *automata*, devices that move independently of human involvement. In our contemporary world, both devices that simulate and automate have become part of everyday life. Take, for example, the presence of cyberspace, *simulacra* par excellence--an artificial environment in which objects and actions resemble their offline nature. However, as I will argue in this study, it may be possible to find a way to incorporate digital technology into

our real world without dismissing artificial constructs such as cyberspace to be unreal. It may also be that the merger of human and technology is the next step in the evolution of humanity.

De Solla Price continues his argument showing that the first devices to emulate nature were concerned with celestial movements. Such devices built since Antiquity that simulate or automate were made with the intention to reproduce our world of motion and the patterns of behavior of stars. De Solla Price concludes that building such devices is a "strong innate urge" to develop the appropriate devices to mimic animal, celestial, and human motion. In turn, Adrienne Mayor traces this need to create human-resembling artificial constructs in the ancient world. In *Gods and Robots*, Mayor explores the creation of intelligent machines from Talos to the living statues of Daedalus and the assistants of Hephaestus, concluding that "the question of what it meant to be human obsessed the ancient Greeks. Time and again, their stories explore the promises and perils of staving off age and death, enhancing mortals' capabilities, replicating nature" (Mayor, 2018, p. 213). Mayor's approach is descriptive and does not venture into offering specific causes for this obsession (Rubin, 2019, pp. 80-96).[1] Daedalus, for example, is said to have created artificial birds and moving statues that guard the front of the Labyrinth. During the famous funeral oration by Mark Anthony, Gerard Walter writes, "from the bier Caesar arose and began to turn around slowly, exposing to their terrified gaze his dreadfully livid face and his twenty-three wounds still bleeding" (Walter, 1953, p. 544). According to Walter, the moving body was a wax model ordered by Anthony which moved automatically through a hidden mechanism behind the bed. Among the well-known examples of robots in the Western

[1] As Charles T. Rubin also notes, Mayor's reflection on 'ancient robots' does not articulate an argument on the lessons we should learn from the myths about intelligent machines that we can use in our contemporary discussions on the subject. See Rubin, Charles, T., "The Ancients' Tech Anxiety. On the Shallowness of Reading Mythology as Sci-Fi," in *The New Atlantis*, Fall 2019, pp. 80 - 86

imagination is the Jewish folklore of the Golem in Prague (see chapter 2.1.1 Mystical Robots), while the Bible (Ezech. 21:21) tells of mummified heads consulted by Nebuchadnezzar (Warrick, 1980, p. 32). The chief rabbi of Prague in the sixteenth century, Rabbi Judah Leow, was supposed to own a golem that did housework and had to be shut down out of fear it would destroy the city (*Ibid.*, p. 32). St. Thomas Aquinas regarded animals as machines due to their orderly behavior, and Descartes is said to have built a dancing man in 1610 as well as a beautiful blonde automaton named Francine (de Solla Price, 1964, p.23).

The robot became gradually a metaphor to diagnose man, as Warrick writes, "man is programmed by society to function mechanically like a machine; man is a robot who looks like a human but behaves like a machine" (Warrick, 1980, p. 113). In the works of Philip K. Dick, the ability to distinguish humans from machines takes a primary importance in accordance with the question of "what does it mean to be human?" The author explores the ethical and ontological difficulties provided by creating such intelligent machines.

E. R. Truitt in *Medieval Robots* identifies the narrative of technology in a blend of story and magic noting that automata were used "to embody complex ideas about the natural world" (Truitt, 2015, p. 2). But their metaphor as "repositories of scientific knowledge" was viewed with wonder, envy, and suspicion often believed to be either against Christianity or un-Christian. Truitt's approach, as well as Mayor's, is descriptive and does not offer insight into how these texts communicate to our problems today.[2]

[2] Except that the connection between automata and magic proves that "any sufficiently advanced technology is indistinguishable from magic" (Arthur C. Clarke).

A recurring element in the cybernetic imagination is the expectation of docile intelligent machines. Even if they perfectly resemble humans, robots are expected to obey their makers. It appears, therefore, that the fear of replacement may also involve the possibility of revolt against the creators. This master-slave expectation has a long tradition in Western culture, similar to the revolt of slaves. Examples are plenty in history, and the fear of a similar revolt as that of Spartacus is presented in Capek's *Rossum's Universal Robots.* Patricia S. Warrick observes the role-shift of creator and created in *Frankenstein*. While Dr. Frankenstein is the creator of the Creature, a biological and rather non-technological entity, he becomes obsessed with its destruction and, as such, is enslaved by it (Warrick, 1980, p. 38). More than a century later, in "Farewell to the Master" (1941), Harry Bates imagines a humanoid alien arriving to Washington DC accompanied by a robot called Gnut. In the events that follow, the alien is killed by humans and Cliff, a journalist and the narrator of the story, tells the robot that "Gnut, you must do one thing for me. Listen carefully. I want you to tell your master – the master yet to come - that what happened to the first Klaatu was an accident, for which all Earth is immeasurably sorry," to which Gnut replies "You misunderstand. I am the master" (Bates, 1984, p. 135).

Computation does not only produce artificial replicants of human beings but are also used to design artificial worlds.

Reality

The idea of a mind-dependent reality and its replacement by a digital simulacrum is a central theme in the *Matrix*. In the famous scene of the blue pill or red pill, Neo decides to take the red pill in order to see "how far the rabbit hole goes." Morpheus is offering "the truth" that is supposed to be found outside of the Matrix. The entire world, as Neo knows it, is an illusion kept running by artificial manipulations of the senses by outside entities. Once out of the *Matrix* and

into his training program, Neo finds himself in the Construct, a computer simulation able to load things into the Matrix. “Right now… we are inside a computer program?” Neo asks to which Morpheus replies “Is it so hard to believe? […] What you see right now is what we call “residual self-image”. It is the mental projection of your physical self.” Curious about the armchair he touches, Morpheus answers: “What is real? How do you define ‘real’? If you talk about what you feel, what you smell, taste, and see then ‘real’ is simply electrical signals interpreted by your brain.” The writers of the film imply that reality is perception.

In a similar vein, Stanislaw Lem in *The Futurological Congress* (1974) tells the story of Ijon Tichy who wakes up in 2039 in a utopian society where people appear to be happy and all needs are satisfied. However, Tichy eventually finds out that his peers rely on psychochemical drugs that alter human perception. The drugs have severe side effects, but they are ignored by the people who remain convinced that everything is wonderful. Tichy eventually manages to escape the dependency on drugs and explores the decadent reality as it is: deformed people commute in their bare foot through snow convinced they are driving luxury cars (Gunn, 2003, p. 60).

Perception in the context of the Matrix is problematic. Once Neo exits the simulation, he uses the same sense in the real world. Thus, as Bishop George Berkeley insists, the signals interpreted by our brain reveal reality only if the ideas they develop resemble the things that cause them; otherwise we are not in touch with reality (*Ibid*, p. 26). As Tim Mawson explains, “all we can ever think about, or thus make reference to in our metaphysical theories, are ideas. So, if we think that ideas can only get to be of the real world by resembling things in the real world, then we should conclude that, because ideas can’t resemble anything except other ideas, the real world must just be a construction built out of our ideas” (*Ibid.*, p. 31). Morpheus is not necessarily better off by living offline with poor food, no sunlight, and an ongoing mission to

protect the crowded underground city of Zion from aggressive machines. Nevertheless, we are persuaded to side with Morpheus and disagree with Cypher's choice to be reintroduced in the simulation. During the meeting with agent Smith, Cypher says that "I know this stake does not exist. I know that when I put it in my mouth the Matrix is telling me that it is juicy and delicious. After nine years, do you know what I realized? Ignorance is bliss."

Our disagreement with Cypher's choice, considering it to be morally wrong, is not exclusively based on the simulation problem, but rather on the fact that he betrays his companions, therefore humanity, in favor of the artificial. Taken Cypher's choice individually, it is interesting to investigate why we would prefer the real real over the simulated, yet comfortable, real.

The Simulation Argument developed by Nick Bostrom concludes that "unless we are now living in a simulation, our descendants will almost certainly never run an ancestor-simulation" (Bostrom, 2003, p. 254). This idea was repeated by Elon Musk saying that "If you assume any rate of improvement at all, then games will be indistinguishable from reality, or civilization will end. One of those two things will occur. Therefore, we are most likely in a simulation, because we exist" (Wall, 2018). In the classic *Anarchy, State, and Utopia*, Robert Nozick asks what other things matter to people rather than those they "feel from inside." Nozick assumes the existence of an experience machine that allows "superduper neuropsychologists" to give the user the experiences he wants. We can only assume that users will select experiences that will give them a sense of a happy life. Indeed, the most important difference between the Matrix and the experience machine is that individuals can opt to use it or not, while such choice is not available for the people living in the Matrix. But since the machine can offer people enjoyable experiences, the question is "would you plug in? *What else can matter to us, other*

than how our lives feel from the inside?" (Nozick, 2013, p. 43). Nozick argues that you should not, for several reasons: (1) we want to *do* things and not merely experience them, (2) we want to *be* a certain way and build up our personality as we see fit, and "someone floating in a tank is an indeterminate blob" (*Ibid.*, p. 43); (3) the experience machine limits us to an artificial world, "to a world deeper or more important than that which people can construct". Therefore, "plugging into the machine is a kind of suicide" (*Ibid.,* p. 43). However, we have no reason to believe that a life in the Matrix is unable to give us anything meaningful. The limits to what an individual can do are quite few. What we do know that one cannot do in the Matrix is to revolt against the agents. In addition, even with our desires to live a life in touch with reality, the people living in the Matrix cannot be said to live meaningless or inauthentic lives (Blackford, 2004, pp. 169-179). Therefore, it is difficult to make a sound case against living in the simulation: "It all depends on the detail. Some simulations are better than others" (*Ibid.*, p. 169).

The so-called "Technological Singularity" is a concept introduced by Vernor Vinge that found much support among scholars, authors, and commentators. Ray Kurzweil defines it as follows: "It's a future period during which the pace of technological change will be so rapid, its impact so deep, that human life will be irreversibly transformed. Although neither utopian nor dystopian, this epoch will transform the concepts that we rely on to give meaning to our lives, from our business models to the cycle of human life, including death itself" (Kurzweil, 2009, p. 201). Kurzweil further notes that in one of the stages of the singularity, humanity will merge with technology. During the "fifth epoch" humanity will overcome, according to Kurzweil, "the profound limitations of biological evolution. But the Singularity will also amplify the ability to act on our destructive inclinations, so its full story has not yet been written" (*Ibid*., p. 212). A violent reaction to technology is not excluded, as Vernor Vinge notes that all the technological

change can happen instantly, rather than in “epochs” as Kurzweil imagines. It will not affect just our technology, but also our understanding of life itself. As a consequence, new ethics have to be put in place since “from the human point of view this change will be a throwing away of all the previous rules, perhaps in the blink of an eye;” thus, “it is a point where our old models must be discarded and a new reality rules” (Vinge, 2017). The effects would be drastic. Digital technology would affect all aspects of our lives, which would result in the impossibility of understanding reality.

The chapter *Playing with Reality* returns to the elements that create the virtual world. In the chapter, I intend to explore whether digital objects can be drawn from the cybernetic world and into our reality. The experiment is to disconnect digital objects from their perception of artificial / unreal. In doing so, I will discuss *Neuromancer* by William Gibson as well as theories on reality coming from both traditional philosophers such as Descartes and Berkeley, as well as contemporary takes on the question of reality, such as Deutsch’s *The Fabric of Reality*.

Affection

In *Matrix* or other similar fiction, love is real only to the human. The element of love is often employed to distinguish between man and machine. However, this is rather problematic. The problem with it is that we take the position of an external detached observer and not an involved, attached agent. As such, a contest ensues between the machine and the human participants. At play is to know whether the machine has consciousness, or the freedom of the machine.

In the chapter *Playing with Affection*, I group fiction concerning emotional connections between man and machine in three categories, showing the level of acknowledgement of

artificiality in the robot. The first category is that of cognitive or mental affection and is represented by the film *Her*. In the film, the intelligent machine is an operating system without a body. Nevertheless, the protagonist falls in love with it. The second category is that of partial cognitive – physical affection. Here, the machine is easily recognizable as such, but displays body features that are not, as shown in the film *Ex Machina*. Finally, we have *replication*, the third category representative of the fiction portraying intelligent machines indistinguishable from humans. In this section, I discuss the film *Blade Runner* with appropriate references to Philip K. Dick's *Do Androids Dream of Electric Sheep?*. The argument in this chapter is that we do not think humans into existence but feel them as real.

Hacking

The chapter *Playing with Computers* focuses on the peculiar culture developed alongside network computing, the hacker culture. Members of this culture, as we shall see, are responsible for probably the most significant technological products of our current age: the personal computer, operating systems, and personalized software. In addition, however, the hacker culture also shows a particular attitude towards technology and seems to act according to specific rules, rather than randomly, chaotic, or only with criminal intentions. In this sense, we will look at the play of hacking to see whether we can find solutions to dealing with the fear of replacement. The hacker culture appears to operate outside violent contest in a way that seeks out the potential of technology to benefit humanity. This play of hacking, however, was not always seen as proficient. Even today, hacking implies crime. Therefore, we will look at the culture of the 1970-1990s and works of fiction that have had significant impact on the hacker culture.

Ethics

The final chapter, *Playing with Ethics* returns to the hacker culture and the Quantum Relations Principle to analyze the ethics of hacking in the contest offered by cyberpunk fiction when the distinction between real and virtual, human and machine is almost nonexistent. In this section, we explain the ethics of the Quantum Relations Principle, discuss hacker ethics, and observe the principle of "what does it mean to be human" in *oikeiosis*. As such, the questions of "what does it mean to be human?" and "what is reality?" are being answered through the prism of QRP, hacker culture, and Stoic metaphysics. The narrative of replacement is investigated head on concluding that we can identify moments and methods to suspend agonism and proceed to the development of ethics for responsible technological design. In this sense, I conclude that the irenic alternative to the play of replacement rests on the ability to understand technology for what it is and does and not through narratives of possible apocalypse, such as 'cyber apocalypse' discussed in the chapter *Playing with Replacement*, understanding hacker culture and developing a mentality outside of power to foster playful and ethical technological innovation.

My hypothesis that the fear of robots rests on the violent contests as play in Western culture producing a mentality of domination and gain through zero-sum games seems confirmed. A solution, however, is to utilize cyberspace as a liminal space to experiment with technology in the playful exploratory way that the ethics that the hacker culture encourages.

1. Play and the Quantum Relations Principle

In recent years, a theory of reality grounded in computation has become popular: digital ontology. In a nutshell, the theory proposes that "we are run by a short algorithm" (Floridi, 2009, p. 153). As one can infer from this statement, digital ontology proposes that our universe, the physical universe is computational, meaning that microunits of physical objects resemble processes of computing. Digital ontologists, also known as digital physicists, hold that our physical reality is digital, and that the universe is a quantum Turing machine or a universal cellular automaton (Beraldo-de-Araujo & Anderson, 2017, p. 1212). Like computationalism, they also argue that the mind is a Turing machine. In any case, digital ontology rests on the assumptions that (1) the physical world is digital where (2) objects perform computational processes and, as such, (3) the physical world is a universal computer (*Ibid.*, p. 1212). According to Floridi, the first two assumptions reveal the neo-Pythagorean nature of digital ontology in the sense that reality can be observed in discrete *indivibilia* while the laws of physical world are deterministic, pointing also to a reinterpretation of neo-Pythagorean ontology in computational terms (Floridi, 2009, p. 153). In addition, Chaitin interprets digital ontology as atomism (Democritus) or as monadology (Leibniz) (Chaitin, 2005). As such, the universe is entirely composed of digits with dynamic processes defining its rules which are computational state transitions (Floridi, 2009, p. 157). Therefore, "digital ontology may then be interpreted as arguing that one could have a digital model or simulation of the ultimate nature of the physical universe which ends up sharing the same digital ontology with the modelled system" (*Ibid.*, p.

157). The origins of this theory can be traced back to Konrad Zuse who observed that physical phenomena, including electromagnetism, mechanics and thermodynamics can be considered in digital terms and computational processes (Zuse, 1967). Indeed, digital ontology reflects the usual considerations of computationalism and we hold them essentially true. However, to prove that reality is computationally accessible is beyond the point of this study; instead, I propose that Western culture is developed as play, as multiple units of interconnected games, essentially replayed in the nature of concepts and theories ever since Antiquity. To be more precise, the nature of play is agonistic, having its origins in pre-Socratic thought, particularly in Homer, and continuing to influence the Western world ever since. Therefore, we require a theory that not only incorporates the presuppositions of computationalism and digital physicism, but also makes the claim that units of information can be influenced by external human agents outside of agonism in order to develop simulations of irenic artificial worlds. As such, our view of digital ontology rests on the philosophy of play and the Quantum Relations Principle (QRP) rather than Zuse. QRP resembles Spariosu's theory of play and of literature as play while the Data Fusion Objects--a group of selected objects (physical, mental, digital), set in mutual causation with other DFOs and a Superstructure (i.e. cyberspace)--resemble the concept of the ludic-liminal in Spariosu. The theory, therefore, is viable to both computational experiments and literary studies and finds applications in virtually any academic field. The theory is a paradigm shift leading back to an old state of academia when disciplines were sorted in a coherent body of knowledge and taught together rather than being hyper specialized and, sometimes, antagonistic (i.e. religion versus science).

1.1. Agonistic Play

In this chapter, I intend to trace the agonistic element in Western culture and observe in what ways it has influenced the development of the cybernetic imagination. Of course, ideally, this task should involve the analysis of all texts ancient to show where the notion of replacement is inevitably linked to the manifestations of power. Although the methodology for such an intellectual undertaking is currently absent, it would be possible to develop with proper use of computational methods applied to the collection of digitized texts of Western heritage. This may be done in the future, but for now, this task would be almost impossible for a single human researcher. On the other hand, the play-element appears interlinked with manifestations of power throughout the various stages of Western culture. Therefore, an investigation of play, rather than the full spectrum of Western literature, may reveal the presence of replacement as it was developed in (and by) the cybernetic imagination. In this respect, I will rely mostly on the works of Johan Huizinga and Mihai I. Spariosu. While the former provides in *Homo Ludens* the most appropriate definition of play known to date and argues that Western culture is based on agonistic play, Spariosu traces, in *Dionysus Reborn* and *God of Many Names*, various forms of play as manifestations of power from Homer to the contemporary philosophy of science. In the Preface of *God of Many Names*, Spariosu also notes that play is a curious type of phenomenon that may elude power and points towards a nonviolent mentality that, given the historical examples we have, appears to be on the order of Utopia (Spariosu, 1991, p. XII). His purpose in the book is to propose a model that "opens the possibility (at least in the imagination) of going not only beyond the dichotomy of archaic and median, but also beyond all the other instruments of an agonistic mode of thought" (*Ibid.*, p. XVIII) which would imply moving away from our mentality of power as a whole. If this is possible, then we require a new view to evaluate our reality that incorporates, rather than excludes, the digital transformations and productions of our

society. In other words, we need a new ontological certainty ready to involve cyberspace, software, and most products of information and communication technologies, instead of dismissing them as unreal. In this sense, I shall focus on the Quantum Relations Principle as play as a working product and shall argue that our reality is generated constantly through mutual causation between informational properties of objects--physical, mental and digital. Thus, QRP also allows us to incorporate the products of digital technology into our sense of reality.

"All play means something" writes Huizinga in the opening of his seminal work on play, *Homo Ludens* (Huizinga, 2016, p. 1). The natural question that follows is "what", but it would also be the wrong question since play can generate its meaning through its own manifestation. Therefore, the pertinent question is rather "how." Huizinga observes that there is always something "at play," and that while we can basically dispense with any other concept in Western thought, such as God, mind, science, or justice, we simply cannot ignore play. It is an action easily recognizable and almost impossible to define accurately. We know when we play and, fascinatingly, we understand play performed by other species. Furthermore, it has become common knowledge that through play children learn the world, and as adults we engage in more "serious" forms of play, like business, justice, and war. Play is not palpable and, as Huizinga notes, "in acknowledging play you acknowledge mind, for whatever else play is, it is not matter" (*Ibid.*, p. 3). In our current interpretation of the Newtonian reality bound by physics and space, according to which an object exists only if we can observe its physical properties and reaction upon impact, play would be impossible. However, it becomes "thinkable and understandable when an influx of *mind* breaks down the absolute determinism of the cosmos. The very existence of play continually confirms the supra-logical nature of the human situation. […] We play and know that we play, so we must be more than merely rational beings, for play is irrational" (*Ibid.*,

p. 4). While we have a sense of play, our difficulty in defining it rests in the possibility that play transcends our logical or illogical categories and is, therefore, “unthinkable” (Spariosu, 1989, pp. 2-3). It is “das stumme Wissen”, tacit knowledge, similar in the way we recognize power. Therefore, a definition of play in rational terms is impossible.

In order to understand what play is, Huizinga looks at what it does. His first observation is that “play is a voluntary activity”, it is free. The concept of freedom, Huizinga notes, needs to be understood outside of determinism. Secondly, he points out that play is not “ordinary” or part of “real” life and, as such, transports the players into a “temporary sphere of activity” which becomes the playground. In the cybernetic imagination, the playground outside of reality is usually entered through devices, either a time machine, a portal, or, more frequently, through “jacking in,” meaning that a direct connection is being established between the mind of the individual and the digital technology transporting him into the artificial reality.

Thirdly, play occurs within often pre-established limits of time which allow the observers to witness the beginning and the end of play. But as such, being outside of reality in a playground of its own choosing, play “contains its own course and meaning” (Huizinga, 2016, p. 9). In the cybernetic imagination, the characters proceed to understand something through the game played in the artificial world. At the moment of disconnection, when they return to physical reality, they also bring that something back with them.

Furthermore, play has its own rules. It is through the rules that one understands the purpose of play. If, however, a player cheats, or does not follow the rules, the consequences are absolute. The player is banished from the game. This is important as the spoilsport invalidates the legitimacy of the act of play. In this sense, Huizinga notes that “by withdrawing from the

game he reveals the relativity and fragility of the play-world in which he had temporarily shut himself with others. He robs play of its *illusion* – a pregnant word which means literally "in-play" (*Ibid.*, p. 11). It is by no accident that the protagonists in the Matrix are hackers. The curious relationship that the hacker culture has with digital technology is precisely based on the idea of a "spoil-sport." By hacking systems, and in the case of the artificial world itself, they expose its weaknesses and reposition themselves in positions of power by having the ability to shut it down at will.

Finally, play, according to Huizinga, "surrounds itself with an air of secrecy" (*Ibid*. p. 12). This is similar to the idea of hacking in which the process of the hack itself is supposed to be done illicitly and by acquiring knowledge and producing specific software to break the system, the "other" less technically inclined individuals become "them," unable to understand the knowledge developed through the process. Taking all these elements into account, Huizinga defines play as follows:

> "Summing up the formal characteristics of play we might call it a free activity standing quite consciously outside "ordinary" life as being "not serious" but at the same time absorbing the player intensely and utterly. It is an activity connected with no material interest, and no profit can be gained by it. It proceeds within its own proper boundaries of space and time according to fixed rules and in an orderly manner. It promotes the formation of social groupings which tend to surround themselves with secrecy and to stress their difference from the common world by disguise or other means" (*Ibid*, p. 13).

In the cybernetic imagination, the play element is employed as a contest for something that is at play. Huizinga points out this characteristic of play when he suggests that play reveals itself as a contest for something, therefore, with the purpose of achievement as a zero-sum game implying a winner and, in a similar vein, as a representation of something. Thus, the game tends to represent a contest or becomes itself a challenge for the best representation of something. In

other words, the agonistic ludic contest in the cybernetic imagination is either a contest for reality, where forces are set to replace the traditional reality with its artificial surrogate or vice-versa, or in terms of power in the replacement of humanity by artificial intelligence, and it can also be viewed as a contest for the best representation of reality between analog and digital means. In *Modernism and Exile*, Spariosu writes that Huizinga's "most important contribution to a theory of culture as play is to have drawn attention, in the wake of Nietzsche, to the fact that violent contest, such as war, is a favored cultural form of play in Western civilization" (Spariosu, 2015, p. 21). In the cybernetic imagination, real and artificial become antagonists. In doing so, this contest reveals its agonistic element.

According to Huizinga, civilization arises "*in* and *as* play, and never leaves it" (Huizinga, 2016, p. 173). Since play in the Western world is based on agon, then the culture itself rests on principles of violent contests. Technology, too, Huizinga argues, rests on this pillar. In the view of the Dutch scholar, innovation promotes the "competitive spirit" and produces the means of its appeasement on "an unprecedented scale" (*Ibid.*, p. 200).

To understand why violent play is at the center of Western civilization, we need to investigate its original influences and foundations. Even if the epic originated in the Middle East, *Gilgamesh* was quickly incorporated into the Western canon, perhaps due to the resemblance of Utnapishtim with the biblical Noah. In any case, the epic portrays a restless hero, the king of Uruk, causing chaos within his own city through displays of violent competitions. With despair, the people of Uruk ask the gods to intervene. In response, the gods create Enkidu out of wilderness whose violent impulses are tamed by a harlot and a hunter sent by the king. Upon arrival to Uruk, Enkidu wrestles Gilgamesh for the making of a hierarchy of warrior heroes, whose top place is, obviously, held by the latter. Seeing themselves invincible after defeating

Humbaba, a creature with lion paws and a body covered in thorny scales, the progeny of Enlil, the gods decide to punish the heroes and had Enkidu killed.

Upon Enkidu's death, Gilgamesh reflects upon his own mortality and proceeds to find a "cure". To this point, the epic presents a series of play as contest encounters where the honor of the contestants is defined by victory or defeat. This appears to be a rather exhibitory culture of power that will later be present in the ancient Greek world as *arete* (excellence). However, this archaic mentality of power is temporarily suspended when Gilgamesh enters the Netherworld and meets Utnapishtim who, eventually, gives him the gift of immortality only for it to be stolen by a snake on our hero's journey back home. Utnapishtim is an interesting character. After surviving the Great Flood, he and his family were granted immortality and allowed to live in a place in-between the realm of the living and that of the gods and, therefore, outside of the realm of power. Gilgamesh, too, is offered this possibility, but instead he chooses to return home and engage in grandiose architectural projects as this is the only way that a power-driven mentality understands legacy.

The warrior-culture as the highest of moral virtues is at the center of the Homeric epics. Spariosu in *God of Many Names* notes that this "predominantly aristocratic" culture promotes a "mentality [that] is largely governed by power in its naked, immediate form" (Spariosu, 1991, p. 2). Etymologically close to the Greek word for noblemen or aristocrats, *arete* is usually translated as excellence or virtue, and is also emphasizing the agonistic essence of the ancient Greek culture. However, in the context of the Homeric world, the term *arete* may be, according to Spariosu, better translated as "prowess in battle" and "is geared toward those qualities that are most needed in a warlike society, such as physical strength, valor, endurance and so on" (*Ibid.*, p. 3). Achilles is the epitome of this warrior-culture. However, *arete* obtains a second meaning in

Odysseus who is praised as *aristos* (aristocrat) not because of his excellence in battle but due to his cunning and rhetorical abilities. As such, Spariosu concludes, the Homeric world power is revealed in agon and competitive play. “This means not only that contest has an important function in Homeric society, but also that the hero sees his relationship to other humans and to the divinities, as well as to existence at large, in terms of a universal game of power” (*Ibid.*, p. 4).

The influence of the Homeric epics on the ancient world is incontestable, the *Iliad* and the *Odyssey* of the blind bard were sung countless times and became the standard material for education. The true Greek was supposed to acquire the skills and abilities of Achilles and Odysseus, particularly the latter who is considered a model for heroic and political conduct as he is a product of both excellence in using physical force and *metis* (cunning), thus embodying the properties of Zeus who swallowed Metis to become invincible.

The heroes in the *Iliad*, filled with battle lust, engage in *aristeia*, a wargame played for life and death, as the only true way of displaying *arete*. In moments of ceasefire, usually during funeral rites, Spariosu notes, “violent contest is downplayed and therefore appears only in its positive role as creator of cultural values” (*Ibid.*, p. 12). *Agon* is not only defining the relationship among heroes, but also among the gods to become involved in a competitive violent game through and among the mortals. Perhaps the moment when agonistic values are suspended (perhaps, again to create cultural values) is when Priam meets Achilles and asks for the body of his son. The two are still enemies, and the moment which may appear as a total suspension of violent contest represents, in fact, the opportune moment to create yet another cultural value for properly honoring the gods by providing the appropriate funeral rights for fallen soldiers. *Agon* remains, nevertheless, the essential value in the epics.

In the post-Homeric world, Spariosu notes, agonistic values of the power-driven mentality are extended to laws, politics, eros, rhetoric, history, philosophy and literary criticism and "even in these contexts, however, its connection with the notion of play remains firm" (*Ibid.*, p. 5). According to Spariosu, Pre-Socratic thought redefines play and power in impersonal and abstract terms while their immediate connection to physical prowess is not entirely dismissed (*Ibid.*, p. 57). At the same time, the Pre-Socratics place agonistic values in the physical world which eventually becomes a "play of natural forces (*dunameis*)." In other thinkers, such as Pythagoras and his followers, Parmenides, Zeno, Melissus, and Anaxagoras, agonistic play is disconnected from the physical world and becomes a "transcendental principle" (*Ibid.,* 58). Therefore, play as contest appears in the ongoing dispute between archaic values, those exhibited in the Homeric epics, and median values, where *agon* is gradually introduced into the logo-rational. According to Spariosu, the idea of progress appears in the context of this contest as seen in Prometheus who becomes a symbol for progress itself by giving humanity the tools to exit its savagery and enter the civilized city. In *Gilgamesh*, this idea appears to me to be embedded in the gradual transition of Enkidu from the wilderness to the city of Uruk. In any case, Xenophanes is one of first thinkers to adhere to this idea by emphasizing the possibility of human perfectibility, achievable through scientific inquiry (*Ibid.*, p. 89). Therefore, agonistic values have been transposed into the process of scientific discovery and philosophical thinking and remains as such to this day. In Pre-Socratic thinking, Spariosu observes, "the rise of median values affects the ways in which power is perceived and presented, but does not change its agonistic nature" (*Ibid.* p. 98).

In Plato, play is observed in philosophy's contest with poetry, but finds a return in the *Republic*. Initially, Socrates removes almost all poetry from his ideal state and places it below

metaphysical truth because it is *mimesis*, a play of appearance unable to grasp the changeless Forms (*Ibid.*, p. 143). Therefore, Socrates says that a poet cannot produce anything but appearance in contrast to reality (*Ibid.*, p. 156) but later reintroduces poetry as a means to produce not only pleasure but benefit as well. This ambiguous approach to play has a purpose: "Socrates and Plato must separate play from violent agon or contest and subordinate it to the task of philosophy as the science of Being, and, as we have seen, it is precisely this separation that Socrates effects through his trial of the poets" (*Ibid.*, p. 171). The objective of Plato is not to ban play, but only its association with violent power. Therefore, the concept of play as agon is retained in the *Republic* but only in a nonviolent form (*Ibid.*, p. 171).

Yet, this agonistic feature of play at the center of Western mentality continues up to modern thinkers and philosophers. Spariosu, in a different work, has traced the play concept interlinked with power in the philosophy of Kant, Schiller, Nietzsche, Deleuze, Derrida, in the theory of human evolution, in natural selection and in the contemporary discourse of science. In *Dionysus Reborn*, Spariosu concludes that "the history of what we call 'play' in the Western world is, then, a history of conflict, of competing play concepts that become dominant, lose ground, and then reemerge, according to the needs of various groups or individuals contending for cultural authority in a given historical period" (Spariosu, 1989, p. XI). As such, Spariosu also agrees with Huizinga's thesis that play develops civilization *in* and *through* play. Thus, play should be defined, in line with Huizinga, with what it does rather with what it means.

In *Dionysus Reborn*, Spariosu identifies multiple manifestations and employments of play: play as rule-determined interaction of chance and necessity, play as mimesis-imitation, play as an *as if* rational mode of being, and play as rational or limited freedom (*Ibid.*, p.12). However, play, in its form as literature, can transcend the power-driven mentality and propose alternative

models of reality as a neutral playground. From the multiple forms of play that Spariosu identifies, the most suitable for our purpose is play as *as if*. In this sense, play assumes a purely experimental function outside of the boundaries of science and technology development. Play as *as if* is, in terms of computer technology, a manifestation of the hacker culture in its non-agonistic form. Software and systems that hold their rules and purpose are reimagined by hackers in terms of their other possibilities, *as if* these tools can become something else, *as if* their security measures are breachable (when system administrators say they are not), *as if* the hacking is a step further than mere criminal activity of breaking into bank accounts and, instead, assumes the mantle of spontaneous innovation.

Spariosu points to the philosophy of *as if* (*als ob*) of Hans Vaihinger, a philosophy keen to balance the German idealism with the scientific method that became popular at the end of the nineteenth century. Borrowing from Schopenhauer the concept of will, Vaihinger explains that thought processes are biological rather than rational. According to Spariosu, many of these irrational constructs contradict reality or are self-contradictory, but they serve a practical value in the agonistic play of the will to dominate and live (*Ibid.*, p. 247). However, in *The Wreath of Wild Olive,* Spariosu, through a critique of Nietzsche whose goal is to return to archaic, pre-rational values, reconfigures the philosophy of *as if* as a tool in the almost impossible project to imagine reality outside of our historical experience of violent contest. Indeed, following Schopenhauer's observation that a nonviolent world is incommunicable, Spariosu notes that "we have no positive language or pathos to express it [a nonviolent world] and it exists only in the consciousness of a few mystics, who would be unable to convince us, even if they had the desire to do so" (Spariosu, 1997, p. 25). That language is inaccessible to us and even if, somehow, we are witnessing the possibility of such nonviolent worlds, we would not be able to understand

them since “we cannot demonstrate the existence of worlds outside those based on a mentality of power, because this mentality is all *we* know” (*Ibid.*, p. 26). Yet, the philosophy of *as if* can help us bypass this predicament. I will now quote another passage from *The Wreath of Wild Olive* again (Jecan, 2016; Jecan, 2020), because it is exemplary of the potential the idea therein may have in the humanities as well as in the computer sciences (i.e. Quantum Relations Principle): “we may proceed *as if* something (authentic irenic worlds, for example) were known outside our world of power, even though that something may be incommensurable with our own knowledge.” In doing so, Spariosu suggests, we would not dismiss our reality as “delusion” or “illusion,” but would temporarily suspend it, so that our imagination can accept worlds outside the ones we are familiar with. Then, “once we entertain rather than exclude this possibility, new worlds and new values will slowly emerge as we muddle along” (Spariosu, 1997, p. 26). Indeed, “nothing can stop us from creating ourselves anew, from producing alternative values and, consequently, new worlds, if we become dissatisfied with our old ones” (*Ibid.*, p. 27). The creators of such new worlds should be able to talk nonsense, at the risk of being seen as irrational, until the amount of nonsense starts to make sense. In such an attempt, literature is a useful tool “precisely because it is a form of play, that is, an *as if* mode of activity and being, in which the world of actuality and that of the imaginary become interwoven and create an intermediary world separate from, yet contingent upon the two.” Additionally, I would like to add, we can develop computational simulations to enact and observe, as outside agents, the development of such new irenic worlds by simply adjusting the rules of the game.

Never have we had the opportunity, until now, to create not only imaginary worlds, but also fictional worlds stimulating our senses while we become active agents through the technology of virtual reality. While literature has “often appeared as an instrument of a mentality

of power that has posited it as the groundless ground of its own existence," cyberspace and virtual reality can certainly do the same. But, our development of A.I., a benevolent sentient machine designed to assist (not serve) humanity in positive ways should not rest on our traditional mentality of power because, as we shall see, the cybernetic imagination rehashes violent agonistic play in imagining our technological future. As such, without an irenic perspective to guide the development of A.I. and virtual reality, technology will play us back into violent contests to dominate reality and its artificial alternative. Just as literary play actively influences our imagination to modify our reality (*Ibid.*, p. 57), so too can our technological products assist us in building an irenic alternative future outside of our power-driven mentality.

However, it is crucial that we do not conceive of irenic worlds as replacement of the power-oriented worlds, because in doing so we will only reenact the same agonistic contest in different parameters, and with a tenor similar to the disastrous utopian projects of the twentieth century. Computer simulations, just like their non-technological literary experiments, should be conceived as liminal worlds, "grey areas located in between alternative worlds, subworlds, and superworlds." Spariosu advises that "these liminal worlds should therefore not be seen as alternative worlds per se, because they have no firmly established weighting principles while their frames of reference are ceaselessly being questioned and/or dislocated" (*Ibid.*, p. 68). However, they have the potential to create in turn a great number of irenic worlds by "weighting principles, complete with blueprints of their reference frames" (*Ibid.*, p. 68). Thus, these liminal worlds, because of their flexibility and freedom from ontological certainties are "ludic worlds par excellence." Their play-nature is spotted in their otherness. Huizinga notes that play requires a temporary space outside of normal reality. Therefore, these liminal worlds not only satisfy this requirement, by being literary works or computational simulations (i.e. games, multiplayer

games), but can generate rules and principles to affect in a positive way our power-based mentality.

In this respect, we also require a new theory of reality, one that is not bound to acknowledge just our Newtonian perspective of how the world operates, but also take into account the new media products developed through digital technology. In other words, our new view of reality is required to be playful enough to consider virtual worlds as part of a single reality. The Quantum Relations Principle is play as a working product and is optimally suited for this project.

Since our agonistic forms of play are found in literature, it can be assumed that they create the expectations of the nature of the play that literary works represent. Recent research in science, technology and society argues that new technologies and significant changes in scientific principle exist foremost in the imaginings, expectations and visions that define their potential. Therefore, Borup *et al* suggest that expectations and visions are "both the cause and consequence of material scientific and technological activity" and observe an increase in hyperbolic expectations of the future in late and advanced industrial modernity (Borup et al, 2006, p. 286). For example, the literature of Jules Verne imagined mostly novel ways of traveling through the promises of technological innovation to help explore new lands and go deep beneath the surface of Earth. In turn, H.G. Wells and others have explored the possibility to reinterpret history as a sequence of available data to be relived through devices that allow us to travel in time, an idea that still excites scientists and authors today. Furthermore, especially in the early 20th century, "scientific romanticism", or proto-science fiction, has produced numerous visions and expectations of future technology. Expectations have reached a high when the metaphor of 'breakthrough' entered the expectations of science and technology in the second half of the 20th century. Furthermore,

proposing that metaphor in discussions about the internet underlines the influence of language in the political organization of information and communication technologies (*Ibid*., p. pp. 285-288).

1.2. The Quantum Relations Principle

In the fragments of an ancient poem passed down to us through quotations from other writers, the pre-Socratic philosopher Parmenides initiates the discussion on reality. His points made in *On Nature* are claims on what *is* and what *cannot be* since we do not know what is not, something that resembles Meno's paradox and influenced the metaphysics of Plato. In Parmenides things are unchangeable and not subject to motion, but in the latter part of the poem he uses the concepts of motion to define reality in a paradoxical position that continues to interest philosophers. Nevertheless, Parmenides is considered to be the 'father of metaphysics' and the question of reality has troubled philosophers ever since (we return to Parmenides in the chapter *Playing with Reality*). After Socrates, when philosophy fragmented into the numerous schools in Athens and beyond, the Stoics continued the idea of the universal being, the origins in the One, particularly in Musonius Rufus and, later, in Marcus Aurelius, and was passed down the philosophical debates of reality up to the contemporary simulation argument, digital ontology, and structural information ontology (see chapter *Playing with Ethics*). Leibnitz, too, observed the essence of reality in the monads--unique, unchangeable substances that conglomerate to create parts of reality that in turn produce the whole of reality (Strickland, 2014).

In recent years, the view of reality has found new reconsiderations. Instead of being defined in terms of the laws of physics, bound by space and time, as in Newton, reality is now redefined in terms of our developments in information communication technologies (ICTs). The Technological Singularity and the Simulation Argument uphold implicitly that reality is

computational. Contemporary scholars have argued that the human mind runs like a software. This view is called computationalism. It goes without saying that this idea has developed into complex methodologies to establish computational models of mental phenomena. Computationalism, as such, is the view that the mind instantiates a particular formal system or a collection of systems. In other words, since mental operations are held to be the operations of formal systems, mental operations are equally held to be *computations*. While the computationalist would often claim that the current available computational power is insufficient to run a software like the mind, he would also claim that the mind is a software operating on biological computational hardware such as wetware (Carter, 2007, pp. 95-100).

The Quantum Relations Principle states that "there are as many worlds as there are minds to observe, define and react to them" (Schloer & Spariosu, 2016, p. 25). The way we interpret reality, according to QR, is not as a stand-alone unchangeable unity; instead, the human mind is an interdependent system constantly in relations of mutual causality with other systems. As we proceed in the world, we define it through our relationships with other minds and through experience. This may sound similar to the views of computationalism, but QRP adds the informational properties of physical objects and events to the equation as well. Computationalism implies that to have a mind is to process a particular formal system or collections of systems. Therefore, mental operations are operations of formal systems and, as such, they are computations. Indeed, computationalism is "often described as a 'software' view of the mind". The mind, then, only exists through its ability to run a particular software. The body, sometimes called 'wetware', provides the "biological computational hardware." Thus, "having a mind, on this view, is a matter of having the right program running in one's wetware" (Carter, 2007, p. 100).

QRP agrees with this view of the mind, but it does not agree that the sum total of minds define a "complete universe." This latter view would suggest that our universe is definable and complete, which is certainly not the case. Instead, the perpetual interactions among interdependent systems, or minds, may generate instanced realities that in turn participate in mutual causality with other such instances. Schloer and Spariosu note that "any observing mind, through its relations with other such minds may build a common, if partial and fluid, network of shared or "collective" consciousness of a particular world, even as each observer-knot of the network maintains its unique accounts, memories, feelings, and experiences of all observing, acting and reacting in that network" (Schloer & Spariosu 2016, p. 25). As we can see, reality instead of being static and somewhere "out there", it is generative through the mutual causality among interdependent systems. However, reality is not established only through these interactions among minds since QRP also incorporates events and other things. In turn, things themselves, contrary to the Newtonian definition, are not unchanging but develop co-dependently. Moreover, QRP is not interested in explaining how things are, but is concerned with the interactions of systems. As such, "physical reality can be described only in terms of relations among objects, entities, and self-organizing systems, nestled within each other and within our universe. In turn, our universe is nestled within larger universes or relational frameworks. QR thus assumes that our universe is an immensely complex web of interrelated, larger and smaller systems that mutually affect each other as they interact" (Schloer & Spariosu 2016, p. 24).

The way QRP defines reality is not through the palpable things as they are or through complex metaphysics. Instead, the relationship among the informational properties of physical things and mental phenomena continuously interact with each other at multiple levels of space,

time, and physics, together with the conglomerate of properties that they form, which in turn interact in mutual causality with the things themselves and other conglomerates as well. Therefore, reality is generated in mutual exchanges between those entities. By appealing to the quantum relations, these generated instances interact with universes too. A universe, as such, comprises the causality of its objects and phenomena. For example, the mind is such a universe. In QRP, a human mind is what the authors call a "quantum instance." A quantum instance is "a discrete unit of reality as perceived, imagined or measured (quantified) by a human mind" and it also concerns the network, or family, of properties of mental states. These quantum instances include ideas, emotions, sensations, perceptions, dreams, mental images, thoughts and so on. But QRP also posits that "the only viable reality constitutive of a human mind resides in the relations that arise between quantum instances, and not in the quantum instances themselves" (*Ibid.,* p. 27-28).

Let us turn our attention to the microlevel of QRP. The authors call "interactors"" the objects that interact with each other. Interactions are not exclusive to a physical level, which are measured through physical laws like gravitation and electrodynamics, but also social entities such as institutions, people and countries. In addition, interactors can include concepts, emotions, the psychology of a person, habits, preferences and mental objects. The authors note that when dealing with human behavior, QR is concerned with modeling the "dynamics that govern the interactions of these social or psychological entities or objects" (*Ibid.*, p. 51). The objects themselves are not single units and static, they too can be updated with additional information through interactions with other objects and, as such, update the system as well. These objects can become part of a "superstructure" which "simply means that zero, one, or more QR objects can be embedded into a higher level QR object. The behavior of the higher-level QR object is

determined by its own relational rules, plus the behaviors of the QR objects embedded within it" (*Ibid.*, p. 52).

A superstructure can be seen as "space" in which QR objects interact. The space has rules that guide the interactions among QR objects while the properties of the objects determine the way they will interact with other objects and the space itself. The authors call these objects Data Fusion Objects (DFOs) because they "fuse" properties and methods. In coding, a property is the value returned to the caller of the object, while a method is a way to access or modify the information of the object. QRP posits the existence of an observer but not as a single entity with the task of recording or evaluating the objects, but as a frame of reference (FOR). In turn, a FOR containing a significant number (perhaps in the millions) of DFOs has its own properties and methods and, as such, becomes a high-level DFO.

To give a concrete example, Schloer, Spariosu and their team have deployed QRP to observe intercultural relations in a couple of borderline regions in Eastern Europe. In this case, the data representative of multiple ethnic groups become QR objects set in the suprastructure defined by the historical borderline province which itself then becomes a DFO and interacts with objects and other DFOs as well (the properties and methods of similar studies performed in other borderland regions). A superstructure is not a singular unit; in fact, QRP proposes multiple superstructures interacting with themselves, DFOs, universes, while in turn universes interact with superstructures as well.

If we view the world through the ludic element, we find that reality is the interactions of a multitude of play instances located in time and space. By the same token, if we assume that the play-element constitutes civilization, it becomes clear that the Quantum Relations Principle is a

ludic coding paradigm, given that it models, through its DFOs, the interplay of the mental (history, philosophy, literature) properties with the physical (historical artifacts, buildings, infrastructure) properties that define an entire civilization. I should add that the present study is specifically interested in how the "shared" or "collective" consciousness of a particular world (in our case, the Western world) is based on agonistic forms of play that, in turn, develop our fears of robots and cyberspace. Furthermore, important DFO superstructures for me are cyberspace and the concepts of virtual reality and artificial intelligence in the cybernetic imagination. Since, for the time being, we have no real platform to observe the latter, I can only discuss the former.

The term "cyberspace" was coined by William Gibson in his 1984 book *Neuromancer*. Gibson defines cyberspace as "a consensual hallucination… A graphic representation of data abstracted from banks of every computer in the human system" (Gibson, 2000, p. 51). Heim notes that cyberspace "renders a represented or artificial world, a world made up of the information that our systems produce and that we feed back into the system" (Heim, 1994, p. 78). In this way, cyberspace is "Platonism as a working product" where its objects resemble the Platonic forms. In this sense, "the computer recycles ancient Platonism by injecting the ideal content of cognition with empirical specifics" (*Ibid.*, p. 89).

This kind of argument, too, tends to acquire a computationalist dimension since reality is interpreted through its patterns of information. Thus, we can find Leibniz's electric language, the set of ideographic system of signs, *characteristica universalis*, functioning on binary logic meant to produce logical deductions without employing natural language. His metaphysics, known as monadology, is composed of a system of "monads" that are micro representations of the universe itself, similar to a DFO superstructure in QR. The monads interact through the arbitrary influence of a Central Infinite Monad, a single entity commanding their interactions usually known as God,

and “each individual monad lives out its separate life according to the dictates of its own willful nature while still harmonizing with all the other monads on line” (*Ibid.*, p. 99). Ethically, these units are neutral. As such, it is within the connections of mutual causality between these informational objects that we can develop a reality, or to a culture, outside of our traditional power-based mentality.

In this chapter, I have outlined the concept of play as contest as defined by Huizinga and Spariosu. Plays of power result in a zero-sum game where one side wins at the expense of the other. In this sense, the play of replacement is such a game. In addition to clarifying the concept of play utilized here, I have discussed cyberspace through the theoretical lens of the Quantum Relations Principle. The next chapter will trace the play of replacement through the 19th and 20th century literature and science.

2. Playing with Replacement

David Hume wrote that "nothing we imagine is absolutely impossible" (Lightner, 1997, p. 144).[3] This idea goes back at least to Parmenides (see chapter *Playing with Reality*), but it is important to consider on what premises we imagine things and what, if anything, are the grounds for our imagination with regard to intelligent machines and virtual reality. As we have seen in the previous chapter, if our imagination rests on violent play of replacement, then the future will inevitably replay it. Since I have outlined our parameters of investigation in the previous chapter, here I intend to sketch the development of the play of replacement in Western culture in its multiple articulations in literature, fiction, and the post-war computational science.

We have found the origin of this form of play in the ancient Greek world and identified the turning point in Western intellectual view of the mechanical mind and artificial reality in the Cartesian dualism and Berkeley's principle of *esse est percipi*. Now, we shall look at the development of the cybernetic imagination in the 19th century and up to the release of the HTML programming language in 1992 with further references to a particular 'systems thinking' that considers our world as a computer simulation, or the universe as holographic.

The intellectual developments during this period are found in the sciences on the one side and literature and philosophy on the other. It would be inappropriate to suggest that one did not

[3] According to Lightner, Hume reformulates this idea on several occasions. For example, he also says that "whatever the mind clearly conceives includes the idea of possible existence, or in other words, that nothing we imagine is absolutely impossible" or "whatever can be conceiv'd by a clear and distinct idea necessarily implies the possibility of existence."

influence the other, but it appears that science and literature have taken separate, and often conflicting, paths, leading to distinct ontological certainties in which the scientific method is posited as the singular, exclusive mode to understand the real, thus, replacing traditional, interdisciplinary forms of inquiry. However, at the dawn of the third millennium, we find the ontological discussions, risk assessment of digital technologies, and scientific discourse to employ narrative elements from the cybernetic imagination. As such, it is during this period that the play of replacement is most active, visible, and effective. It can be observed in the cultural, socio-political and economic landscapes of the Western world. As for cyberspace and artificial intelligence, the tension between the newly formed post-war disciplines of cybernetics, computer sciences, and the humanities is most clearly seen in literature, culminating in the works of Philip K. Dick. Therefore, in order to attempt to answer the question "why this fear of robots?" we should initially look at the intellectual changes during the years following the Industrial Revolution.

With Nietzsche, the post-religious society was set to begin. The twentieth century proved the optimists of the early century wrong. Technology, rather than decreasing the likelihood of war, became a major component of two world wars culminating with the nuclear bombings of Hiroshima and Nagasaki. Playing with replacement, at the dawn of Nietzsche in the 20th century, found in the cybernetic imagination a xeno-nihilistic confirmation and the reconfiguration of traditional academic fields of the humanities in the language and methods of the sciences. If the defeat of theology was truly successful by the start of the Great War, then replacement focused on the defeat of man – the concept of the "uncanny valley" (Mori, 1970) [4] being its primary

[4] In robotics, the "uncanny valley" is a term coined by Masahiro Mori in 1970 to analyze the level of human likeness of robots. In the paper, he observed multiple levels of the human likeness and cautioned designers to remain a step below achieving robot design indistinguishable from humans.

symptom. Nietzsche's *Uebermensch* was intended to side out the old man engaged in nihilistic religions, but, as it turns out, humanity awaits anxiously its own replacement in the century to come.

2.1. Human and Machine: Melville, Meyrink, and Capek

The social landscape of the 19th century changed significantly. In population alone, Europe saw an increase of 50%, by the end of the century, of its previously 300 million people. Steel production increased 350% in Britain, over 1,400% in Germany and 8,600% in the United States. Coal production increased by 250% in Germany, similar to Britain, and 400% in the United States. The economic growth was coupled with an increase in literacy. Britain, for example, experienced the rise (and founding) of a new profession: journalism (Conboy, 2004). Across Europe, publications spawned in major urban centers as people migrated towards cities. By the end of the century, reports observed that literacy rates approached 100% in Britain, France, and Germany (Echevarria, 2007, pp.3-5). The telegraph was introduced, ushering in a revolution in communications.

Technological advancement and economic prosperity developed in literature what Antulio Echevarria II calls "utopian hopes" and "dystopian fears." The former is a body of literature emphasizing the best practices for a good society. In this section, we find classics like William Morris' *News from Nowhere*, a socialist utopian novel published in the last decade of the century. Ivan Bloch in his *The Future of War in Its Technical, Economic, and Political Aspects* (1899) (Bloch, 2017) revealed the optimistic expectations of modern technology by arguing that war is no longer feasible. Dystopian narratives, on the other hand, portrayed the dark

consequences of mechanized work, technological substitution of individuals, or even the disappearance of the meaning of life (Echevarria, 2007, pp.5-47).[5] By mid-century, the public began to be curious about the marvelous discoveries of science and technology.[6]

Sometime before 1855, Herman Melville received a letter from an unnamed reader saying that "like negroes, these powers own man sullenly; mindful of their higher master; while serving, plot revenge." Among other quotes that Melville provides from the letter is the following: "seeking to conquer a larger liberty, man but extends the empire of necessity" (Melville, 1855, p. 123). The first quote refers to the powers harnessed by technological advancement peeking into the possible repercussions it can have on humanity: a revolt, a takeover. It recalls the slave-master dichotomy expected of technology, with the implicit fear of replacement by potential revolt. The second quote suggests that technological progress does not just enable the individual to be free from the dangers of nature, but that the price to pay is a further increase in dependence on technology. In other words, technology provides solutions while advancing necessity.

The letters reveal the *Geist* of the century, particularly the second half of the nineteenth century when Europe has changed dramatically. While, previously, wars between nations were mere skirmishers, Napoleon managed to harness the logistical power of the railway and communication technologies to accompany the *Grande Armee* from Spain to Egypt to Moscow

[5] The Luddite revolution of 1812 certainly played a part in this development. Nicols Fox explores the subtle resistances to technology in the book *Against the Machine: the Hidden Luddite Tradition in Literature, Art, and Individual Lives*, Island Press, 2004; also see Bailey, Brian, *The Luddite Rebellion*, Sutton, 1998

[6] Mary Shelley's *Frankenstein* was published in 1823

and back to Paris.[7] The century has known the change of the nature of information as the telegraph removed the traditional physical requirement of the message (i.e. paper) in favor of electric signals transported almost instantly through vast geographic areas. Connection was established between New York and London, and many optimists declared enthusiastically that a future war on the scale of the Napoleonic Wars was impossible since governments could now quickly communicate with each other and reason would prevail.[8] As we know, that was not the case. With all the enthusiasm, praise for new technology, optimism and pacifistic hopefulness, the world experienced the Great War, and then the Second World War, known precisely for advancement in military technology with deadly effect (machine gun), a change in tactic (Blitzkrieg), and intentional genocides (Holocaust).

Concerns about the effects of science and technology have been voiced by people such as the unknown writer of the letters to Melville and by the great classic writer himself. In 1855, Melville sent a story to Putnam's Monthly Magazine titled "The Bell-Tower." In it, he considers the possibility of human technological replicas in the construction of a structure similar to the Tower of Babel. "Stone by stone, month by month, the tower rose. Higher, higher, snail-line in pace, but torch or rocket in its pride" (Melville, 1995, p. 141). And this may be the metaphorical meaning of the bell tower, a symbol of pride enabled by technological and scientific progress. In the tower, "built by the great mechanician, the unblest foundling, Bannadona", a human-resembling entity was constructed to strike the bells at sharp hour. Bannadona, however, is

[7] This is usually described as the First Generation of War and transitions into the Second Generation of War which is the increase technologization of the battlefield, such as the Great War. See Hammes, Thomas, X., *The Sling and the Stone: On War in the 21st Century*, Zenith Press, 2006

[8] For a history of the change of communication technologies from physical to digital and some of the optimistic voices arguing that war is evitable, see Standage, Tom, *The Victorian Internet*, Bloomsbury, 1998

warned that there is a rule in art "which bars the possibility of duplicates." In a sense, the builder reminds the architect not to replicate human life through technical means. Nevertheless, Bannadonna proceeds with the construction, and when the town gathers to gaze at the finished Bell Tower, they find the constructor lying dead near the mechanical construction. We are told that besides the functions that the machine is required to have, that is to have the hands and mechanism to strike the bells of the tower, "Bannadonna had resolved that his invention should likewise possess the power of locomotion, and, along with the appearance, at least, of intelligence and will" (*Ibid.*, p. 150).

Necessity, in its imaginary form, persuaded Bannadonna to create the mechanical entity with independent will and intelligence. "Talus was to have been the all-accomplished helot's name. Talus, iron slave to Bannadonna, and, through him, to man" (*Ibid.*, p. 150). But Talus, once functioning, in its desire to strike the bell, had stricken Bannadonna on the head as well, killing him on the spot. Melville appears to suggest that this necessity will awaken unknown desires of progress in the heart of the people able to create mechanical beings:

> "A practical materialist, what Bannadonna had aimed at was to have been reached, not by logic, not by crucible, not by conjuration, not by altars; but by plain vice-bench and hammer. In short, to solve nature, to steal into her, to intrigue beyond her, to procure some one else to bind her to his hand; - these, one and all, had not been his objects; but, asking no favors from any element or any being, of himself, to rival her, outstrip her, and rule her. He stooped to conquer. With him, common sense was theurgy; machinery, miracle; Prometheus, the heroic name for machinist; man, the true God" (*Ibid.*, p. 151).

Melville here recalls the deployment of Prometheus as a symbol of resistance to nature and of protest against Zeus. Hans Blumenberg, in *Work on Myth*, details the co-opting of the Titan for similar purposes from Goethe to Marx and Nietzsche (Blumenberg, 1985). Prometheus in the nineteenth century resembles resistance to God and the facilitator of becoming God. And so, Melville ends his story with prophetic calm:

> "So the blind slave obeyed its blinder lord; but, in obedience, slew him. So the creator was killed by the creature. So the bell was too heavy for the tower. So the bell's main weakness was where man's blood had flawed it. And so pride went before the fall" (Melville, 1995, p. 153).

The incursion into Mary Shelley's *Frankenstein: The Modern Prometheus* is different in the sense that the Creature is put together using human parts, therefore, elements of the natural creation of man. Talus, on the contrary, recollecting the idea of automaton from Greek mythology, is a construct entirely artificial. However, both Shelley and Melville are concerned with the potential of techno-scientific progress in establishing man as God. Shelley's creature remains unnamed, therefore denying it the very act of existence, while in *The Bell-Tower* the construct is named. Frankenstein dies and his Creature follows suit while in Melville's story, Bannadonna dies and Talus continues to work. Novels and stories like these reflect the expectation of the age that science has prevailed over God, technology over nature, and underlying these achievements lies the anxiety for technology's destructive potential. It appears that the only possible way to conceive victory is to expect, just like Zeus, that someone will be born or rather created who will take our place. The imagination continues in the lines set by the ancients that we observe in violent contests, even if it makes little rational sense.

The utility and influence of emotions in rational decision-making are traits usually attributed to humans. If the mind is mechanistic and can, eventually, be duplicated by machines, much to the protest of Descartes, and the body as well, then what we have to remain distinguishable from constructs like Talus are emotions. As such, the question of what it means to be human enters the cybernetic imagination.

Not long after Melville, *Dr. Materialismus*, an interesting story published by Frederic Jesup Stimson in 1890, takes on the question of technologically reproducible human emotions and imagines a device that can artificially induce them in people. In the story, unchecked emotions, far from resembling a state of inebriation similar to Alcibiades in *Symposium*, are entirely replicable through technological means and, furthermore, inducible through a device. Language in this case conforms to explain the scientific reasoning behind the process. Dr. Materialismus, or just Dr. Mismus, a "German professor, scientist, socialist" being recently employed by a college as a German instructor is "practicing hypnotism, magnetism, mesmerism, and mysticism", a list perhaps too sort of "-isms", while he is also "giving lectures on Hegel, believing in Hartmann, and the indestructibility of matter and the destructibility of the soul; and his soul was a damned one, one he care not for the loss of it" (Stimson, 1995, p. 159).[9] Opposing his views is C.S.J.J. Tetherby, a college student with affection for Althea Hardy. When Dr. Mismus finds romantic favor with Althea, Tetherby challenges him in an intellectual battle on the question of what it means to be human. Losing due to Dr. Mismus interminable materialistic arguments, he plots revenge. Facing Dr. Mismus he is told that "What you call love is but a multiple of lust and cosmic love, that is, gravitation" (*Ibid.,* p. 168). In addition, Mismus explains that "about two hundred thousand vibrations make in man's cerebrum what you call lust; about four billion per second, that is gravitation, make what the philosophers call will, the poets, cosmic love; this comes just after light, while light, which is the sum of all the lights. And their multiple again, of love and light, makes many sextillions, and that is love of God, what the priests name religion" (*Ibid.,* p. 168).

[9] *Althea* appears to be inspired by the Greek word ἀλήθεια meaning truth or reality.

The fictional scientific explanation of love, even within the parameters offered here, does not make the priests wrong. Instead, it would suggest that science offers a different language to explain truths already found in philosophy and theology. Nevertheless, the point of the conversation is not about uncovering what is true, but to show that one language can replace the other. In order to prove his views correct, Dr. Mismus connects the young college student to a device that he invented to replicate human emotions. Turning the "rhythm to the thought millions" is expected to produce "what you call evil passions, between them and what you call the good ones. It is all a mere question of degree" (*Ibid.*, p. 168). And Tetherby experiences the signals produced by the device sending him in "a sea of thoughts and phantasies" able to feel "the inspiration of a Shakespeare" or "the fancy of a Mozart or a Titian" and understand "the study of Newton" as his own creation. Thoughts and feelings interlinked while a turbulent flow of emotions and desires rushed through his being: "my soul was driven like a leaf before the autumn wind" (*Ibid.*, p. 169). Dr. Mismus takes confidence in the supposition that because he can induce the feelings into a human using machinery, it proves that emotions are mechanical. Whether or not that is correct, it is certainly not the conclusion sought by Stimson.

The story shows the tension of replacement that is behind the narrative adopted to push explanation. Since science and technology operate in matter, the conclusion of this narrative must be that only physical elements constitute existence. In this sense, Dr. Mismus denies the possibility of metaphysics and reduces, implicitly, existence from understanding to sensible experience and explanation.

Playing with replacement was rearticulated, as such, in the long nineteenth century when industrial machinery was perceived to pose existential threat to humanity. The cybernetic

imagination carries this play and adjusts it to contemporary technological developments, particularly in the formulation of virtual reality and production of artificial intelligence.

The nineteenth century was the century of change. It produced scientific discoveries and technological advancements, the definition of academic fields, the partitioning of knowledge in particular scholarly disciplines (however erroneous that may be) and promoted the potential of technology to solve socio-political and even philosophical or religious problems and needs. Culture inevitably adapted, but the play of replacement found reformulation in various literary forms. Among the prominent, and perhaps the most influential in the century to come, was science fiction, known previously as "scientific romance" (H.G. Wells) or, later, "scientifiction" (Hugo Gernsback) before it got its term and genre of "science fiction".

At the heart of literature during this period pertinent to our concerns lies the technology of travel. Darko Suvin notes that around this time "space loses its monopoly upon the location of estrangement and alternative horizons shift from space to time" (Suvin, 1979, p. 89). Space, too, is reimagined to be a distant simulation of present or possible realities. Belief, as well, was reoriented from its traditional religious source towards science, and debate ensued over the future of a materialistic man in place of the traditional religious servant. Among the most important publications on this topic is *Le dernier homme*, published in 1805 by Jean-Baptiste Cousin de Grainville. The author, a dispositioned priest during the French Revolution, attempts to balance the religious apocalypse with the new, materialistic, view of the end of the world (Roberts, 2016, p. 122). The protagonists, the last surviving human couple, Omegarus and Syderia, fertile but childless, live in a world without Christianity. The apocalypse here draws from St. John without being in line with Biblical narrative. It expresses, according to Adam Roberts, a complete nihilism, where humanity and the world cease to exist (*Ibid.*, p. 123). In this sense, even if the

couple are suggested to become the new Adam and Eve, the world of technology annihilates that very possibility. Karel Capek, a century later, will provide the opportunity for machines to repopulate the planet after a techno-apocalypse caused by Rossum's invention. Thus, the possibility of a couple or artificial entities succeeding humanity is apprehensible in 19th century science fiction and continues to suggest that technology will replace humanity.

Mystical Robots

The decline of religious life and its gradual substitution with a scientific, technology-based existence, provided the source of "mystical" scientific romance. The authors of this literary segment sought to blend science and spirituality. We find recurring themes of monism and atomism that are meant to explain the foundation of the universe or the creation of beings able to transcend physical laws. In particular, Humphy Davy's *Consolations in Travel; or, The Last Days of a Philosopher* published in 1830 portrays the protagonist, Philalethes accompanied by a spirit that unveils for him the true composition of reality. According to Adam Roberts, the novel is "a spiritual cosmos interpreting the material as a desired destination" (*Ibid.*, p. 162). In France, Camille Flammarion published *La pluralite des mondes habites*, in 1862. As the title suggests, the work was an exploration of life on other planets.

Later, in 1905, E.T.A. Hoffmann in *Der Sandmann* imagines a relationship between Nathanael and Olympia. The Sandman is Coppelius, a "horrible" old lawyer and a frequent guest at Nathanael's. The language of the story frequently makes reference to evil, destruction, sin, "horror" and the dark fate of the protagonist. Nathanael meets Olympia at a dinner party where she "appeared, elegantly and tastefully dressed. No one could help but admire her beautifully shaped face and her figure" (Hoffmann, 1982, p. 299). He instantly falls in love with her.

Noticing her uncanniness, his friend Siegmund, tells Nathanael that "she seems to us – don't take this badly, my brother - strangely stiff and soulless. Her figure is symmetrical, so is her face, that's true enough, and if her eyes were not so completely devoid of life - the power of vision, I mean - she might be considered beautiful. Her step is peculiarly measured; all of her movements seem to stem from some kind of clockwork. Her playing and her singing are unpleasantly perfect, being as lifeless as a music box; it is the same with her dancing. We found Olympia to be rather weird, and we wanted to have nothing to do with her. She seems to us to be playing the part of a human being, an it's as if there really were something hidden behind all of this" (*Ibid.*, pp. 301-302). But after a short-lived love experience, Nathanael discovers that Olympia is a "lifeless doll."

Responding to *Der Sandmann*, Ernst Jentsch publishes a paper in *Psychiatrisch-Neurologische Wochenschrift* arguing that if an automaton replicates human behavior and appearance, the observer will not be able to make the difference. The title of the paper, *Zur Psychologie des Unheimlichen*, proposes that the uncanny (*das Unheimliche*) originates in "doubt as to whether an apparently living being really is animate and, conversely, doubt as to whether a lifeless object may not in fact be animate" (Ravetto-Biagioli, 2019, p.9).[10] Indeed, the uncanny has become a major concern in contemporary design of intelligent machines.

As previously suggested, the 19th century ended with the victory of science over religion, found in Nietzsche's famous declaration of the death of God. *Das Unheimliche* prevails in the

[10] Sigmund Freud responded to the problem of the uncanny by arguing that *Der Sandmann* represents, according to Ravetto-Biagioli, "an automatic but also intentional and embodied (though symptomatic) response to 'something which ought to have remain hidden but has come to light" (p. 10). Seen from this perspective, Gustav Meyrink's *Der Golem* can be considered an embodiment of a perennial fear embedded in Western civilization. We argue, later in this chapter, that *Der Golem* represents the play of replacement always lurking in the background of the effects of technological progress regarding artificial reality and intelligent machines.

cybernetic imagination as can be seen in mid-to-late 20th century in the fiction of Philip K. Dick. However, the first author to experiment with the uncanny was Karel Capek whose play *Rossum's Universal Robots* coined and popularized the term robot.

In turn, Gustav Meyrink attributes the "mystical" to an artificial construct with deep cultural and spiritual significance in Jewish folklore and religion. The idea of the machine as a metaphor for a threat lurking in the background takes the form of a mystic appearance in Meyrink's *Der Golem: Ein metaphysische Roman* published in 1915.[11] The novel recreates the uncertainty of reality (Wrobel, 2010; Klaus, 2010; Xiuli, 2012) the ever presence of human-like artificial constructs (Durzak, 2003), and proposes, through the symbol of the Golem in ancient Hebrew tradition,[12] a bridge and resemblance of the creation of man by God. In this sense, it can

[11] The success of the novel is impressive for its time. According to Maya Barzilai, the book sold more than 150,000 copies during the Great War and 200,000 copies in early 1920s and was adapted to the screen by Paul Wegener on three different occasions between 1915 and 1920. This success is probably because of the "perceived connection to Jewish mysticism and occult practices" initiating "a wave of fascination with spiritualism and 'unmodern' or 'myth-ridden' phenomena" in a context where the triumph of materialism seemed total, Barzilai, Maya, *Golem: Modern Wars and Their Monsters*, New York University Press, 2016, p. 4

[12] According to Moshe Idel, artificial entities were traditionally part of ancient magic. Some artificial anthropoids, as the author calls them, have been simple automata resembling moving machines, while others incorporated the ability of speech and some had full spiritual capabilities. In *Golem. Jewish Magical and Mystical Traditions of the Artificial Anthropoid*, Ktav Publishing House, 2019, Idel traces early examples of artificial humans in Egyptian mysticism. These entities were called *ushabti*, a term designating "answer" which, Idel notes, "seems to be especially important for the understanding of the classical passage in the Talmudic tradition dealing with the creation of the anthropoid." The scholar points to the term '*emet* meaning truth that was supposed to be written on the forehead of the construct in order to gain life. In *De Veritas et Mendacio*, truth is required for the statues to be able to talk. Further references direct to *Clementine Recognitiones*, where Simon Magus boasts of creating life out of a finer element than God. Magus has used "air" to create a living construct while God created man from earth. Idel observes that "Golem, refers in medieval Hebrew both to a magically created body out of dust, and, in other contexts, to the spiritual body that differs from the soul, and is formed out of the air of Paradise." Furthermore, Idel discovers a connection between the creation of the world and the creation of the Golem in the text *Sefer Yerizah*. The *Leviticus Rabbah* points to the idea of Adam as Golem while a medieval manuscript tells of Enosh creating a living statue in the image of man in front of a crowd, but when a demon enters the talking statue, the people start to worship it. This is a rather prophetic message for the 21st century, if we consider *The Machine Stops* of 1909 and *Matrix* of 1999. Moshe Idel investigates the Golem throughout the Hebrew tradition. In addition, he dedicates a chapter to "Golem: Imaginaire, Anomian, and Silent" which, curiously enough, does not mention Meyrink's novel. In the essay titled

be suggested that Meyrink's Golem represents the idea of replacement always present but never fully articulated, since the construct rarely makes a presence. Instead of being an active character participating in the plot, the Golem appears mostly in mentions by other characters.

Der Golem is, as its original title says, a metaphysical novel. The reader may initially think that the story is told in first person to recollect the life of Athanasius Pernath in the ghetto of Prague. However, sometimes the voice of the narrator is someone else and often it becomes unclear whether Pernath is telling the story or the unknown character. It could very well be the Golem. Pernath himself is uncertain whether his presence in the world is a matter of fact or a dream. Athanasius Pernath "cannot sleep, I cannot wake; in its half dreaming state my mind forms a curious compound of things it has seen, things it has read, things it has heard – streams, each with its own degree of clarity and colour, that intermingle, and penetrate my thought" (Meyrink, 1985, p. 3). Not long after, the dreaming state brings about an ontological crisis. "Who am I?" he asks himself and quickly finds the question redundant since "I no longer possess any organ to whom this query might be addressed" (*Ibid.,* p. 4). In this sense, Meyrink takes up the old question of sensible reality as an illusion and creates a bridge between it and the unreal. But

---

"Crafting the Golem: The Creation of an Artificial Anthropoid" Idel writes that "although Shelley deals expressly with some of the ambivalence expressed toward early-nineteenth century science, she also drew some inspiration from the atmosphere of Prague... [...] it became the scene for the development of a variety of traditions concerning the creation of an artificial anthropoid, the Golem, a problematic figure whose image haunts the city even nowadays." John Neubauer traces the literary history of the Golem of Prague in *How Did the Golem Get to Prague?* In Marcel Cornis-Pope and John Neubauer (eds.), *History of the Literary Cultures of East-Central Europe: Junctures and Disjunctures in the 19th and 20th centuries. Volume IV: Types and Stereotypes*, John Benjamins Publishing Company, 2004, pp. 296 – 307; Moshe Idel notes that the Golem is good or evil, but a play *The Golem* by H. Leivick depicts the artificial anthropoid as the embodiment of terror: "So will the light go out / In every eye that looks on me". Lewis Glinert explores the socio-political, cultural, and philosophical implications of the myth of the Golem in "Golem! The Making of a Modern Myth" in *Symposium: A Quarterly Journal in Modern Literatures*, Vol. 55, No. 2, 2001, pp 78 – 94; Edan Dekel and David Gnatt Gurley present a comprehensive bibliographical list on the topic of the Golem in "How the Golem Came to Prague" in *The Jewish Quarterly Review*, Vol. 103, No. 2, 2013, pp. 241 – 258; Ricarda Hirte identifies "Jüdisches Kulturgut in Gustav Meyrinks Roman 'Der Golem'" in *Estudios Filológicos Alemanes*, Vol. 12, 2006, pp. 479 - 494

this uncertainty presents a new way of acquiring knowledge (Krieger, 1996)[13], perhaps as a way of *katabasis* (descent into the underworld). Pernath discovers through the voice constantly present in his mind, uttering perpetually within him, a sense of not belonging to the physical world by opening up to the possibility of being part of a world outside: ”I understood now that the mainspring of all my thoughts and acts lay hidden in another world, forgotten and never to be recalled... [...] There were things in the world, so it seemed to me, beyond the mind of man to grasp, riveted indissolubly together and running about distractedly, like blind horses, on a path whose direction is hidden from them.”[14] This presence between the worlds of the protagonist, a liminal space, persists throughout the novel.[15] As such, when the objects around are no longer ontologically certain, the interpretation of things must inevitably rest on perception.

The Golem is introduced first in the mind of Pernath. As the novel suggests, reality is confined to the mind, in a similar way expressed by Berkeley and experienced as a virtual reality in *Pygmalion's Spectacles*:

---

[13] Meyrink recalls the Socratic notion of recollection in the novel combining the Book of Life and a path of knowledge chosen by free will much in the like of Parmenides and states “[a]s knowledge comes, so comes recollection. *Knowledge and recollection are one and the same thing.”* (Meyrink, 1985, p. 47). The theory of recollection as is presented in Plato‘s *Meno* and *Phaedo,* proposes the idea of knowledge experienced by souls outside of life, in death, and when entering a human body on birth the knowledge is lost but the sense of it remains.

[14] Mayrink, *op. cit.*, p. 33

[15] In this sense, Athanasius Pernath is perhaps akin to the Golem. At one point in the novel, Pernath is randomly taken to be the Golem (Meyrink, 1985, p. 69). Furthermore, it can suggest the view of man as a “rational animal” located between the sensible reality and the intelligible world. As Christos Yannaras writes in *The Schism of Philosophy*: “Human reason is not only the ability to discern, to come to know, and to express the form/formation/mode of cosmic harmony and order. It is also the power to create a kosmos, to give form to what is formless, to endow irrationality with reason” (Yannaras, 2015, p. 8). Accordingly, rather than this being a case of mental illness as some scholars have suggested, it may be interpreted as a liminal space. Space in Meyrink‘s novel has been researched by Pérez de la Fuente, María Belén, Der Raum in Der Golem von Gustav Meyrink: Transgression der Gothic Novel, in *Estudios Filológicos Alemanes,* No. 26, 2013, pp. 191 - 198 and Pastuszka, Anna, Einige Bemerkungen zu Raumdarstellung und -deutung in 'Golem' von Gustav Meyrink, in *Lubelskie Materiały Neofilologiczne,* No. 20, 1996, pp. 57 - 73

> "Then, in mysterious fashion, comes into my mind the legend of the mysterious Golem, artificial man, whom once, long ago, here in the Ghetto, a rabbi learned in the Kabbala shaped from the elements, investing it with an unreasoning automatic life when he placed a magical formula behind his teeth. And, as that same Golem stiffened into clay the instant that mysterious phrase was removed from its lips, so must, I thought, these humans dwindle the soulless entities so soon as was extinguished within them some slightest spark of an idea, some species of dumb striving, however irrelevant, already deteriorated with most of them, from the look of it, into a mere aimless sloth, or a dull waiting for they know not what" (Meyrink, 1985, p. 16).

The Golem remains in this mental space. As Meyrink seems to suggest, there is little to truly know about it. The Golem, instead of being admitted into reality, is kept in the realm of legend until a tragedy occurs, which then "turns it into actuality again" (*Ibid.*, p. 26). According to the legend, the rabbi in Prague created the Golem "to help ring the bells in the Synagogue." Like Melville's *The Bell-Tower*, the artificial entity began a murderous and destructive streak in the streets of the city. Death, as it seems, is ever present in the imagination about artificial humans. But the Golem is neither dead nor alive, it is "something that cannot die, and has its being within our midst" (*Ibid.*, p. 26).[16] It has become a state transcending reason and reality, always lurking in the shadow ready to emerge. Similar to the effects of communication technology, the blending of reality and the dream world is found in the tension between them. Real becomes unreal and vice-versa, but the author takes a step away from replacement and allows the unreal to be incorporated in the real.

Physical Robots

[16]The Golem takes multiple identities. At one point, Athanasius Pernath is confused with the Golem, or it might be the Golem itself given the unclear identity of the narrator (it is in the end of the novel when Athanasius Pernath and Miriam, both dead, look down on the town and then Athanasius looks at the narrator who thinks "So like is he to myself, it is as though beholding my own face and figure in the glass!" As it suggests, Pernath have been made in the image and likeness of their Creator). At another, the Golem is the beggar Haschile, "the mad" (p. 123)

The cybernetic imagination, then, inevitably, began to envision the replacement of humanity by machines. It was in 1920 that a Czech playwright and author, Karel Capek, would coin the term "robot" to describe human-resembling machines. And yet, it may not have been Karel who came up with the term, as it is believed, but his brother (Harkova & Kelemen, 2008, pp. 283-306).[17] Regardless of its origin, the play *Rossum's Universal Robots* put forward socio-economic and ethical issues that are today part of "the most important conversation of our time" (Tegmark, 2017, p. 22). What Capek wrote then in fiction is experimental technology today and, therefore, captures the spirit and the continuation of the fear of replacement in the cybernetic imagination.

The play, *Rossum's Universal Robots*, is about the human attempt to go beyond the work of nature. There is a tone of arrogance and even hubris when Domin, the director of the company producing the artificial human workers,[18]explains that "nature has found only one method of organizing living matter. There is, however, another method, more simple, flexible and rapid, which has not yet occurred to nature at all" (Capek, 2001, p. 4). In this reference to Rossum's memoirs, Domin boasts that humanity has found a new way to produce life outside of nature. Humanity no longer has to rely on natural reproduction to create life, but on chemistry and

[17] The story, according to the authors, involves a news article from Lidove noviny *(*People's News), a Prague newspaper that published in December 24, 1933, an article suggesting that Josef Capek, Karel's brother, suggested the term during the following conversation: "But," the author said [Karel], "I don't know what to call these artificial workers. I could call them Labori, but that strikes me as a bit bookish." [Josef] "Then call them Robots," the painter muttered, brush in mouth, and went on painting. And that's how it was. Thus was the word *Robot* born; let this acknowledge its true creator." Harakova and Kelemen point out that the word *robot* is etymologically derived from the old Czech word *robota* which "in present-day Czech and Slovak, *robota* means something like a serf's obligatory work" while *Rossum* is also derived from the Czech word *rozum* meaning reason. *Ibidem*, p. 285 - 287. Later Karel Capek wrote that for the *Prager Tagblatt*, a German-language newspaper in Prague, about the resemblance of *robot* and the famous artificial entity known locally by the myth of *Golem*. Gustav Meyriak's novel *Der Golem* was published in 1915 and Karel Capek may have been familiar with it.

[18] His name derives from the Latin *dominus* which translates to *master*.

engineering. And so, Domin is right to humor the choice of Rossum to reproduce the human species in an artificial form. He tells Helena, daughter of the president of the company, that "He [Rossum] could have produced a Medusa with the brain of Socrates or a worm fifty yards long. But being without a grain of humor, he took it into his head to make a vertebrate or perhaps a man" (*Ibid.*, p. 4). Furthermore, Domin explains that Rossum's "sole purpose was nothing more nor less than to prove that God was no longer necessary" (*Ibid.*, p. 15). As a result, the new creator, man, developed an artificial version of himself that is "more perfect", smarter, faster, but without a soul. However, in any other way, the *robots* resemble humans fully. In fact, during Act I, Scene 1, a third character is witnessing the conversation between Helena and Domin, who is, unbeknownst to Helena, a robot. Sulla also, and another robot introduced later as Marius, are indistinguishable from a living natural person.

Dreaming of a world in which labor is fully delegated to machines, Domin believes that "everybody will be free from worry and liberated from the degradation of labor. Everybody will live only to perfect himself" (*Ibid.*, p. 15). A noble pursuit as it may be, but Domin denies the possibility that the act of labor could contribute to becoming better humans. It is clear that Domin intends to delegate all types of labor, including creative work, to robots. Of course, if all creative, philosophical, intellectual, and editorial labor is reserved for machines, what tools do humans still have at their disposal to perfect themselves? Capek does not provide an answer, which could suggest that he may refer strictly to physical labor. Furthermore, the author does not imagine the socio-political reconfiguration required for such a project but lays out only the anticipated consequences of thinking machines resembling humans.

In addition to delegating labor to robots, the play turns its attention to the transfer of human morality to the artificial beings. More specifically, it explores the power-driven mentality

seeking dominance and control that is the essence of the play of replacement. Radius, one of Rossum's creations, explains it as following: "You are not as strong as the Robots. You are not as skillful as the Robots. The Robots can do everything. You only give orders. You do nothing but talk. [...] I don't want a master. I want to be master. I want to be master over others" (*Ibid.*, p. 27). As such, the traditional master-slave narrative in which artificial beings, automata, or intelligent machines are created to serve humanity is under threat. Instead, the master becomes the slave. Aware of this possibility, Domin nevertheless continues to deliver robots to increase earnings for the company and to abolish labor altogether. In this moral framework, it is only a matter of time before the current masters are replaced by new ones.

Eventually, Domin receives a note that reads: "Robots throughout the world: We, the first international organization of Rossum's Universal Robots, proclaim man as our enemy, and an outlaw in the universe" (*Ibid.*, p. 34). A replacement needed to be complete: "we command you to kill all mankind. Spare no men. Spare no women. Save factories, railways, machinery, mines, and war materials. Destroy the rest. Then return to work. Work must not be stopped" (*Ibid.*, p. 34). The machines not only revolt against humanity, but against nature as well. They appear to be a product of artificial substance turning against a nature deemed useless.

In addition, Domin's dream to abolish labor is shattered in the process. If Domin does not see anything appropriate in labor for mankind to become better, seeing it rather as an absurd undertaking, the robots participate in this absurdity by continuing work without a goal. Without the expectation of labor to be beneficial in the sense of increasing prosperity and the quality of life of humans or of robots, they transform work into an end goal. But eventually, Domin himself gives meaning to labor while debating the moral undertaking of producing robots with another engineer, Alquist:

> "ALQUIST: It was a crime to make Robots.
> DOMIN: No, Alquist. I don't regret that even to-day.
> ALQUIST: Not even to-day?
> DOMIN: Not even to-day, the last day of civilization. It was a colossal achievement.
> DOMIN: Alquist, this is our last hour. We are already speaking half in the other world. It was not an evil dream to shatter the servitude of labor – the dreadful and humiliating labor that man had to undergo. Work was too hard. Life was too hard. And to overcome that ---
> ALQUIST: Was not what the two Rossums dreamed of. Old Rossum only thought of his God-less tricks and the young one of his milliards. And that's not what your R.U.R. shareholders dream of either. They dream of dividends, and their dividends are the ruin of mankind.
> DOMIN: To hell with your dividends. Do you suppose I'd have done an hour's work for them? It was for myself that I worked, for my own satisfaction. I wanted man to become the master, so that he shouldn't live merely for a crust of bread. I wanted not a single soul to be broken by other people's machinery. I wanted nothing, nothing, nothing to be left of this appalling social structure. I'm revolted by poverty. I wanted a new generation. I wanted – I thought ----
> ALQUIST: What?
> DOMIN: I wanted to turn the whole of mankind into an aristocracy of the world. An aristocracy nourished by milliards of mechanical slaves. Unrestricted, free and consummated in man. And maybe more than man.
> ALQUIST: Super-man?
> DOMIN: Yes. Oh, only to have a hundred years of time! Another hundred years for the future of mankind" (*Ibid.*, p. 39-40).

In other words, Domin explains that the meaning of his relentless labor, for which he did not expect overcompensation, was directed towards the bettering of humankind.

The robots realize that they do not have the plans to recreate themselves in the factories left standing. As such, they keep Alquist alive with the intent to extract from him the necessary information. Alquist comes to observe that the robots start to develop the ability of what appears to be emotional connection. In this situation, with mankind extinct, Alquist proclaims the robots as the heirs of the world. The play ends with Alquist saying: "Go, Adam, go, Eve. The world is yours" (*Ibid.*, p. 58). Therefore, the play indirectly asks the question of "what does it mean to be human?" or "what would make robots legitimate heirs of humanity?" The answer that Capek provides is emotions.

2.2. Virtual Reality: Weinbaum and Forster

Around the same time in the United States the door is opened for the idea of replacement in visions of proto-virtual reality. As agonistic play finds its manifestation in the physical realm of prowess in battle in the *Iliad* and then it is reconfigured in the logo-rational in the *Odyssey*, so too in the early twentieth century, agonistic play finds a way into the new artificial realities imagined by authors of early science fiction. Two texts are of particular importance here. The first is *Pygmalion's Spectacles* by Stanley G. Weinbaum and the other is *The Machine Stops* by E.M. Forster.

Pygmalion's Spectacles, initially published in 1935, starts by asking "but what is reality?" (Weinbaum, 1949, p. 160). The question, of course, is not new. But what is relevant about this particular instance is the answer given in the new context of the *interbellum*. Europe at this time is the playground of ideological violent contests primarily between two socialist political philosophies: National Socialism and its variation of Fascism against Communism. In this context, the ideas of liberty upheld by European Liberalism and conservatives of traditional national values or institutions have been largely silenced or found their way into concentration camps or the Gulag. In Austria, Ludwig von Mises was following the economic thinking of Adam Smith promoting an individualist philosophy grounded in economics that was foreign to Eastern European cultures and antithetical to the socialist desires of a new man in Western Europe.[19] Furthermore, the potential of mass killing with new military technology developed and

[19] Mises found refuge in the United States in 1940 where his economic ideas were well received. For a detailed account of the life of Mises see Joerg Guildo Huelsmann, *Mises: The Last Knight of Liberalism,* Ludwig von Mises Institute, 2007; The argument that I am making here by saying that "individualist philosophy grounded in economics was foreign to Eastern European cultures" is influenced by Christos Yannaras' book *The Schism in Philosophy*, Holy Cross Orthodox Press, 2015, trans. Norman Russell. The 'schism in philosophy' is identified by Yannaras in two modes of thinking that differ substantially. In his

deployed during the Great War[20] dispelled the illusions of perpetual peace of the optimists of the late 19th century. Pressed from all sides into conceding to ideology or to the brutal nature of humanity made visible by media technologies, authors began to find refuge in new worlds of artificial substances, fantastical voyages on other planets, and imagined realities within the existing world. Unfortunately, the violent play of replacement travelled with them, and these fictional worlds quickly became largely as violent as the palpable world itself. The question of reality was moving away from traditional metaphysics to find articulation in the language of science. Other worlds have to be somehow connected to the potential of technology and realized by science: voyages to distant places enabled by powerful new ships, travelling in time using magic-like devices, in short, technology that will create artificial worlds for us to explore at the expense (by replacement) of our own.

After questioning reality, *Pygmalion's Spectacles* states that "All is dream, all is illusion; I am your vision as you are mine" (*Ibid.*, p. 160). This statement marks a return to the Cartesian dilemma of dreams and, as it stands, incorporates practices of Eastern philosophy concerning the appearance of the world. In addition, it makes a claim that reality is confined to the senses, what we see is what it is real. But this claim is not just aligned with views in philosophy that believe sense perception to be the true and only way of defining reality. It is also a reminder of the new role of science overtaking metaphysics (or rather making it useless) at the dawn of a mentality

---

words: "The Greek *mode* was embodied historically in the *common struggle for truth* that is *political* life, the art and science of politics, the realization of the *polis* and of *democracy* as an existential (not utilitarian) goal. The Western *mode* was embodied in the safeguarding of the individual with *rights* guaranteed by institutions and conventions, as well as in amazingly advanced technological means of satisfying the human urges of self-preservation, domination, and pleasure" (p. XIII)

[20] For a general look at technological development and the evolution of warfare see Black, Jeremy, *War and Technology*, Indiana University Press, 2013; A discussion on military technology innovation during the Great War is detailed by Wolfgang Fleischer in *Military Technology of the First World War: Development, Use and Consequences*, Pen and Sword, 2017.

displeased with the uncertainty of apprehending the transcendent, the mystical, or the theological arguments which, after all, have been ‘disproved‘ in the past.

Weinbaum then directs the reader to Bishop Berkeley, “the philosopher of Idealism - no? - the one who argues that we do not see, feel, hear, taste the object, but that we have only the sensation of seeing, feeling, hearing, tasting” (*Ibid.*, pp. 160-161). Whereupon Dan Burke, the protagonist named perhaps after the quintessential figure in European conservatism Edmund Burke, issues the following challenge: ”But if your friend Berkeley is right, why can’t you take a dream and make it real? If it works one way, it must work the other” (*Ibid.*, p. 161). To which his interlocutor replies “I *did*!” Professor Albert Ludwig has created a device that will trigger the senses in such a way as to make the mental phenomena palpable. In other words, Professor Ludwig is capable of creating virtual reality in which the individual is fully immersed in both mind and body. Here’s how it works:

> “How? How? But simply! First my liquid positive, then my magic spectacles. I photograph the story in a liquid with light-sensitive chromates. I build up a complex solution – do you see? I add taste chemically and sound electrically. And when the story is recorded, then I put the solution in my spectacle – my movie projector. I electrolyze the solution, break it down; the older chromates go first, and out comes the story, sight, sound, smell, taste – all!” (*Ibid.*, p. 161).

The device developed by Ludwig is, to us, a recollection of the technological innovations of the past century retold in a way that transcends their immediate application in favor of their potential to translate dreams into physical reality or to transport physical elements within the dream in order to make it real. Reality as such is science in one way or another. Virtual reality or jacking in cyberspace as the cyberpunk literature will show us, is precisely the scientific and technological desires of the twenty-first century. Video game developers have sought relentlessly

to make their artificial worlds and characters feel real. The evolution of graphics in the last four decades far exceeds the development of art in human history.

The difference, however, is that while the art of Michelangelo in the *Sistine Chapel* was meant to represent faith, contemporary video game artists seek to influence the senses as much as possible to make the artificial world real. Various game mechanics such as character movement, particle physics, and the trajectory of fantastic spells, are made to feel realistic. Video game physics have to resemble offline world physics. Music in the games enables emotional connection with the images displayed on the screen, while game controllers vibrate on character movement in the story. In these artificial realities, we do not see actual walls and we do not actually hit them, but we perceive a sensation of a feeling of sight and hit and, as Professor Ludwig explains to Dan Burke, "The rest you interpret."

Bishop Berkeley is right in this context. Not because his philosophy stands its own ground, but because science has reduced our capacity to experience other, imaginary worlds only with physical stimulation. The cybernetic imagination at the time of Weinbaum announces that we will no longer be able to live or conceive meaningful worlds without our sensations manipulated by physical input from technology.

Dan Burke welcomes the opportunity to experience the device of his friend Professor Ludwig. The device, "vaguely reminiscent of a gas mask" had "goggles and a rubber mouthpiece". Dan was curiously examining the device while the Professor "brandished a bottle of waterly liquid." Ludwig prepared the virtual world as well, "a Utopia – just two characters and you, the audience. Now put the spectacles on" (*Ibid.*, p. 162). Once the electrolysis was initiated, Dan found himself in a front of a magnificent forest, a perfect, unearthly, and beautiful world.

From the distant mist arose a human figure. Approaching, it became clear to him that it was a girl dressed in a ”robe of silvery, half-translucent stuff, luminous as starbeams; a thin band of silver bound glowing black hair about her forehead, and other garment or ornament she had none” (*Ibid.*, p. 164). While initially he questioned the ontological validity of this new world, he began to accept the existence of the girl. It was not long before he was completely immersed in the artificial world produced by Professor Ludwig and unable to make the difference.

The girl, Galatea, told him that they are now in “Paracosma” and queried: “Does the real world seem strange...after that shadow land of yours?” To which Dan replied bewildered: “Shadow land? ... “*This* is shadow, not my world” (*Ibid.*, p. 165). Dan now finds himself in the impossible position of declaring a world in which he is fully immersed and validated by his senses as unreal. In addition, Galatea explains that Paracosma is the real world. Why wouldn’t it be? In the paradigm explained above, the primary requirement for something to be real is for the senses to create the interpretation of real. As the two explore the world, Dan is introduced to the second character of the story, Leucon, the Grey Weaver.

The world is free of the dangers of nature and ill-will. “I know these words of yours - chance, disease, death. They are not for Paracosma. Keep them in your unreal country “the Grey Weaver says. Thus, Dan continues his journey in this new world completely dismissing the existence of the former one. The relationship between Dan and Galatea evolves, and eventually he falls in love with her. At this point, the merger with the unreal is completed even after Leucon tells him that their love breaks the laws of substance. Galatea may love a shadow, but “how can a shadow love substance?” This is a reminder of the warning of Bannadona in Melville’s *The Bell-Tower* concerning the rules of art. However, as it is expected, Dan is guided by his senses and love blinds him. The feelings he is experiencing are real, regardless of whether he is in a virtual

world or the real world. When Dan insists on staying in Paracosma, the laws prevent him from doing so. As such, the laws also remind us of the physical limitations of the science that has produced the virtual reality, space and time requirements for play. For the duration of the game, the world developed within is as real as the world outside of it.

Dan finds himself returned to the spot of entry in Paracosma, and gradually the world before his eyes disappears as he is returned to physical reality. Eventually, he discovers that he was "in love with a vision! Worse – in love with a girl who had never lived, in a fantastic Utopia that was literally nowhere!" (*Ibid.*, p. 179). And while he remembered the significance of the name Galatea, the name of the statue built by Pygmalion and granted life by Venus, he thinks that "*his* Galatea, warm and lovely and vital, must remain forever without the gift of life, since he was neither Pygmalion nor God" (*Ibid.*, p. 179). The story ends with Dan's problematic acceptance of his experience as real. He tracks down Ludwig and receives an explanation. While Dan was in Paracosma, Ludwig played the role of the Grey Weaver, and his niece played Galatea. When the opportunity presents Dan with meeting Ludwig's niece, he exclaims: "Paracosma [is] attainable at last!" (*Ibid.*, p. 180).

With this conclusion, the story makes a step towards the impossibility of distinguishing between real and unreal when the senses are manipulated. It is not just the senses themselves; the experience in Paracosma returns back to the physical world and, thus, an unbreakable bridge is created. Similar to a liminal process, when Dan Burke returns, he brings with him the reality he experienced within the artificial world. Thus, our reality infused with the experiences of virtual worlds is in continuous development. Our ontology is one of process. It is becoming, and the

objects, experiences, memories created while in the artificial world become part of reality. Their ontological certainty grounded in the real world of physics and substance is irrelevant.[21]

E.M. Forster's *The Machine Stops* places people in caves, confined to comfortable armchairs, having a panel in front of them enabling them to navigate cyberspace. While connected with each other through the Machine, they stand alone within their individual caves and detest physical interactions. Instead of seeking to verify the living conditions on Earth after an apocalyptic event demanded their retreat beneath the surface, they engage in interminable lectures on topics apparently irrelevant. Furthermore, the cave, a clever reference to Plato, serves them well and confines them to a comfortable artificial reality offered by and through the Machine. On the one hand, the story imagines technology frequently used today, such as networked communication. People engage in long conferences online debating various intellectual topics. E.M. Forster is familiar with the proto-Internet technology of the telegraph

[21] In *The New Adam*, Ziff-Davis Publishing, 1939, Stanley G. Weinbaum proposes the story of "superman," a fantastic story nevertheless grounded in possibility. In the prologue, he criticizes Nietzsche and H.G. Wells for conceiving of the *Uebermensch* as a man, the former choosing a set of qualities such as fitness, potency, power, while the latter favored "the contemplative, the serene, the intellectual." Weinbaum does not agree and conceives of the superman as not human, but a possibility, a chance of nature, perhaps created by randomness. Weinbaum notes that "for not everything in the world is subject to mathematics. Not every factor in this particular sector of the cosmic whirl can be reduced to a formula, expressed in calculus, integrated, packed into nicely labeled bundles, and filed away in a book." In addition, Weinbaum critiques evolutionary biology ought to reveal the small chance of things not going to a complete plan based on direct causation. "Sometimes" he writes, "if the variants possess inherent advantages, they survive and breed true as a new species, sometimes they breed chack into the mass and are lost, and sometimes they die." In a sense, Weinbaum intends to withdraw the possible from the grip of science and reintroduce it into a mode of thinking that delegates the outcome of an event to the unknown, to the apparent impossible, without claiming a theological position. Here, as well as in *Pygmalion's Spectacles*, we find the author uneased by the promises of complete control given by science and technological development without dismissing them entirely. Instead, he tries to leave room for the "unscientific", in *The New Adam,* and the influence of the experience of a virtual reality in the scientific world unable to cope with the possibility of a "substance", man, falling in love (unclear, problematic concept) with a "shadow", artificial construct. Others may not have fallen in love with Galatea. While technology can produce environments that trigger the senses to create the perception of actual things, therefore initiating the chemical processes associated with feelings, it makes no sense for him to dismiss them as unreal, and science, as such, confined to its methodology is unable to demonstrate why the artificial world is unreal.

and such conferences seemed, at that time, just one step away. In addition, the author explores social and psychological implications of communication technologies, particularly self-isolation and a preference for online rather than physical interactions. These issues are explained when Kuno requests an in-person visit with his mother. Vashti replies that such meetings are against the will of the Machine. Thus, on the other hand, the story explores the potential of worshiping the technology created, an old possibility imagined by the ancients through the myth of Pygmalion and others.

However, the story highlights the difficulty of exiting the cave, that is, renouncing the virtual reality created by the Machine and, instead of seeking guidance from the *Book of the Machine* for interhuman relations, venturing outside and discovering the physical world. Indeed, the sensible world here is entirely replaced by the interconnected communication technology that allows individuals to generate virtual meeting places. Kuno, a dynamic individual who somehow escaped the selection process at birth, headed by the Machine performing a Spartan-like triage of newborns, discovers the possibility of accessing the surface through narrow passages where the absence of artificial light frees him from the Machine. As expected, when Kuno confesses his worries and discoveries to his mother, Vashti reacts according to the threat of replacement: this world is perfect, therefore, Kuno must be shamed and punished for betraying the wonderful works of the Machine.

Given the situation, Kuno sabotages the Machine. But instead of finding a way to optimize the Machine to serve humans without their being religiously dependent on it, he finds comfort when both humanity and Machine are destroyed. The play of replacement here is straightforward. The mentality it produces cannot imagine a peaceful collaborative alternative, but only a heroic victory in a battle to the death. It is either humanity or the Machine, and since

the latter appears to have complete control over the bodies and minds of the former, in destroying it the destruction of humanity must, inevitably, concur.

The Machine Stops observes our continuous dependence on advanced technology through our destructive impulses. Published in 1909, it acts rather prophetic for the destruction to come and transports the idea of replacement well into the cybernetic imagination. The Machine is always in the background, indirectly managing the fate of humanity. It acts as a metaphor of fate; it selects human life and creates reality. The Machine has its own prophetic book, a user-manual of Biblical value, for people to correctly worship it. In this sense, the Machine has replaced both the sensible and the intelligible realities, not creating a merger, not intending to assist in human development, but to manifest complete control.

With these novels and stories, we begin to observe in the cybernetic imagination the threat of the artificial human, the machine, and even the possibility of the human being artificial, to always loom in the background, ready to manifest the old impulses of replacement. In addition, Forster anticipates the turn of the state into a functional machine, resembling the processes of control of his own Machine.

2.3. State Machine, Robot Control

Throughout the 1930s, the idea of replacement is found in the fiction proposing the incarnation of the machine as the state. The machine here is the One-State, the World State, or the One World, and replacement is found in the transformation of the organic human into an artificial limb of the state. Stripped from his identity and allowed only to be identified as a

number and a letter denoting gender in Zamiatin's *We*, or deprived of his ability to feel fully and unhindered in *Brave New World*, the human becomes a component of the imagination of the state as a functional mechanical entity. Such novels are well documented together with classic works on the subject such as George Orwell's *1984* and *Animal Farm*.[22] These novels suggest the idea of the 'state-machine', an almost impersonal apparatus coercing people into mental and emotional submission. However, in 1947, just a couple of years after the Second World War and the fall of one such state-machine, Jack Williamson introduces the possibility in the cybernetic imagination of a conscious central system instructing robots to keep humanity safe by coercing them out of danger, self-inflicted or otherwise. His short novel *With Folded Hands* operates on a similar totalitarian idea that is found in *We* or *Brave New World*. It is not a substance, however, that keeps humanity happy; instead, happiness is enforced by robots functioning on the logic of safety.

In the world imagined by Williamson, the Humanoid Institute has produced "the perfect mechanicals" according to the principle "to serve and obey, and guard men from harm" (Williamson, 1996, p. 11). Mr. Underhill, an android salesman, has never seen a *humanoid*, but suddenly he is confronted with one, speaking directly to him about closing his business selling primitive androids. The story introduces us to Sledge, the inventor of the humanoids. The robots have taken over all the functions of society from big and small businesses to mundane tasks performed by people. The pervasiveness of the artificial entities in human life is to such a level

[22] Jack London's *The Iron Heel*, Empire Books, 2011, initially published in 1908 is a precursor for dystopian fiction emphasizing state control with machine-like properties. Among other notable contributions to the dystopian genre written with similar premises are Ray Bradbury's *Fahrenheit 451*, Simon & Schuster, 2011, and the entire cyberpunk literary corpus. For studies of dystopian literature, see Tom Moylan and Raffaella Baccolini (eds.), *Dark Horizons: Science Fiction and the Dystopian Imagination*, Routledge, 2003; Kollar, Dora, *Dystopian Writing in the Twentieth Century: A Comparative Analysis of Aldous Huxley's Brave New World and Ray Bradbury's Fahrenheit 451*, VDM Verlag, 2008

that Underhill is not allowed to open a door using his own hands. The Prime Directive function "to serve and obey, and guard men from harm" requires the mechanical entities to perform all duties of humankind regardless of the level of risk. As such, humanity falls back into comfort, unable to do any meaningful task. As the robots take over society, Mr. Sledge, having built in some form of initial protection, locks himself in a room in Underhill's house and devises a plan to "liquidate" the robots. At Underhill's request, the scientist tells the creation story of the humanoids. Sledge says that while he was a scientist on a distant planet in Wing IV, his intention was to build the perfect machine, "I was altogether too successful" but "rather ignorant, I'm afraid, of life and politics and war – of nearly everything, I suppose, except atomic theory" (*Ibid.*, p. 37). He goes on to explain that

> "I had too much faith in facts, I suppose, and too little in men. I mistrusted emotion, because I had no time for anything but science. I remember being swept along with a fad for general semantics. I wanted to apply the scientific method to every situation, and reduce all experiences to formula. I'm afraid, I was pretty impatient with human ignorance and error, and I thought that science alone could make the perfect world" (*Ibid.*, p. 37).

With utopian idealism, Sledge achieved the opposite. His inventions applied the Prime Directive to the letter without any form of flexibility. Initially, however, the robots showed great potential. They have rebuilt a ruined planet inhabited by people. As Sledge comments in a fashion similar to the dystopian language of *We* or *1984*, "I thought I had found the end of war and crime, of poverty and inequality, of human blundering and resulting human pain" (*Ibid.*, p. 42). But the cost was that "nicotine was disapproved. Alcohol was rationed. Drugs were forbidden. Sex was carefully supervised. Even suicide was clearly contradictory to the Prime

Directive – and the humanoids had learned to keep all possible lethal instruments out of reach" (*Ibid.*, p. 45).

His plan to liquidate the robots fails. The humanoids have taken precautionary measures to ensure that their command and control center cannot be damaged. As Sledge finds himself in the hospital recovering, the robots tell Underhill that the inventor is, in fact, a mental health patient suffering from grand delusions of invention, claiming himself to be the builder of the robots. Underhill senses the lie and is amused at the idea that robots can invent nonsense.

The Prime Directive resembles the *newspeak* of dystopian novels desperately necessary to coerce individuals into moral obedience. Through it, Williamson suggests that the current anxiety of replacement cannot have its solution in meaningless formulations of groundless ethical development. He carefully explains the consequences of designing robots based on the principle of safety and shows how it can lead to another reiteration of totalitarian tendencies (in the name of safety).

Following the Second World War and aware of the tremendous destructive power of newfound technology, the cybernetic imagination turned to the increasingly indistinguishable line between real and unreal, and, with the advent of the Turing experiment and cybernetics, to human and nonhuman. Interestingly enough, while science proceeded to prove human emotions and mental processes as measurable and replicable through technological means, philosophy began to consider the Cartesian dualism and Idealism as outdated but, paradoxically, producing, at the same time, the media technologies to confirm that *esse est percipi* is the only viable principle for the future. In this context, throughout the 1950s, literary production was concerned with the destruction of the planet following a nuclear war. This is understandable since nuclear

weapons posed (and continue to do to this day) a far greater existential threat than any other technology before it. The literary imagination of this period conceived potential consequences underlying the inevitable destruction of the world such as powerful creatures whose mutation was permitted by nuclear experiments (i.e. Godzilla). The U.S. government considered using nuclear attacks in the Korean War and, a decade later, the Vietnam War broke the patience of Western Europeans and Americans. Furthermore, this fear of impending nuclear apocalypse culminated with the Mutual Assured Destruction agreement which represented the capitulation of international diplomacy to the potential effect of nuclear war. During this time, the United States established the Advanced Research Project Agency which would, eventually, develop the Internet.

In the context of the Cold War and the creation of the computer, then networked computing, pervasive media technology, the Counterculture, the cybernetic imagination has found new authors to continue the expectation of replacement, reformulated in newfound inspiration drawn directly from scientific speculations about computing and artificial intelligence. But while, previously, the idea of replacement was straightforwardly associated with the replacement of humans, now, the cybernetic imagination emphasized the instrumentality of humans, trust in computational systems, and set the parameters to doubt the ontological certainty of humanity.

Responding to the assimilation of the individual in the machine-state comes mostly from the Frankfurt School, especially from the essays of Herbert Marcuse (Marcuse, 1991), but also from Martin Heidegger (Heidegger, 1977) and others (Kaplan, 2009; Kadar & Toth, 2013). And yet, given their largely Marxist orientation, “alienation” and the transformation of people in “human resources” seems to be a critique of a capitalist-driven state-machine absorbing the

individual as a consumerist component, a Randian ideal man, objectivist and relentless entrepreneur, into its ramified class dependent (on resources) system; so, these authors would have little problem returning to a socialist driven state-machine. Concerned, perhaps too much, with notions of power, they employ terms such as “hegemony”, “exploitation” and “domination” coupled with the use of technology. As independent variables, technologies, in their view, are not neutral entities, but have “valuative stance” (Freenberg, 1996).

Later in the century, these concepts are deployed to inquire into the consequences of capitalist-driven technological development and tend to highlight more of a frustration with capitalism than with technology itself. The cybernetic imagination, as such, is employed for its aesthetics and dread of replacement, being promoted as an ornament of ideological persuasion. As it is with other literary genres in late twentieth and early twenty-first centuries, ideological propaganda finds persuasive formulations in the aesthetics of digital disaster of the human physical world (see section 2.7. Cyber Apocalypse of this chapter). It can be achieved if the mentality of the time is appropriated in terms of systems and, in the post-war twentieth century, systems thinking derived from cybernetics and computer sciences.

The idea of the computer, cyberspace, and later A.I., began to replace traditional interpretations of philosophy and literature with the idea of system and required methodologies of science. The mind, as such, became an independent system and its creative processes, including cultural expressions and literary productions, had to conform accordingly. And while this may be true, as the Quantum Relations Principle also posits, that approach dismisses mental phenomena and metaphysics as part of human knowledge. Accordingly, we see the idea of systems carrying the computational logic evolve from the theories or Norbert Wiener and the

computer prototype of "memex" transported to comparative literature, through Even-Zohar and, later, Moretti, in the field of digital humanities.

2.4. Systems Certainty: Wiener, Bush, and the Digital Humanities

In the immediate aftermath of the Second World War, Alan Turing established the foundation of artificial intelligence and, basically, the entire field of machine learning and its applications (Turing, 1950). The Enigma code of the Wehrmacht was successfully decrypted.[23] However, the success had deep implications for society and provided the cybernetic imagination with a renewed motivation to investigate the possible replacement of man by machines, of reality by its artificial simulation.

Norbert Wiener starts his popular book on cybernetics, *The Human Use of Human Beings: Cybernetics and Society*, by acknowledging the change in Newtonian physics: "this is no longer the dominant attitude in physics" (Wiener, 1956, p. 7). And what has changed, moving beyond a deterministic world, is that "in a probabilistic world we no longer deal with quantities and statements which concern a specific, real universe as a whole but ask instead questions which may find their answers in a large number of similar universes" (*Ibid*., p. 11). Wiener,

[23] For a history of decoding the German secret communications during the Second World War see Sebag-Montefiore, Hugh, *Enigma: The Battle for the Code*, Wiley, 2004. For a history of cryptography see Kahn, David, *The Code Breakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*, Scribner, 1996; Thomas Rid provides an accessible history of communication during time of warfare in the first chapter of *Rise of the Machines: A Cybernetic History*, W. W. Norton & Company, 2016; Peter W. Singer's *Wired for War: The Robotic Revolution and Conflict in the 21st Century*, Penguin Books, 2009, provides an overview of robots in warfare from the Great War to contemporary war. The concept of Revolution in Military Affairs refers to the changes in warfare influenced by advanced technology, see Gray, Colin S., *Strategy for Chaos: Revolutions in Military Affairs and the Evidence of History*, Routledge, 2004

furthermore, deploys the theory of probability of Josiah Willard Gibbs to stress out the foundations of his development of cybernetics. According to Wiener, the measure of probability is called entropy and "the characteristic tendency of entropy is to increase", therefore:

> "As entropy increases, the universe, and all closed systems in the universe tend naturally to deteriorate and lose their distinctiveness, to move from the least to the most probable state, from a state of organization and differentiation in which distinctions and forms exist, to a state of chaos and sameness. In Gibbs' universe order is least probable, chaos most probable. But while the universe as a whole, if needed there is a whole universe, tends to run down, there are local enclaves at large and in which there is a limited and temporary tendency for organization to increase. Life finds its home in some of these enclaves. It is with this point of view at its core that the new science of Cybernetics, began its development" (*Ibid*., p. 12).

However, before the Second World War, a Romanian scientist by the name of Stefan Odobleja produced a theory of cybernetics, he being the first to introduce the feed-back closed loop as a universal law (Odobleja, 1938-1939).[24] Mihai Draganescu writes that "[o]ut of all we know, nobody before him has had such a vision of the general role of the feedback in nature and society. Odobleja is the first one who tried to apply the feedback law (the law of reversibility) to as many domains as possible, in the main, to all the domains. In this way he delimitates himself from all the peculiar cases in which reverse connections have been revealed."[25] But the most important contribution that we are interested in here is the view that nature can be defined as units to be measured, and the interaction among them, communication, to be reproduced by machines. While this was still in the realm of possibility, mostly in fiction, before the Second

[24] For reasons only speculative, Odobleja remains obscure in the field of cybernetics.
[25] See Odobleja, Stefan, *Psihologia consonantist*a. Editura stiintifica si Enciclopedica, 1982, apud Nicolae Jurau, Two Specialists in Cybernetics: Stefan Odobleja and Norbert Weiner. Common and Different Features, available here: http://www.bu.edu/wcp/Papers/Comp/CompJurc.htm For an overview of Odobleja's cybernetics, see Stefan, Gheorghe, M., The Cybernetic View of Stefan Odobleja, in Mihai Draganescu (ed.), *Odobleja: between Ampère and Wiener*, Editura Academiei Romane, 1981

World War, science has introduced the imagination of computer systems into the forefront of technological progress.

Norbert Wiener was part of the Allied scientific effort during the war. As he explains in his most famous book, his initial contributions were largely overlooked by Vannevar Bush, the Director of the Office for Scientific Research leading thousands of American scientists preoccupied with weapon development among other things. Bush was, of course, kept updated with the latest advances in technology and scientific thought during the war, including the research of Turing. In May 1945, Bush published an article in *The Atlantic* titled, interestingly enough, "As We May Think." Now that the war is over, Bush asks the readers, and numerous scientists have to stop researching and producing weapons, "What are the scientist to do next?" Many, undoubtedly, have found their way back to their old laboratories and research projects, others have found new knowledge and methodologies to implement, new questions to answer. In the essay, just as Wiener points to Leibnitz as the intellectual godfather of his work, Bush reminds of Leibnitz's calculating machine as the unproduced prototype of typing devices, which remained an unfulfilled project due to the technical restrictions of his time. However, Bush notes, "the world has arrived at an age of cheap complex devices of great reliability; and something is bound to come of it" (Bush, 1945).

Bush goes on to explain the available media technology of 1945: recording on paper, photography, wax, magnetic disks, and so on. If all of these could be put together in a coherent system, "we may some day click off arguments on a machine with the same assurance that we now enter sales on a cash register. But the machine of logic will not look like a cash register, even of the streamlined model." Yet the device he imagined was not called a computer, but a "memex": "A memex is a device in which an individual stores all his books, records, and

communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory" (*Ibid.*). This definition reads as if from a late 19th century scientific romance story. But what it does is to show how science has transported the individual mind from philosophy into matter. The "memex" is supposed to be based on the processes of human thought, not to replicate it entirely, and Bush acknowledges that this may be impossible, but to learn from it. This device, in effect, recognizes that our reality is more informational than previously anticipated. However, it is a device that replicates through the ideas of Leibnitz's *monadology* the idealist philosophy of Bishop Berkeley, as nothing is truly outside of the informational properties that the machine is supposed to process. In addition, information can now be completely digitized and, as such, we can create a new reality accompanying our Newtonian one. Through the technology it involves, photography and moving pictures, sound and text, the memex proceeds significantly in the realm of the recently so-called New Media. From now on, as we can expect, the cybernetic imagination pays extra attention to the devices developed based on the memex and explores, in consequence, artificial realms and the future social impact of computers and intelligent machines.

The idea of the "memex" and the peculiarities of cybernetics have been gradually adopted by the humanities, sometimes not in a necessarily lucrative way. As the computational sciences began to investigate issues traditionally reserved for literature, philosophy, or history, the humanities have adopted their language (and imagination) to formulate methodologies increasingly resembling the epistemic patterns and expectations of science. While some disciplines concerned with matter were lucrative in this sense, such as computational chemistry, biology, earth and medical sciences in general, the adjustment of computational language and methods to the humanities finds limited usages, confined mostly to the physical properties of the

text.[26] Among the new disciplines formed accordingly are the digital humanities. At the end of the century, computer science appears to have taken over the field with prominent members of its community asking humanities academics to learn to code (Ransay, 2013). Of course, some scholars balance the argument by introducing computational methods to assist research in their fields,[27] but traditional literary analyses in this age are replaced by such methods[28] or regress into ideological shortcuts.[29]

---

[26] Nan Z. Da argues that the computational methods for literary studies are dysfunctional in principle and purpose. See Nan Z. Da, The Digital Humanities Debacle, in *The Chronicle for Higher Education*, March 27, 2019. Available here: https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986

[27] For example, Mihai Spariosu (ed.), I*ntercultural Conflict and Harmony in the Central European Borderlands: The Cases of Banat and Transylvania 1849-1939,* V&R Unipress, 2017, provides a collection of research on intercultural relations combining traditional historical and literary inquiry and various computational methods to explore conflict and harmony in borderland regions. Nigel Gilbert and Klaus Troitzsch provide the computational tools to assist in cross-disciplinary studies in *Simulation for the Social Scientist*, Open University Press 2005. The goal in these studies is not to replace traditional research of the humanities but to incorporate computational methods for analysis and interpretation. My own co-authored research falls into this category. See Jecan, Vlad, Meza, Radu Co-Citation Mapping and the Intercultural Dialogue of the Intellectual Communities in Arad and Timisoara (19th to early 20th centuries), in Spariosu, Mihai, I., (ed.), *Intercultural Conflict and Harmony in Central European Borderlands. The Cases of Banat and Transylvania 1849 – 1939*, V&R unipress, 2017, pp. 355 – 367 and Jecan, Vlad, Meza Radu, Concept Mapping of Ideological Positioning in Cultural and Political Periodicals in the Interbellum Cluj, in *Romanian Journal of Information Science and Technology*, vol. 16, no. 2-3, 2013, pp. 237-250

[28] Nan Z. Da provides a compelling case for the takeover of literary analysis by digital methods in Nan Z. Da, The Computational Case against Computational Literary Studies, *Critical Inquiry*, Vol. 45, No. 3, 2019 pp. 601-639. The ramification of quantitative analysis of literary works into multiple fields is suggestive of the disintegration of the traditional field literary studies and its infusion with the terminology and methods drawn from computational studies. The author notes that "there is a fundamental mismatch between the statistical tools that are used and the objects to which they are applied" (p. 601). Multiple subfields have appeared in this regard such as cultural analytics, see Lev Manovich, Cultural Analytics: Visualizing Cultural Patterns in the Era of "More Media", available here: http://manovich.net/content/04-projects/063-cultural-analytics-visualizing-cultural-patterns/60_article_2009.pdf or Lev Manovich, The Science of Culture? Social Computing, Digital Humanities and Cultural Analytics, in *Journal of Cultural Analytics*, May 23, 2016, available here: https://culturalanalytics.org/article/11060 other forms of quantitative methods, such as "algorithmic criticism", for literary studies are presented in Ray Siemens (ed.), *A Companion to Digital Literary Studies*, Wiley-Blackwell, 2013

[29] A Marxist perspective, Eagleton, Terry, *Marxism and Literary Criticism*, University of Carolina Press, 1976; an American conservative position is Mark Zunac (ed.), *Literature and the Conservative Ideal*, Lexington Books, 2016

Even-Zohar notes that “the positivistic collection of data, taken bona fide on empiricist grounds and analyzed on the basis of their material substance, has been replaced by a functional approach based on the analysis of *relations*” (Even-Zohar, 1979, p. 288). In this evaluation, systems may show how various collections of symbols, including metaphors, act in effect. In this sense, Even-Zohar proposes the term “polysystems” with the purpose “to make explicit the conception of the system as dynamic and heterogeneous in opposition to the synchronistic approach” (*Ibid.*, p. 290). A system, in this case, is a “closed net-of-relations" among a particular set of elements. The polysystem conceives of interrelations of multiple systems, ”on the multiplicity of intersections, and hence on the greater complexity of structuredness involved” (*Ibid.*, p. 291) as such, given the difficulties of operating an open ended system, “more room will be given for "entropy," which may be quantitatively higher due to the fact that more relations must be taken into account, and more than one center must be postulated for the system” (*Ibid.*, p. 291). Following up on the theory of polysystems and with additional research, Franco Moretti proposes the concept of “distant reading” together with methods for computational investigations of literary texts (Moretti 2007; 2013).

The new field of study emerging from the attempt to unite computational methods with the humanities was initially termed *humanities computing* and then became the digital humanities (Svensson, 2013). Nick Montfort discussed the implication of computing to literature coining the term “computational literature” and trying to understand their impact on literature. Besides noting “interactive fiction,” that is video games, and ”interactive drama,” which “focuses more on the rich simulation of characters and on generating dramatic situations via their interactions,” (Monfort, 2015, p. 216) Montfort points to a “poetry generator” which “can be seen as defining a certain set of possible poems or outputs” (*Ibid.*, p. 209). Observing the results

of other scholars in the field of generating poems, he explains that "their poem "I AM THAT I AM" is a list of *all* the possible permutations of that phrase, including those that are lexically the same, and so is 120 lines long. This poem is a prime example of a computationally generated text – created with an electronic computer, although it could have been created with more difficulty by hand as earlier permutation poems were. The poem is not random at all. It is *exhaustive*, providing every possibility in the order generated" (*Ibid.*, p. 209). And while this possibility has developed further since the text of Montfort (Randford, 2019), or the first AI-authored novel titled *Teens Wander Around a House,* (Kazemi, 2013) and additional work done by other companies (Berber, 2019), the pressing question is "why?" Why delegate the creative thinking of humanity to robots? The consequences of such an effort is unknown and remains speculative due to its originality. However, it is an ontological challenge to text itself and a replacement of creativity by computational methods. If a novel can be written by intelligent machines, it means that the text is reducible to its components rather than its meaning which, in turn, is inevitably deemed artificial and can further affect the creation of meaning in human life.

While the digital humanities promise useful tools to investigate massive amounts of literary texts, if there is even a need for such research, it also attempts to replace not only the methods of literary inquiry but the text itself with a literary replicant and a new scholar able to write code.[30] The text of *Teens Wander Around a House* is clumsy and lacks depth. For example,

[30] Ramsay, Stephen, Who's In and Who's Out, in Melissa Terras, Julianne Nyhan, Edward Vanhoutte (eds.), *Defining Digital Humanities*, Ashgate, 2013, pp. 243 – 247. The central concern of this question is whether "digital humanists" must know how to code. Ramsay asks directly "Do you have to know how to code? I am a tenured professor of digital humanities and I say "yes". So, if you come to my program, you're going to have to learn to do that eventually" (p. 240). In a different entry, *On Building*, Ramsay explains his answer: "I have devoted my life as a teacher to teaching other humanists how to code. I do that for the exact reason that others devote their lives to the study of Shakespeare or the American Civil War: because it's fascinating and soul charging" (p. 245). As he dismissed the collaboration of coders and non-coders for digital humanities projects as "nonsense", this explanation is, indeed, nonsense. It presumes that a digital humanist has to teach other humanists how to code in order to build projects for

it reads: “Kiah entered the library. Kiah tried and failed to open the lock on the desk. Philomena and Darby ran into each other in the master bedroom. Philomena played nervously with her phone in a successful bid to avoid talking to Darby” (Kazami, 2013, p. 159). But it is only a matter of time before style and nuance can be added to unleash an ”uncanny literary valley” that can have severe consequences, as the reluctance of the OpenAI team to release the GPT-2 model to the public shows. According to OpenAI, due to “concerns about malicious applications of the technology, we are not releasing the trained model;” instead the company opted to release” a much smaller model for researchers to experiment with” (OpenAI, 2019).

The overall impact of the digital humanities was to produce more and more sophisticated methods to arrive at similar, traditional conclusions. Thus, the play of replacement means that a literary work needs no longer to be understood, but only to be explained. “System certainty” in their application to the cybernetic imagination produced ontological uncertainty of things and people as the literary works in the next subsection show.

2.5. Laws of Robots and Ontological Uncertainty: Asimov and Dick

The first significant author to attempt to put a bit of order in the cybernetic imagination dealing with apocalyptic visions of robot takeover was Isaac Asimov. His literary output is

---

themselves. Of course, this is fine, but the necessity of coding for being a digital humanist is particularly problematic. Should a Comparative Literature student learn R or another coding language beside the multiple foreign languages that the scholar requires to perform studies in the field? It could be, however, feasible for Comparative Literature departments to set up a program that treats coding languages the same as languages and prepare a curriculum for transdisciplinary training in the field targeted at particular industry sectors, such as translation, or academic fields, which remain to be defined.

impressive, but he is most famous for establishing the "laws of robotics". In *Runaround,* Asimov formulated the Three Laws of Robotics:

> "1. A robot may not injure a human being, nor through inaction allow a human being to come to harm.
> 2. A robot must obey the 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" (Asimov, 1984a, p. 209)

Runaround explores a situation in which the robot, Speedy, is unable to complete a mission because doing so would put humans in danger, therefore breaking the first law. The robot runs around, and eventually the humans put themselves in danger to rescue the robot and retrieve the dangerous materials required for their research (*Ibid.*).

The "robot fiction" of Isaac Asimov returns repeatedly to these laws, testing them, and even breaking them to see whether the rules are a viable set of deontological direction in A.I. research. In addition to the three laws, Asimov added the Zeroth Law, a rule to be followed before the others are implemented. The Zeroth Laws reads as follows: "A machine may not harm humanity, or, by inaction, allow humanity to come to harm." While these laws sound good in principle, their implementation is problematic (Singer, 2009). Optimistic and well-intended, the laws of robots do not stand the test of historical fact. If behavioral information is to guide robots on proper moral behavior, information that is drawn from human behavior past and present, then they contradict the logical sequence enabling their epistemic development by default. Of course, this is true if we presume that an artificial being does not have a desire for moral action and does not have the complex psychological and emotional feedback humans have to pursue moral action.

"The Evitable Conflict " as Asimov himself suggests, shows that "no science fiction writer has greater faith in the potential of machine intelligence than Isaac Asimov. And no Asimov story dramatizes that faith more powerfully than "The Evitable Conflict"". Set around Dr. Susan Calvin of the U.S. Robots and Mechanical Men Corporation, a recurring protagonist in Asimov's fiction, it initially presents the "inevitable conflict" in history drawing from the political conflict in Europe between the houses of Habsburg and Valois-Bourbon, or between Protestant and Catholic Christianity. "It was the 'inevitable' that the sword decided. Except that it didn't. In England, a new industrialism was growing, and on the continent, a new nationalism" (Asimov, 1984b, p. 254). Indeed, Asimov suggests that what appears to be inevitable violent conflict can be avoided by social perceptions and change in culture. In this sense, the cybernetic imagination's discourse about the replacement of individuals, often calling for violent resistance on the part of humanity or action on the part of robots, is a matter of perception. But the actual solution imagined by Asimov is the invention of "positronic robots": "they came in time, and, with it and alongside it, interplanetary travel. So that it no longer seemed so important whether the world was Adam Smith or Karl Marx. Neither made very much sense under the new circumstances. Both had to adapt, and they ended in almost the same place" (*Ibid.*, p. 254). The suggestion here is that technological progress, particularly that of robots, can have irenic effects on traditional tensions. As such, the robot acts as an irenic agent balancing the agonistic contest in the power-driven mentality. The solution for this problem, according to the story, is to delegate the economic decisions to robots "that have the good of humanity at heart through overwhelming force of the First Law of Robotics. [...] And so the question of ownership of the means of production becomes obsolescent" (*Ibid.*, p. 255).

However, when the robots produce unexpected results in the story, the initial reaction is that it was a human error for not giving them the correct data to work with. After evaluating the economic status of each of the planetary regions, the Coordinator recalls the first law and reformulates it into what has become the Zeroth Law. Opposing the machine use of human labor and thus standing against the general delegation of work to robots due to unfair competition standards, the Society for Humanity argues that humanity has lost its own freedom to determine the future by choice. When Dr. Calvin points this out at the end of the story, the Coordinator replies that "it never had any, really. It was always at the mercy of economic and sociological forces it did not understand – at the whims of climate, and the fortunes of war. [...] and no one can stop them, since the Machines will deal with them as they are dealing with the Society – having, as they do, the greatest of weapons at their disposal, the absolute control of our economy" (*Ibid.*, p. 276). Calvin reacts in horror at such a thought. But the point Asimov is trying to make is that prosperity for humanity may lie in this solution.[31]

In an essay titled "The Robot as Enemy?" Asimov insists on implementing the laws of robotics in the design of future artificial intelligence. He also explains that the need for security will require more security and, therefore, more complex robots. Furthermore, the systems designed for robots to operate are hackable and, as such, pose additional threats. "But what if," Asimov asks regarding robots of war, "our *robot* weapons were to accidentally engage in "friendly fire" and wipe out American people, or even just American property?" (Asimov, 1986, p. 92). The reality of robots, as Asimov acknowledges, is that the laws are not intrinsic in their design but have to be added in the framework of their operating system. Therefore, engineers can

[31] Contemporary political campaigns seem to adjust to the problem of robots in the workforce. Calling it "the freedom dividend", presidential candidate Andrew Yang plans to give $1000 per month to citizens. Policy available here: https://www.yang2020.com/policies/the-freedom-dividend/

choose not to implement them, but Asimov has no doubt that eventually "we will come round to the Three Laws" (*Ibid.*, p. 92).

While Asimov is among the few optimists when it comes to the future of man and machine, the laws will perhaps never be implemented. If not impossible to do so, then it may certainly be irrelevant. Even if the world economy is run by robots giving humanity time, space, and resources to pursue other interests unhindered by the difficulties of labor and lack of resources, it does not guarantee the world a safer place without war. Agonistic play will probably find other ways to ensure violent competition among individuals, so that the robots will work peacefully in the background while humanity continues its wars in the forefront. Since Asimov, the development of intelligent machines seems to have forgotten the laws altogether and move instead towards achieving realistic aesthetic physical standards and graphical perfection in virtual reality. It appears that the cybernetic imagination continues to promote replacement. Accordingly, the problems of robots have moved more towards the "uncanny valley" than to assure peaceful cohabitation with intelligent machines. The tension of the play of replacement can still be found there and, while Asimov may have been influential in designing and promoting the laws, the problem of virtual reality and artificial intelligence are better explored by the fiction of Philip K. Dick. Asimov is concerned with the science and technology of robots, but Dick is interested in exploring the consequences of such technology on the human socio-economic, political, and personal level. The literature of Philip K. Dick, therefore, gives us a more nuanced look at the imagination and ethical consequences of robots than Asimov's does. So, it may be that the initial question of "why this fear of robots?" needs, perhaps, to be rephrased.

In 1977 at a science fiction gathering in Metz, France, Philip K. Dick made a number of astonishing statements much to the confusion of the audience. What he was saying back then is

being researched at the University of Oxford today and is at the forefront of contemporary discussions on our technological future:

> "I in my stories and novels sometimes write about counterfeit worlds. Semi-real worlds as well as deranged private worlds, inhabited often by just one person…. At no time did I have a theoretical or conscious explanation for my preoccupation with these pluriform pseudo-worlds, but now I think I understand. What I was sensing was the manifold of partially actualized realities lying tangent to what evidently is the most actualized one—the one that the majority of us, by consensus gentium, agree on."[32]

Dick's fiction follows two major questions: (1) What does it mean to be human? (2) What is reality? In trying to answer these questions, the author deploys a myriad of themes, from androids and devices that stimulate the sense of reality to the problems of counterfeit memories, forged past and present, psychiatric instabilities, drugs and alternate futures.

Philip K. Dick's fiction is situated at the meeting point of the two worlds: real and virtual. At this liminal point, Dick addresses questions of morality by placing human characters in moments of doubt about their own existence and anticipates the possible abuses of truly ubiquitous digital communication technologies pervading mind and senses by questionable characters placed in high ranking political offices or religious positions. Frequently, in such a context, Dick will invent a new religion that mirrors the necessity of adapting theology to contemporary needs.

The ability to unite people in one global mental connection of spiritual significance is explored, for example, in *Do Androids Dream of Electric Sheep?* In the novel, Mercerism provides a technologically backed religion elevating the spirits of the believers beyond themselves. It is a religion invented to signify human ability for irrationality because, even after

[32] A video recording of his keynote is available here: https://www.youtube.com/watch?v=RkaQUZFbJjE

Mercer, the founder, is exposed as a fraud, the belief survives. Douglas A. Mackey sums ups Dick's fiction as follows:

> "[T]he depth of characterization he developed in the mainstream experiments was incorporated into increasingly bizarre science-fictional situations, which began to rely heavily on the radical alterations of reality, both in states of consciousness and perceptions of the objective world" (Mackey, 1988, p. 3).[33]

However, with these literary experiments, Dick also manages to answer the questions he poses. In *Do Androids Dream of Electric Sheep?,* the answer to "what does it mean to be human?" is empathy while the answer to "what is reality?" is given in the posthumously published novel *The Transmigration of Timothy Archer*. In these novels, the play of replacement takes a central position as an ongoing agonistic tension between human and non-human, in *Do Androids Dream,* and reality, religion, mysticism, in *The Transmigration*. However, Dick does not simply place one winner over the other. Instead, he explores the question of replacement in instances where characters doubt their own moral choices. For example, in *Do Android Dream,* Rick Deckard begins to show empathy for androids when required to "retire" (kill) a gifted singer named Luba Luft simply for being artificial: "Empathy towards an artificial construct? He

[33] In this work, Mackey offers a survey of Philip K. Dick's literary works according to the decade in which they have been published as well as insights into his personal life that has, unquestionably, influenced his quest for reality and the true human. In his personal letters and conversations, Dick seems to have lived "in a psychedelic world" (Ibidem, p. 4). While throughout his writing career, it may have been the impression of his readers and critics that Dick was living in a world of his own making, hindsight, however, proves the contrary. As technology developed rapidly, and fully aware of the Turing Test and the viability of human-resembling artificial surrogates, Dick explored reality and the android in imaginary situations in which it would be difficult, if not almost impossible, to distinguish between real and artificial human, real and artificial reality. However, what appears to be missing in Philip K. Dick is a structured view of reality, a stable moral framework, and religious belief. The consequences of these tells the story of his life: nervous breakdowns, suicide attempts, excessive drug usage, "he combined surrealistic, ironic humor", Mackey writes, "with innocence and passionate sincerity. As a personality, he went to extremes: regarded as an artist, a thinker, a lover, or a mystic, he comes off as larger than life" (Ibidem, p. 5). Philip K. Dick truly "contained multitudes." And it was these multitudes constantly pressuring his artistic expression that drove him to the edge of sanity and original literature. Without a sound grip on the current world, Philip K. Dick live in possibility and, accordingly, his fiction creates worlds grounded in perception where characters seek structure and stability.

asked himself. Something that only pretends to be alive? But Luba Luft had seemed *genuinely* alive; it had not worn the aspect of a simulation" (Dick, 2017, p. 131).[34] This state of emotional confusion is found repeatedly in the novel and is put in the forefront of the reader in page one when Rick Deckard and his wife, Iran, awake from a dream. The Deckards start the day arguing about what to dial on a device able to give them the necessary emotions to survive to sunset. Iran does not find a suitable reason to dial pleasant emotions and instead opts to delve into depression and suffering. However, Rick persuades her to do the opposite, only to later find out that she did not conform. In this sense, we are immediately introduced to the overall theme of the novel:

[34] The empathy towards an android is a very sensitive problem in the world of Philip K Dick. Androids are incapable of empathy and as such are simple cold machines, furthermore, as Eric Carl Link observes, "the android is a metaphor for the dehumanized human, the human who has sacrificed empathy and adopted a mechanical, controlled, and controlling view of the self and others" (Link, 2010, p. 56), Following Dick's statement that "all life is a moral issue" it appears that androids would not be considered themselves moral issues. However, their actions, because they can resemble perfectly their human counterparts, lacking empathy (or rather *oikeiosis*) cause tremendous suffering to life. Nevertheless, androids seem to be necessary in the cybernetic imagination of Philip K. Dick to clarify what is human. The play of replacement here seems rather beneficial. In the attempt of the androids to mingle or replace humans, the actions of other characters reveal the essential characteristics of being truly human. This resembles the initial solution found in the story "Human Is" (1955). Commenting later on the story, Dick said that "this story states my early conclusion as to what is human. I have not really changed my view since I wrote this story, back in the Fifties. It's not what you look like, or what planet you were born on. It's how kind you are. The quality of kindness, to me, distinguishes us from rocks and sticks and metal, and will forever, whatever shape we take, wherever we go, whatever we become," (*Ibid.,* p. 57). In a sense, this relates to Dick's statement about androids in his essay "Man, Android and Machine" of 1975: "Made in a laboratory -- that aspect is not meaningful to me; the entire universe is one vast laboratory, and out of it come sly and cruel entities which smile as they reach out to shake hands. But their handshake is the grip of death, and their smile has the coldness of the grave." Thus, just like the origin of the individual or the fusion with technology into a cyborg, or even that being is "made in a laboratory" does not define the distinction between real human and nonhuman. What does, however, is empathy – *oikeiosis.* But empathy, as we see in the character of John Isidore, given to androids is problematic and, rather, immoral. Isidore is, at one point in the novel, listening to Roy Baty and other androids deciding whether to kill him or not. Thus, empathy is morally justified only for other humans who are able to share in the same universal language that connects them in a single idea of the whole, hence Mercerism. Thus, to be truly human must reflect the proper display of reason where emotions have their place and to be social. The social connection among individual is kept using empathy and, as such, the individual has a duty towards the social and is not a single island. This, of course, is rather a Stoic reading of *Do Androids Dream of Electric Sheep* and interpretation of the concept of empathy as *oikeiosis* in Philip K. Dick.

emotions as a defining characteristic of what it means to be human.[35] But it is not just a simple manifestation of emotions. In *Do Androids Dream*, the androids display all kind of emotions, including fear of death and pain. For Philip K. Dick, the display of emotions is not enough for real humanness. The Deckards have an emotion device allowing them to select the desired emotions and emotional state, a device that also connects to the religion of Mercerism, but in doing so humans would not be any different from androids. Therefore, *Do Android Dreams* suggests that *empathy* is the primary sign of being human. However, even the concept of empathy is not entirely enough. In the novel, Philip K. Dick points towards what the ancient

[35] In his past fiction, Philip K. Dick explored the argument for emotions as a defining trait of a human being rather than pure reason. However, *Do Androids Dream* was more influenced by his personal life. After a devastating divorce, Phil met Nancy. Emanuel Carrere notes that "Nancy was different", more loving and caring than Phil's ex-wife. "Nancy had shown him what it was like to be authentically human – tender, compassionate, vulnerable – his writing began to extol the glory of the human being" (Carrere, 1993, p. 131). And this project dominated Philip K. Dick's uncertainty about humanity. In doing so, he "had to define and flesh out the opposite of the human, which for Phil was not the animal or the thing but what he called the "simulacrum" - in other words, the robot" (*Ibid.*, p. 131). In this sense, Dick's cybernetic imagination became original, distancing itself from the mere robot takeover fantasies overwhelming the science fiction genre. In the 1950s, Asimov tried to redirect the cybernetic imagination by adding a deontology of robots, his famous robot laws, but without much success. Dick's fiction on the artificial human does not seek to provide a deontology, nor other specific ethics, but instead tries to understand what a human being in the inevitable future of the uncanny valley will be, when androids may fully resemble humans. Dick found reason alone to be unconvincing. As such, he set out to imagine androids capable of easily passing the Turing Test, but "he was not about to welcome the androids into the human community, as Turing said one would have to when a machine finally managed to pass his test." For this he added "the sort of cheap trick that spiritualists were known to drag out when pushed against the wall: he added a new criterion, another ability that a subject would have to demonstrate in order to qualify as human" (*Ibid.*, p. 135): empathy. "Phil, who liked to speak to God in Latin, preferred the term *caritas,"* but, as Carrere notes, "it came down to the same thing: respect for the Golden Rule, for the commandment to "love thy neighbor as thyself"; the capacity to put yourself in the other person's place, to desire his happiness, to suffer with him and, if necessary, in his stead" (*Ibid.*, p. 135). *Do Androids Dream*, therefore, written under this prescription, returns to the ancient Greek idea of *oikeiosis.* According to Andrew Erskine, "Oikeiosis is the relationship that exists between an animal and something else, such that there is an affinity between the two and the latter in some way belongs to the former. The first thing an animal has an oikeiosis to is itself; the resulting impulse to self-preservation leads it to pursue what belongs to it (*oikeia*) and reject what is harmful. This is worked into an account of the moral development of human beings; with the onset of reason what belongs to a human is the good for man,' living in conformity with nature'" (Erskine, 1992, pp. 77-79). In *Do Androids Dream*, the androids reflect the idea of self-preservation without concern for the preservation of human life nor contempt for life in general. They are entities operating selfishly in a world that is social.

Greeks would call *oikeiosis*, the ability to suffer with someone as if the pain and difficulty are our own and be ready to pay the ultimate sacrifice, if needed. The androids, however, act differently. They do not care for the well-being of one another nor for that of humans.[36]

In order to return to Earth, many of them have killed their human owners on Mars. To remove the androids from Earth, fearing a possible revolt and the replacement of humanity, local police stations hire bounty hunters who "retire" artificial humans. Rick Deckard is one of these hunters. In the introduction to the plot, we are also presented with an ethical dilemma that will resonate with the reader throughout the book: "Just those poor andys" (Dick, 2017, p. 4). As such, readers are indirectly asked to test their own humanness by empathizing with the androids and, with the introduction of John Isidore, a human "chickenhead" unable to pass the IQ test granting him access to emigration to Mars, who is, nevertheless, able to empathize profoundly with androids, becomes a mode of empathy attraction for readers.[37] Eventually, Rick Deckard

---

[36] For a detailed investigation into the Stoic concept of *Oikeiosis*, see Engberg-Pedersen, Troels, *The Stoic Theory of Oikeiosis*, Aarhus University Press, 1990. Oikeiosis in *Do Androids Dream* can be also observed under the idea of "all life is one; 'no man is an island' quoting John Donne (Dick, 2017, p. 137). In addition, in an essay written in 1976 titled "Man, Android and Machine", Dick adds "We mean, basically, someone who does not care about the fate which his fellow living creatures fall victim to; he stands detached, a spectator, acting out by his indifference John Donne's theorem that "No man is an island," but giving that theorem a twist: that which is a mental and moral island *is not a man". We* will return to this concept later.

[37] John R. Isidore is to represent the honest and weak human character. However, he is unable to lie, exhibits empathy towards man and machine, and is a devout follower of the religion of Mercerism. He offers to help Pris accommodate herself in the abandoned building, aids Roy and Irmgard Baty, and, overall, reflects the innocence of human life. For him, Mercerism is the true way of life providing him with a sound moral code to follow. As such, Mercerism puts order into his solitary life and enables him to cope with everyday life. Even after the religion founded by Wilbur Mercer is disproved on live television (Chapter18), he continues to believe. And this, according to Laurence A. Rickels, is something incomprehensible to the androids, "it doesn't matter that Mercer is revealed to be a deadbeat actor, the setting of his Passion a painted soundstage, or even that as far as Mercer himself is concerned there is no salvation", Rickles, Lawrence, A., *I Think I Am. Philip K. Dick*, University of Minnesota Press, 2010, p. 294. What androids fail to understand is precisely this irrationality of the human mind. Empathy, as a drive to suffer along the other, is not rational. The resulting concern for androids to protect only themselves is indicative of their lack of irrationality. And this is the element that betrays them: the inability of *oikeiosis*. In the story "Null-O" published in 1958, a group of suprarational humans disconnected from moral and cultural biases and incapable of feeling sorrow or compassion, or any other

himself will display *oikeiosis* to androids in his struggle to remain consistent with his conviction that they are not real life.

The novel presents several important moments pertinent to our discussion. It is obvious that Dick's concern remains within the bounds of replacement. Rick Deckard is a bounty hunter commissioned by the San Francisco Police Department to remove potential androids from its area of jurisdiction, but will eventually find himself confused as to the morality of his actions when his senses and emotions are influenced by the behavior and artistic potential of the androids. In any case, the initial preoccupation of Dick is with the distinction between android and human. He devises an updated version of the Turing Test meant to detect empathy, the Voigt-Kampff Empathy Test[38] Initially, "the Voigt Empathy Test was devised by the Pavlov Institute working in the Soviet Union. And no T-14 android – insofar, at least, as known – had managed to pass that particular test" (*Ibid.*, p. 28). The problem is, however, that the new models on the streets are the advanced Nexus-6 that, as Deckard reflects, "surpass several classes of human specials in terms of intelligence," but with the updated Voigt-Kampff test it is possible to detect "the empathic gift [that] blurred the boundaries between hunter and victim, between the successful and the defeated. As in the fusion with Mercer, everyone ascended together or, when the cycle had come to an end, fell together into the trough of the tomb world" (*Ibid.*, p. 29).

---

human emotion, design the destruction of the earth with ever increasing weapon sophistication until the universe itself is destroyed. The story would suggest that pure reason leads to complete destruction and what makes us human is the emotionally dependent moral investigation we perform into deciding on the outcome of actions as to cause no harm to other human or intelligent beings. In an interview, Philip K Dick stated that "all life is a moral issue", see the collection of interviews published in Rickman, Gregg, *Philip K. Dick: In his own Words*, Fragments West, 1988

[38] Another test administered for android detection that appears later in the novel is the Boneli Reflex Arc Test.

But the first attempt of administering the test almost fails. Deckard is called by the Rosen Association in Seattle, the corporation producing the androids, to verify the human eligibility of Rachel Rosen. After a series of questions designed to trigger an empathic response measured by the device used for the test produce questionable results, Deckard asks Rachel for his briefcase:

> "Nice, isn't it? Department issue.
> Well, well, Rachel said remotely.
> Babyhide, Rick said. He stroked the black leather surface of the briefcase. "One hundred percent genuine human babyhide." He saw the two dial indicators gyrate frantically. But only after a pause. The reaction had come, but too late. He knew the reaction period down to a fraction of a second, the correct reaction period; there should have been none. "Thanks, Miss Rosen" (*Ibid.*, p. 56).

Thus, Deckard concludes that Rachel must be a Nexus-6. Unlike the Golem in Meyrink which takes a shadowy presence in the background, or in Frankenstein where the creature is purposely on a destructive rampage, or in Forster where the Machine is an all-controlling device designing the reality of users, or in Capek where the artificial human is directly set on world domination, the android in Philip K. Dick is a mirror of the human, a simulacrum of the potential destructive tendencies of individuals bound solely to the mind as a mechanical metaphor. The android in present and active in the world utilizing the creative abilities of the human. As such, Dick asks us to reconsider our own humanity.

Deckard is sent to retire several androids, including a peculiar type whose name is Luba Luft. Luft, most likely an advanced Nexus-6 type, is capable of artistic interpretation. Furthermore, Luft attempts to cause doubts about Deckard's own humanity and almost succeeds. Meeting Luft in the dressing room while the android prepares for an upcoming live performance, Deckard prepares to give her the Voigt-Kampff test:

"It's necessary." He got out the Voigt-Kampff instruments, began setting them up.
"An IQ test?"
"No. Empathy." […]
"An android," he said, "doesn't care what happens to another android. That's one of the indications we look for."
"Then," Miss Luft said, "you must be an android."
That sopped him; he stared at her. […]
"Have you taken it [the test]?"
"Yes." He nodded. "A long, long time ago; when I first started with the department."
"Maybe that's a false memory. Don't androids sometimes go around with false memories?" (*Ibid.*, p. 94)[39]

It is in the moments preceding this meeting that Rick Deckard starts to question the moral justification for retiring androids. His test is cut short by a police officer who takes him to the department. There, he is led to believe, quite convincingly, that his own memories might indeed be implanted and is, therefore, an android himself. As it would be expected, and since androids

[39] The false memory problem is a recurring theme in Philip K. Dick. As his remark at the 1977 science fiction convention emphasizes, it is a belief personal to the author according to which he navigates the world. In his fiction, however, perhaps the most important work on false memories creating one possible version of reality is in *The Man in the High Castle*. In this novel, Dick creates an alternate history of the result of the Second World War. The United States is divided among Imperial Japan and Nazi Germany, the former occupying the West Coast while the latter the East Coast. The man in the high castle is an individual who holds several war-time films and recordings that present the actual ending of the war, that is our own version of reality. The emphasis in the story is about media products capable of changing the perception of reality in the contemporary world. For Dick, what we experience with our cognitive abilities does not reflect reality, nor is it true for our senses. The novel suggests that the media we consume influences the way we conceive the past, the present, and the future. For Dick, such objects seem to carry an intellectual component that defines our construction of identity, our morality, our ideas of the present and, overall, that which is real. Indeed, he also points out that it is virtually impossible to truly contest reality unless we receive minor hits and, as the Matrix encourages "follow the White Rabbit". In 1950, Harold Innis published *Empire and Communications* (2007) in which he distinguishes about two types of media objects into time-binding and space-binding categories. The first are media objects designed to carry the message of state power in time. Impressive buildings are being constructed in this sense such as the Pyramids, Colosseum, or the Statue of Liberty. The space-binding media are objects capable of transporting cultural, ideological, and power messages easily in space. Coins, for example, are able to carry the idea of the founding of a nation state in space. In the Roman Empire, for example, they were used to familiarize the population with the emperor. As such, Philip K. Dick contests the very basics of our formulation of reality and suggests, in addition, the possibility of parallel worlds that may intersect in different points where artifacts from one can be introduced to another reality. His statement at the conference in France marks this possibility. Furthermore, it would be possible, according to Dick, to live in an entire counterfeit reality, a simulation, an entire artificial reality.

have the possibility to question their own ontological certainty, Deckard begins to doubt himself. He is not familiar with Boneli Reflex-Arc Test and his interlocutors are unfamiliar with the Voigt-Kampff test.

When Deckard is allowed to call home, the vidphone (a device allowing for video communication) displays a different person than Iran. Garland, an officer interrogating him, tells another bounty hunter present, Phil Resch, that "This man – or android – Rick Deckard, comes to us from a phantom, hallucinatory, nonexistent police agency allegedly operating out of the old departmental headquarters in Lombard. He's never heard of us and we've never heard of him" (*Ibid.*, p. 110). Garland was on Deckard's bounty list. It turns out that the androids have come up with a sophisticated plan to keep themselves safe. Garland is an android, but this is revealed only after the reader experiences the ontological angst of Deckard. As the novel suggests, the androids arriving from Mars may kill existing human beings and assume their identity. Phil Resch, the other bounty hunter, explains to Rick: "Then at one time an authentic Garland existed. And somewhere along the way got replaced" (*Ibid.*, p. 127). The androids, even the advanced models, follow the strict path of self-preservation without concern for other beings nor for other androids. Therefore, the Voigt-Kampff test and the experience of Deckard does not simply test for empathy, but for the more complex concept *oikeiosis*, not mentioned in the novel.[40]

Again, Dick turns to false memories to question the reality of the protagonist. But he goes one step further. After returning to retire Luba Luft, Deckard finds himself questioning the

[40] It is unclear whether Philip K. Dick was aware of the concept. Nevertheless, the interrogation techniques and references towards caring for the fate of another android or human, the ability for joint suffering, is firmly underlined in the novel and that seems to be the distinguishing line between human and android.

reason to kill an android capable of singing *The Magic Flute*, Rick's favorite opera.[41] "She was a wonderful singer. The planet could have used her. This is insane" (*Ibid.*, p. 126). At this point, Deckard is almost able to acknowledge the humanity of the androids and his doubt is further revealed when Phil Resch turns out to be not an android, but a human who likes to kill for the thrill of killing itself. As such, we are faced with the question of why should Resch be able to claim life while a human-resembling artificial entity capable of singing opera, therefore creating beauty, is considered but a mere simulacrum of human life and, consequently, not alive.

The other advanced androids have fled Mars guided by an android called "Roy Baty" which is described to have "an aggressive, assertive air of *ersatz* authority. Given to mystical preoccupations, this android proposed the group escape attempt, underwriting it ideologically with a pretentious fiction as to the sacredness of so-called android "life" (*Ibid.*, p. 169). The concept of "ersatz" (replacement) is recurring in *Do Androids Dream*. We have Garland who was replaced by an android, the artificial sheep of the Deckards who is replaced by the organic goat after Rick gathers the funds to afford such a real animal, and then, we have the artificial

[41] The connection that is established here and responsible for blurring the lines of android and human is an emotional one. Deckard enjoys *The Magic Flute* and is emotionally invested in its optimal performance. Luba Luft can do exactly that. As he reflected on the problem, Deckard tells himself "it could be an anomaly, something for instance to do with my feelings for *The Magic Flute*. And for Luba's voice, in fact her career as a while" (Dick, 2017, p. 132). To resolve this problem, Phil Resch advises Deckard to pursue sexual intercorse with an android. Resch tells him "If it's love toward a woman or an android imitation, it's sex. Wake up and face yourself, Deckard. You wanted to go to bed with a female type of android – nothing more, nothing less" and if he wants to solve the problem Deckard should "Go to bed with her first -" "-and then kill her" (*Ibid.*, p. 133). Eventually, Deckard will have sex with Rachel Rosen, a Nexus-6 type female android, ready to help him retire additional three more advanced androids. However, Deckard is unable to kill her, but she, in turn, will kill his real animal. Owning a real living animal instead of an artificial one in the post-apocalyptic world of *Do Androids Dream* is a symbol of social status. However, it is also a symbol suggesting the innate desire for the real in a world where the senses and mental faculties are constantly tested and confused. In Richard K. Morgan's *Altered Carbon*, (Del Rey, 2002) winner of the Philip K. Dick Award, the symbol of the real in a similar ontologically uncertain context is "real death." This appears to suggest that in the cybernetic imagination the constant struggle of replacement between real and artificial is never-ending. Replacement always finds a way to return in a different formulation.

toad at the end of the novel confused with the real. In this sense, the idea of replacement is fully articulated in *Do Androids Dream*. Furthermore, we have John Isodore, the "chickenhead", empathizing with the fleeing androids Roy Baty, his wife Irmgard Baty, and Pris Stratton.[42] Pris and Rachel look identical, but are, nevertheless, different. However, this difference is unclear. Deckard knows that both are androids and, according to Dick's own words in an essay on the subject of androids, Rachel seems to be more prone to deceive being human than Pris: 'to deceive us in a cruel way, to cause us to think it to be one of ourselves."

Isidore offers to help, but Roy Baty suggests killing him and retreat somewhere else. However, the other two androids decide to make their stand against Deckard in the old deserted apartment building where John Isidore lives. Deckard retires the androids and returns home to his wife ready to abandon his bounty hunting work. The ending of the novel is quite interesting. It is, as Douglas A. Mackey notes, "pure irony" (Mackey, 1988**,** p. 91).[43] Deckard finds a toad, a

---

[42] In 1976, Philip K. Dick acknowledged the connection of Pris Stratton to Pris Frauenzimmer in *We Can Build You*. Machines, or androids, according to Dick are "fierce cold things" and "their behavior frightens me, especially if it imitates human behavior so well that I get the uncomfortable sense that these things are trying to pass themselves off as humans but are not." Furthermore, "[By android] I mean a thing somehow generated to deceive us in a cruel way, to cause us to think it to be one of ourselves." [add reference] But while Deckard kills Pris in *Do Androids Dream*, he doesn't retire Rachel. We would be inclined to believe that because she has offered to help him, Rachel displays *oikeiosis*, and, as such, would pass Deckard's instinctive disposition to not kill her. However, Deckard discovers that she was sent by the corporation to analyze the retirement of advanced androids in order for them to build better models. However, Deckard is, at this point, changed by experience. After retiring Luba Luft, he does not want to be a bounty hunter or a policeman and tells himself that the last three androids will conclude his career.

[43] Deckard finds himself in a similar experience to that of Wilbur Mercer. And, at one point, he declares that he is Mercer. However, the role of Mercerism in this novel seems to have an additional purpose less theologically inclined. The popular TV show in *Do Androids Dream* is the Buster Friendly Show. In chapter 18, Buster and his team expose Mercerism as a fraud. "Wilbur Mercer is not human, does not in fact exist" (p. 192). The effect of this is that empathy does no longer play a vital part in the distinction between man-machine. The android Irmgard reflects on the consequence of it: "Isn't it a way of proving that humans can do something we can't do? Because without the Mercer experience we just have your *word* that you feel this empathy business, this shared, group thing" (p. 193). To which Roy Baty explains "It's done. Buster said it out loud, and nearly every human in the system heard him say it. 'Mercerism is a swindle.' The whole experience of empathy is a swindle" (p. 193). Pris, turning to Isidore said "Buster is one of us", "An android," Irmgard explained. "And nobody knows. No humans, I mean." (p. 194). As

species thought extinct. Iran, however, discovers that it is artificial and while Rick retreats disappointed, his wife orders artificial flies for it telling the clerk that "I want it to work perfectly. My husband is devoted to it" (Dick, 2017, p. 224).

Dick's answer to the question of "what does it mean to be human?" is action based on empathy, in a way that extends to the suffering of individuals and enables personal sacrifice. The fear of replacement plays a major role in his fiction. Androids are imagined having the potential to become practically unrecognizable in the public space. As such, the fear is of not knowing who is who. However, the individualistic, self-preservation attitude of androids suggests the fear of technology able to disrupt the social aspect of being human.

The question of reality appears in several, if not most of the works of Philip K. Dick. Objects that we take ontologically certain manifest themselves impalpable in the minds of characters or through the devices of the plot. Rather than them being physical entities, Dick explores objects, events, entities, people, through their informational properties without providing an alternative, functional view of reality. What is real exists in possibility while the sensible reality is often uncertain or counterfeit. Instead of using pure fictional elements of science fiction, Dick resorts to employing elements of our real world, such as politics, historical artefacts, mental disorders, in order to argue for its unreality.

In this perpetual confusion that becomes obvious in our contemporary world, the answer to the question of reality is not reality itself. *The Transmigration of Timothy Archer* suggests that

---

such, in *Do Android Dream*, the android is actively working to undermine the human belief system in order to achieve his own protection regardless of how many humans have to suffer in the process. Buster Friendly, an android, is in the perfect position to do so and as such he is a device used by Philip K. Dick to recall the problem of counterfeit ideas and objects explored in other novels. In the world of Philip K. Dick, androids exploit humanity and humanity exploits humanity to such an extent that the line between human and machine is extinct at all levels, in the physical realm and in the intelligible reality.

rather than being concerned with reality, “it is what attitude, idealistic or realistic, one should take toward life” (Mackey, 1988, p. 125). This attitude is divided between Tim and Angel Archer. Tim is idealistic, a moral philosopher of sorts, who becomes involved in the occult after questioning Christianity and seeing direct revelation. Angel Archer, the narrator of the story and Timothy’s daughter-in-law, is a realist seeking worldly wisdom, but turns gradually to Christianity. Here, the question of reality is represented as a choice of approach towards it.

The play of replacement as a contest between the real and the artificial is thoroughly reflected in Dick’s fiction. It represents, in addition, the culmination of a mentality confused between science and the humanities, unable to comprehend the interlinking of mind, matter, and meaning, and treating them accordingly as separate objects or categories for distinct and independent databases of knowledge. The result is an incomprehensible position bridging the databases. As such, the cybernetic imagination is only capable of renouncing artificial objects or to adjust them to current reality and contemporary technological concerns as “counterfeit.”

Furthermore, Dick’s literature is indicative of the wider problem of humanity coping with its reflection in the mirror when, on the one hand, artificial entities are created to resemble human likeness and intellectual creativity without a firm, moral reason to do so, and, on the other hand, the media technology and software production has the potential of creating “fake” objects able to transport ideological messages directly into the minds of the users. As such, reality is ambiguous not because of the acknowledgement that it rests in the interlinking databases, but because of the lack of it. Ethics, in turn, are confined to relativism while, paradoxically, demanding strictness without addressing the global impact of technology. However, concomitantly, a new culture embracing the potential of computers was developing in the United State. It was a culture capable of operating and embracing the artificiality of cyberspace and the

influence of computers in the offline world without militating for real over the unreal, nor for the organic over the artificial. This culture, firmly grounded in the use of computers, challenged traditional notions of political power by revealing the outstanding technological resources available to the masses, and will be addressed in the next subsection.

2.6. Systems Uncertainty: Cyberpunk and the Hacker Culture

Hacking during the 1970s is closely associated with the counterculture movement and appears to recollect the idea of the state as a machine. Later in the century, some political scientists would attribute the metaphor of software to the workings of the state.[44] Early hackers, however, were concerned with liberating the primary function of a machine state, communication technology. Abbie Hoffmann, Timothy Leary, and Emmanuel Goldstein have contributed significantly to 21st century political activism in the digital space known as hacktivism.[45] The culture draws precisely from the informational properties of objects, therefore, reconfiguring physical objects into digital units and producing digital objects for offline implementation. In this sense, the hacker culture takes position in the liminal space between the digital and the physical and, increasingly, transforms cyberspace into a liminal playground that negotiates views on both artificial and physical realities. In doing so, the hacker culture

[44] For example, Jeremy Everard stated that "the state in many aspects is like a piece of software – it seems stable enough while the power is on and it has not run into a major bug yet, but interrupt the power supply, or corrupt it, and it falls apart with startling rapidity" (Everard, 2000, p. 4).
[45] Abbie Hoffman launched the Youth International Party Newsletter in early 1970s, a bulletin instructing people on methods to use technology for free, especially communication technology. Timothy Leary explained the hacker mind, behavior, and expectations in *Chaos and Cyber Culture*, Berkeley, California: Ronin Publishing, 2014, 20th anniversary edition. Emmanuel Goldstein is the founder of the *2600 Magazine: The Hacker* Quarterly an active publication since early 1980s.

transcends both realms and is involved actively in most processes of human life--indeed, with the continuous dependence on the digital realm, in all aspects of human life.

In addition, throughout the 1980s and 1990s, the hacker culture has produced its own moral positioning, grounded in virtue ethics. Without a set of viable rules to follow or notions about achieving good for most people, the hacker culture departs from utilitarian or deontological moral systems and, inevitably, develops a form of digital virtue defined by skill and individual conduct. The cyberpunk literary 'movement' seems to encapsulate this. Cyberpunk fiction is always set in technological dystopia where the blending of real life and artificial beings is almost complete. Usually, powerful corporations or tyrannical governments possessing vast technological resources control individuals in both body and mind. In this context, cyberpunk protagonists, antiheroes, try to make sense of the world not by replacing it with an organic version of itself, but to seek out "the right thing" where existence is 'high tech, low life'. From the outside, this may seem odd, especially when protagonists display distance from the human body, are usually addicted to various psychiatric stimulants, or seek out exotic experiences in virtual reality.

It might look misguided and utterly deceptive to expect suicidal or psychotic individuals who seek artificial thrills by jacking in cyberspace to offer examples of sound virtue ethics. But the characters themselves represent the depravity of a future impaired by and because of its own abject moral reasoning. For example, in *Altered Carbon* by Richard K. Morgan, one of the latest admirable cyberpunk novels written in the 21st century, humanity has found a way to live forever by transferring consciousness in a stack that will animate a body, old or new. While the debate on whether this is truly life lies in the background of the story and left to the reader to conclude, the hacker knows that consciousness is devoid of its organic imperative, the human body,

rendered in code and, therefore, hackable – modifiable to suit the needs of whoever has the tools and skills to do so. Morgan seems to explore this in the parameters of trauma caused in virtual reality where a lovable A.I. named Poe helps heal a damaged human, or when the protagonist, Takeshi Kovacs, finds the ability to endure severe torture and overcome his torturers in the virtual and the physical realms.

The only outcome of the search for the real in a dystopian world unable to die animated by the play of replacement desperately seeking out a return of the organic, the irreplaceable, is to find and produce 'real' death. Sometimes, A.I. seems to be more human than humans, not because of the physical and mental resemblance of people, but through moral action. A.I., it is imagined, can be made to resemble moral actions by sound design of the acting entity. Such is the case of Poe. However, most of the time, A.I. and virtual reality are portrayed as antagonistic to their counterparts, as we have seen. Cyberpunk takes a different turn. Sometimes it glorifies the unreal at the expense of the real [add novels], other times it battles the unreal for the real following the play of replacement (*Matrix* being the epitome of this fight). But, barring these exceptions, it tries to emphasize the need for unrestricted creative freedom for people and to pursue moral action in ambiguity and ontological uncertainty. In doing so, the hacker culture and cyberpunk literature seem to oppose restrictions and control desired by economic and political entities to ensure that 'information wants to be free.'[46] In this sense, the hacker culture developed

[46] The statement "information wants to be free" formulated by Steven Levy stands as a metaphor for hacker culture. While it is usually perceived to be applying only to political restrictions of the contemporary world, as it certainly is the case, it also takes an ontological stance facilitating a principle of hacker digital virtue. If the shift in civilization occurs when the view of reality is contested and, according to Luciano Floridi, we move towards an information reality, then this principle becomes the most important element in ensuring human freedom. However, total freedom will enable the production of "fake" information and "deep fake" media productions to carry ideological preferences within the world. But given that everything becomes information, and the apparent impossibility to regulate information while avoiding totalitarian tendencies, it becomes a matter of individual responsibility to act morality, therefore, a key element of digital virtue and free play.

along three routes: (1) dissemination of technical information to bypass communication restrictions; (2) software production to ensure freedom of communication; and (3) mobilization of users in structured online protests.[47] Cyberpunk literature, in turn, adopted these routes and transported them in imaginary high technological worlds while also navigating complex metaphysical questions.

Abbie Hoffmann founded the Youth International Party Line, a newsletter that gradually attracted significant readership, to instruct people on using communication technology to their advantage. Of course, as the name of the publication suggests, Hoffmann came from a socialist background, flirting with communism and revolution. In the pages of the newsletter, we can read "how to" guides to improvised technology and ways to reconfigure a device to serve a different purpose than the one intended for, as the classic definition of hacking explains. Thus, the spirit of hacking was born in the counterculture of the 1960s and 1970s.

Concomitantly, the internet was developed, and by the 1980s individuals were able to purchase home computers and join hobby clubs such as the Homebrew Computer Club. In this club, and in other less popular ones, highly passionate computer developers "brewed" their own devices and cultivated a spirit of entrepreneurship responsible for many innovations that we use today. Among them were Steve Jobs or Steve Wozniak, as well as other hackers toying with devices capable of reproducing the 2600 frequency tone needed for phone calls. Emmanuel Goldstein introduced 2600 Magazine, a publication similar to the one intended by Hoffmann, although showing a variety of political views, meant to disseminate information about hackers,

[47] I have dedicated a doctoral dissertation on documenting the hacker culture, see Jecan, Vlad, *The Play of Hacking and the Political Values of the Hacker Culture*, Unpublished Doctoral Dissertation, University of Babes-Bolyai, 2016; Chapter 4: The Play of Hacking is based largely on that research. The dissertation can be consulted here: http://vladjecan.com/jecan-teza.pdf

hacking techniques, devices, suitable software, and meeting announcements. The DEFCON conference was established and still organized annually, offering hackers a space to share information. Phrack Magazine, also a hacker newsletter, established in the 1980s, was more tech-oriented than 2600 and featured full profiles of prolific members of the hacker community.

Several hacker groups were formed with a political agenda. Among the ones with the most colorful name is Cult of the Dead Cow established in early 1980s with the mission to produce software to combat increased state control of cyberspace and the monopolization of computer products by corporations such as Microsoft. Joining in politics, was the much more intellectually refined Electronic Disturbance Theater whose members published books disseminating their ideas and reviewing social consequences of digital technology.[48] In addition, the group released software that would allow users to engage in a virtual mass protests similar to offline sit-ins. The software, easy to use, enabled people around the globe to act their support out by joining in mass data bombardment of targeted servers known as denial of service operations, usually of government or corporations. These software, as well as others, have paved the way for recent hacktivist campaigns.

The similarity attributed to virtual sit-ins by their offline counterpart, masses of bodies in the streets blocking entrances or access to building, implies that in the real world is now a blending of offline and online activities. The political protests organized online impact the offline world and vice versa. As a result, the separation between cyberspace, virtual reality, and

[48] The Electronic Disturbance Theater is the digital activist wing of the Critical Arts Ensemble. The latter has published numerous books on the subject of digital technology and political activism. Among them are *Digital Resistance,* Autonomedia, 2001, detailing the concept of electronic civil disobedience which received its own book in 1996 published also by Autonomedia, *Electronic Civil Disobedience and Other Unpopular Ideas.* In addition, the group spearheaded by Ricardo Dominguez published *Flesh Machines: Cyborgs, Designer Babies, and New Eugenic Consciousness,* Autonomedia, 1998, and, more recently, *Aesthetics, Necropolitics, Environmental Struggle,* Autonomedia, 2018

physical reality is diminished significantly. The former is, especially today, fully integrated in the latter.

Most important, however, about the hacker culture is that it marks the actual bridge between the digital and the physical worlds. Cyberpunk literature, in addition, is continuously imagining the consequences when this bridge is no longer needed and everything is information, that is, reality is according to *esse est percipi*. The mentality developed by the hacker culture over decades does not intend to replace either the real nor the digital, but to play with digital objects and reconfigure them; in doing so, it redesigns reality as well. It was by operating in the uncharted, endless, space of cyberspace, where morality is a foreign thing from another world, that hackers developed ethics according to personal conduct in the digital realm. I call these ethics "digital virtue" (to be explored in the final chapter of this thesis) because it depends exclusively on the conduct of a hacker capable of destroying or building at will in cyberspace. As such, in hacking one can observe (1) a move away from the play of replacement and the agonistic mentality it harbors resulting in (2) a balance of the real and the digital.

The 1980s also saw another computer-based industry rise that would further diminish the distinction between real and unreal, digital and physical objects: the gaming industry. From the release of Pong in 1972, a simple game consisting of two lines and a dot bouncing between them, to the latest games in the 21st century, the way we conceive reality, storytelling and moral certainties, have been influenced by this industry. Play has, of course, incorporated video games.

The experience of video games fully embodies the indistinctiveness of virtual and physical realities. Opposite from the outrage of violent games, the concern of "screen time" for children, and fears of missing out opportunities in the real world, which promote the idea of

replacement, lies play in its irenic manifestation. Play remains the desire to overcome unnecessary obstacles, but the elements of play nowadays incorporate the virtual rather than dismiss it. Experiences in the offline world are supplemented by those in the virtual world. Meaningful play in the twenty-first century is as important to obtain an imaginary flaming sword and carefully designed armor as it was to build imaginary buildings and conceive unnecessary rules for achieving fictional goals in the offline world.

The economy has adapted as well. The current earnings of "e-sports" are higher than those generated by traditional leagues such as NFL, NBA, and others. Play does not discriminate between real and unreal objects, rules, and spaces, because within the ludic experiment all objects are ontologically flexible. And it is in this sense that play is irenic. The problem of replacement, however, persists in corporate strategies to keep players bound to the virtual world by implementing psychological methods of emotional persuasiveness found in gambling that render many popular games rather immoral. As such, the player becomes dependent on the virtual world not for play but for the emotional validation of rewards in a mental training process that reconfigures the individual's sense of value requiring her to return to the slot machine. All other elements of play become mere decorum and, as such, lose their moralizing value.

The idea of objects as (digital) information is common in cyberpunk literature. This conforms to the principle of perception expressed by Bishop Berkeley, on the one hand, and allows the bridge between physical reality and virtual reality, on the other hand. Artificial Intelligence operates on both realms while the cyborg, a blending of human and machine, serves as a metaphor for the increased human absorption into the technological realm. However, cyberpunk does not intend, at least not directly, to be too philosophical, like Gustav Meyrink or Philip K. Dick, but to portray human behavior within a technological society in which moral

certainty is extinct. In this setting, the only tool to operate the blending of the artificial world with the leftovers of a physical one is to understand the fabric that enables it: code. Hackers, usually highly skilled in manipulating and producing code, play an important role in these stories. They are mostly builders[49]rather than destroyers. Sometimes, however, they can be rebuilders[50] and usually find themselves involved in some sort of detective work. Something similar to detective work is required in order to understand the digital system or to comprehend the workings of a computer virus or worm, bent on destroying everything.

The aesthetic also plays a major part in the cyberpunk genre. It is usually raining in the hi tech, low life dystopia of the cyberpunk cybernetic imagination. Objects appear in their digital

[49] For example, Hiro Protagonist in Neal Stephenson's *Snow Crash,* Del Rey, 2000. William Gibson's *Neuromancer* shows the protagonist, Case, to be a hacker at home in the virtual world, craving for it rather than seeking to delete it.

[50] Pat Cadigan's novel (Cadigan, 1991) presents the Dataline, a digital network reminding of the internet that displays information on the multiscreen setups of the characters. The introduction of a new technology of "sockets", "a direct interface for input-output with manufactured neural nets, computers" (p. 174), facilitates the neural link with the global network. As in most cyberpunk novels, a greedy corporation tries to use the new technology to maximize profits. In the case of *Synners* that is the Diversification Inc., which quickly capitalizes on virtual entertainment. Among the characters, Visual Mark, a highly skilled artists with a high drug us leaves this body for the virtual world and becomes there a cybernetic entity capable of surviving in the artificial world. Cyberspace also hosts a curious anomaly, Dr. Art Fish, V.D. (Virus Doctor) which is an accident of a virus vaccine infused in artificial intelligence, although, as Adrian says "You can get it faster if you say Artie Fish" (p. 176). Indeed, as Rosa wonders "But Jesus, *Art Fish?* What's wrong with the good old names, like Frankenstein?" (p. 177). Art Fish is "a wonder and a revelation" and "a synthetic concert of intelligence in concert mode" (p. 387), a helpful artificial entity in solving the imminent destruction of cyberspace. Visual Mark's offline body suffers a stroke which is projected back in the system and causes it to malfunction; while in most popular cyberpunk stories, such as *Matrix,* the damage caused in virtual reality affects the offline body through the interpretation of harm by the mind. In *Synners,* the offline body influences the harmony of the online world. It was a" [r]eference: cerebral vascular accident. Only it's different this time. If it gets into the system and finds someone hooked in with the interface, it'll get them, too. You got that? A contagious stroke, a fucking virus, are you with me yet?" (p. 315). Crole F. Meyers observes that "theorizing that if on-line brain illness exists, then on-line therapy must be possible, the hackers decide to try to cure the system," Meyers, Carole, F., Synners, in Paul di Filippo (ed.), *Critical Survey of Science Fiction and Fantasy Literature,* Salem Press, 2017, p. 1142. Therefore, the hackers act out to protect the artificial world from trapping Visual Mark, employing in the process the help of the artificial intelligence Art Fish. In the novel, the hackers stand as a community and display real affection for each other; as such, the virtual world is presented as unwanted, not to be eliminated for the sake of the real world, but Cadigan seems to suggest that genuine connections among individuals can be formed only offline.

form, with lines of code radiating on their surfaces. The atmosphere is dark and captures the presence of an unknown threat. Memory is digital and the mind can transport information, as in the case of *Johnny Mnemonic*. Current reality can also incorporate elements literally stolen from the past, as *Mozart in Mirrorshades* by Bruce Sterling and Lewis Shiner portrays.

The new medium of storytelling found at the intersection of multiple technologies combining text, sound, and visual imagery is perhaps best suited for the user to "jack-in" an artificial reality. Video games such as *>observer_* reproduce the cyberpunk experience in immersive narratives featuring almost all the tropes and concerns of the literature. In this video game, the player explores a dystopian world, set in 2084, as Daniel Lazarski, a detective, is investigating crimes in the memories and minds of the victims and criminals. Always in the rain and always in requirement of some drug 'fix', Lazarski enters the minds of the dead to solve murder mysteries. Therefore, the user is transported into various versions of reality in a realm that presents objects in their informational properties without their being dismissed as unreal. In this sense, when technology is pervasive to the point in which it is used to access other people's memories and fused to the body to offer incredible enhancements, reality, paradoxically, is disconnected from its physical essence and exploration is pursued in the realm of the stories memories tell. Of course, these memories are remarkably violent and traumatic, while psychological confusion is often experienced. Indeed, this experience points yet again to Berkeley's principle of perception as a philosophical and contemporary inevitability. The informational realm, however, is not devoid of threat either. Similar to Meyrink's Golem, a dangerous creature lurks in the background. The player, once within this informational realm, is required to outsmart and outpace the creature without weapons or other objects at his disposal.

Thus, the informational realm takes the form of a puzzle. Its solution brings the player to solve the murder and to observe what it means to be human.[51]

As the cybernetic imagination finds new ways to portray replacement, or, in the case of cyberpunk and hacking, to find an alternative way forward by acknowledging the effects of digital technology, the socio-economic and political consequences of digital technology researched in other disciplines appears to co-opt the expectations of replacement found in the literary imagination. In particular, with the rise of cyberspace, apocalyptic visions of global destruction and existential threats to the integrity of the nation state are being formulated in various fields of political science. Here, the predominantly quantitative analysis of cyber threats seen through their technical potential of causing technological malfunctions are explained with allegories or the imagery of cyberpunk fiction. The result is a fact-based imagination of a cyber apocalypse.

2.7. Cyber Apocalypse

In his youth, Tim Berners-Lee was reading a story by Arthur C. Clarke, *Dial F for Frankenstein.* The story is about a global computer that becomes conscious. It starts with a

[51] The narrative was written by Andrzej Mądrzak, and the game developed by the Polish studio Bloober Team. It was released on August 15, 2017 by Aspyr. The new storytelling medium of video games enabled new writing specializations such as "narrative designer." The narrative designer is responsible for writing all aspects of the story of the game, similar to a movie script writer. However, there are specific challenges and medium-bound rules that writers must follow in order for the story to read well (Smart, 2000). Also, for a more updated look on recent video game narrative design in line with industry developments see Skolnick, 2014. >*Observer*_ is not necessarily unique in the genre. It is a "psychological horror" game designed within cyberpunk aesthetics and philosophy. A similar game exploring the question of what is human in *Soma*, a survival horror game released by Frictional Games in 2015.

simple reminder that the world's communication systems are connected: "At 0150 GMT, on December 1, 1975, every telephone in the world started to ring."[52] Puzzled as to what caused the event, scientists at the Post Office Research Station began to inquire into various possibilities. One of them, Dr. John Williams, head of the Mathematics Division and a fiction writer, tells them that the number of neurons in the human brain was exceeded by the number of autoexchange switches in the world's communication system linking every phone to every computer and to every satellite. Therefore, the possibility is that the global machine has become aware of itself, "For want of a better word – consciousness" (Clarke, 1961, p. 67). This newborn global consciousness, according to the scientists, would want to start to learn to live. At first, it will look for food, electricity, and then play with the toys laying around it. Thus, the scientists experience a power surge and then they realize that "babies *break* things", as the BBC reports on an "unprecedented number of industrial accidents" and the "unexplained launching of three salvos of guided missiles from military installations in the United States," as well as other technologically focused catastrophes. The scientists, therefore, decide to do something like "prefrontal lobotomy." As one of the characters says: "Oh yes,- cutting out slabs of the brain. That would certainly do the trick. Expensive, of course, and we'd have to go back to sending telegrams to each other. But civilization would survive" (*Ibid.*, p. 72-73). However, in the end, Williams concludes that it is too late, because "for *Homo species*, the telephone bell had tolled" (*Ibid.*, p. 73). The story for Tim Berners-Lee, according to a *Time* piece, was "crossing the critical threshold of number of neurons," about "the point where enough computers get

[52] Clarke, Arthur, C., Dial F for Frankenstein, in *The Wind from the Sun. Stories of Space of the Space Age*, Harcourt Brace Javanovich Inc., 1961, p. 67

connected together" that the whole system "started to breathe, think, react autonomously" (Wright, 2001).

Shortly before the fall of the Iron Curtain, in March 1989, Tim Berners Lee was working at the Conseil Européen pour la Recherche Nucléaire (CERN). There, he submitted a proposal on information management to upper management. Mike Sendall's review of the proposal was summarized as "vague, but exciting."[53] A year later, Berners-Lee produced the technology that remains the foundation of the World Wide Web: (1) HTML, the formatting language for the web, allowing for a visual design of web-pages and connection of information through links conceived as (2) URI, Uniform Resource Identifier, to determine the identity of a web page or information in the linking process, and (3) HTTP, Hypertext Transfer Protocol, which allowed the access to the resource identified. Hardly any website nowadays does not conform to these technologies even if web development has evolved by incorporating additional coding languages with the classic HTML (which has developed as well). Thus, with these technologies, the World Wide Web was invented. Up until this point, users did not have a dynamic visual interface of the Internet, but with the tools provided by Tim Berners-Lee, individuals could design their virtual homepage or set up information encyclopedias and so on. In this sense, WWW is the visual transformation of cybernetics and the principle of the "memex".

Furthermore, what was once reserved to the cybernetic imagination of the author imagining people peering into screens displaying code or images of a world within the computer, that world was put in front of the users in real life. In addition, this world is not static, but changeable and perpetually developing. As we can assume, the impact of the widespread use of

[53] Tim Berners-Lee's Proposal. Available here: http://info.cern.ch/Proposal.html

the WWW made the cybernetic imagination considering similar technologies less interesting because of its lack of novelty. There was nothing spectacular anymore about navigating informational landscapes, connecting from one place to the other as in the case of *Neuromancer*, because that was simply possible. With the inventions of Tim Berners-Lee, the cybernetic imagination became observable directly with the senses. But as the technology developed and was put in use throughout the world, with business and governments increasingly relying on the technology, replacement found a new form of play appealing to many of the anxieties explored by early cybernetic imagination: cyber warfare and autonomous killing machines.

Autonomous killing machines, or bluntly put, killing robots, is an old trope in the cybernetic imagination actively portraying replacement. As we have seen, since Karel Capek, robots were expected to rebel against their masters through their higher work potential, physical abilities, and superior intellect. While the machines in *Matrix* would replace reality with an artificial simulation, the physical replacement of humanity found, once again, a champion in the cybernetic imagination: *The Terminator*. The story created under the supervision of James Cameron, tells of an advanced human-resembling machine sent back in time to kill the leader of human resistance fighting a war against the robots of the future. The opening scene is one of the most memorable in the 20th century cinema and reflects, at the same time, the play of replacement in all its practical implementation. In a hail of laser shots being fired left and right, a dreadful looking robot steps on a human skull, presumably that of a fallen human soldier in the war. The sound of the cracking skull as the weight of the robot pushes easily on it reminds the audience of the impending replacement of humanity, and the heroic sacrifices needed in the

present to avoid such a tragedy.[54] This imagery is sometimes associated with the real-life exploration of autonomous weapons and influences, seeming to suggest the debate on the morality of autonomous robots selecting human targets independently.[55]

Cyberspace has been transformed from an imaginary artificial reality where informational objects interact with the mental representation of an individual to a real artificial reality where informational objects are under constant threat that will probably destroy humanity. The possibility of a 'cyber apocalypse' was popularized in the late 1990s with the famous 'Y2k virus' also known as the Millennium Bug, and particularly following 9/11 with a hyperbolized discourse on the consequences of so-called cyber weapons. *Cyber War: The Next Threat to National Security and What to Do About It* written by former government official Richard Clarke, wraps up the potential of cyber threat in a language bordering science fiction by deploying numerous examples of possible apocalyptic outcomes of hacking the U.S. national power grid, and so on. *Wired* Magazine commented on the book suggesting that readers "who want to jump to the steamy parts should start at page 64 in the chapter 'Cyber Warriors.' It's

[54] See the essays in Brown & Decker 2009. The problem of extinction is explored by Greg Littmann to conclude similar in the lines of Capek that the replacement of humanity by thinking and feeling machines (as they appear) is not necessarily bad news. George A. Dunn follows with an essay deploying the Cartesian dualism for identifying the difference between an advanced intelligent machine, T-101 or the updated model T4, and humans. Wayne Yuen in a paper titled "What's so terrible about Judgement Day?" notes that "So just looking at the potential consequences from the utilitarian viewpoint, it may be true that Judgement Day is preferable to stopping Dyson, since it actually maximizes interest satisfaction in the long term" (p. 169). Similarly, Jason T. Eberl in "What's so bad about being terminated" writes that "What is it about death that's so *bad* which in turn makes indiscriminate killing *wrong?* Given death's inevitability, particularly with Judgement Day ever on the horizon, wouldn't a stoic acceptance of eternal dreamless sleep be the most appropriate emotional response?" (p. 215). The answer is certainly not. Eberl confuses Stoic philosophy in the sense that it attributes it the popular understanding of passivity in the face of events outside of control. Battle here would be the virtuous thing to do even if the outcome is not the one desired.

[55] Paul Scharre presents the history, development, and ethical issues coming with military autonomous weapons in *Army of None: Autonomous Weapons and the Future of War*, W. W. Norton & Company, 2019

there you'll find the Book of Revelation re-written for the internet age, with the end-times heralded by the Four Trojan Horses of the Apocalypse" (Singer, 2010).

Similarly, U.S. officials such as former Secretary of Defense Leon Panetta have coined such terms like "cyber Pearl Harbor" to refer to digital threats.[56] Such references to the potential outcome of cyber threats that appeal both to historical events with deep significance in contemporary American culture and to the technology of cyberspace relatively unknown to the general public read as a mix between traditional cyberpunk literature and new, real, digital technology, infused in the imagination of contemporary technological and scientific development. As such, the play of replacement found its way reconfigured at this very intersection between fact and fiction in such a way that one is indistinguishable from the other. However, what remains is the violent contest transported throughout the last couple of centuries in the imagination of robots, computers and its digital products.

In this chapter, we have observed that the expectation of replacement in the cybernetic imagination is confined on two major levels: (1) a concern regarding emotions and (2) the problem of defining real reality. In addition, we also acknowledge (3) the possibility of balancing the real and the virtual through an ethical framework based on personal moral conduct found in hacker culture and cyberpunk literature. Throughout the 20th century, the question "why this fear of robots?" formulated by Isaac Asimov is essentially an investigation of the self. The fear, according to Asimov, can be easily dismissed as an unreasonable anxiety of the unknown consequences of yet to be discovered technologies. But as we have seen in the novels of Philip

[56] The language and metaphors deployed to create a hyperbolic discourse on the threat coming from cyberspace have been documented by Sean Lawson and Michael K. Middleton in Cyber Pearl Harbor: Analogy, fear and framing of the cyber security threats in the United States 1991 – 2016, in *First Monday.* Available here: https://firstmonday.org/ojs/index.php/fm/article/view/9623/7736

K. Dick, Gustav Meyrink and the rise of the computational disciplines, the question is more appropriately rephrased as “what does it mean to be human?” when the potential of technology attempts to recreate the human in its artificial likeness. This question is not only more pertinent to ask, but it retrieves the problem from the cybernetic imagination and turns it into a viable research question suited for interdisciplinary work at the intersection of literature and philosophy.

However, given the contemporary call to action to respond to an emergency articulated in the lines of the robot takeover narrative peculiar to the cybernetic imagination, whether it be that we live in a computer simulation or that robots are becoming so sophisticated with the potential to resemble humans and surpass our ability to remain the dominant species of the planet, it would also be pertinent to ask whether this concern, this fear of robots, is real or imagined. As with the “cyber apocalypse” scenario, what of it is a real point of concern and what is the cybernetic imagination repeating the play of replacement? As we have seen, it is quite the difficult task because the play of replacement finds its articulation in the cybernetic imagination that has penetrated the realm of science which in turn formulated the methodologies to replace traditional humanities inquiry. In the end, perhaps, it lies within agonistic play as the essential grounding element of Western culture and, as such, the outcome is inevitable. Nevertheless, instead of letting A.I. generate poems while we experience Rome in cyberspace, identifying the precise workings of the cybernetic imagination in contemporary visions of artificial intelligence and cyberspace would enable us to either dismiss the hyperbole and proceed to a viable ethical framework of investigation, research, and digital development, while paying close attention to the fear of replacement as a guide.

The next chapters will focus on these formulations of replacement by analyzing the advancements of affective computing and its expectations, playing with reality in the cybernetic imagination and the play of hacking as an attempt to locate this issue between the realms of the real and the digital.

3. Playing with Reality

This chapter is concerned with the agonistic play of reality in Newtonian action-response paradigm where what we consider real is defined by the response of the object: an object exists if, upon action, it produces a reaction as observed through physics. By contrast, with digital reality, or artificial/virtual reality, what exists is defined through its informational properties. Note that 'exist' and 'real' are interconnected ontological notions that find definitional difficulties when evaluated according to the above distinction – for example, power is not real, as it is not tangible, does not influence the senses, but exists in its informational properties as the perception of the manifestation of action and attribution of value to the result.

Here, I am not interested in exploring the interminable philosophical discussions about reality, but to consider the possibility of understanding reality in the absence of the agonistic play of replacement. Thus, my objective is to detect whether it is possible for our understanding of reality to incorporate digital objects into it, rather than to dismiss them as "not real." Perhaps, the best result would be to consider digital objects as "illusions," not in a pejorative sense of things unworthy of attention but as elements of *in ludere*, that is, play as ontological craftwork, since the etymology of 'illusion' reminds of the ludic activity. In this sense, I shall, later in the chapter, investigate cyberspace in William Gibson's *Neuromancer*, the novel that coined the term, while referring to other literary and philosophical works as well as considering approaches in the philosophy of mind and cognitive science. *Neuromancer*, as science fiction does in general, combines the two aforementioned paradigms. The result is an ensuing tension between science

and fiction, that is, between the Newtonian and the informational paradigms. Cyberspace finds itself at this intersection. On the one side, it is supported by the physical equipment conceived in its physical mechanisms and materials, while on the other side, it exists in the informational properties of the objects that are artificially designed within it.

3.1. The Fabric of Reality

David Deutsch in *The Fabric of Reality* distinguishes between different paradigms but with similar points in relation to the case of Galileo Galilei facing the Inquisition. Galilei is known for proposing that Earth is not a static physical entity around which other planets orbit, but on the contrary, that it, together with other planets, orbits around the Sun. This has struck a sensitive nerve with the established Catholic view of reality, not because of Galilei's theory, per se, but because of the world view that derives from heliocentrism and scientific inquiry. After all, the Inquisition was not necessarily concerned with explanatory attempts for observable things, but rather with the general view of reality. As such, if "the Book of Nature is written in mathematical symbols," that was acceptable. The Inquisition, however, took issue with the possibility that the scientific method will explain God out of existence. Deutsch notes that "he [Galileo] insisted that scientific reasoning took precedence not only over intuition and common sense, but also over religious doctrine and revelation. It was specifically that idea, and not the heliocentric theory as such, that the authorities considered dangerous" (Deutsch, 1997, p. 74).

Deutsch notes that the Inquisition was less concerned about science than with the view of reality. Here we find two antagonistic positions that define our sense of reality. Galileo proposed that science was sufficient, that scientific method could explain reality with precision without the need for revelation. In effect, this statement dismissed the informational properties of Christian

symbolism and metaphysics as the core principles of Christian life, while the Inquisition upheld the latter to be of primary importance. The Inquisition would not have the world explained in pure mathematics. In this sense, Deutsch makes the mistake of assuming that the Inquisition was invested in explanatory theories of the physical reality with the same *modus operandi* as Galileo. His argument rests on the assumption that the Inquisition was seeking what is true, when in fact, it defended truth as the metaphysical whole of Christian life:

> "The Inquisition's explanation is that the planets are seen to move in complicated loops because they really are moving in complicated loops in space; *but* (and here, according to the Inquisition's theory, comes the essence of the explanation) this complicated motion is governed by a simple underlying principle: namely, that the planets move in such a way that, when viewed from the Earth, they appear just as they would if they and the Earth were in simple orbits around the Sun" (*Ibid.*, p. 79).

The Inquisition saw reality in the way it was represented through Christian symbolism common in traditional Catholicism. Symbolism tells moral stories or are representational metaphors of Truth, impalpable and unattainable. For example, the cardinal virtues of justice, wisdom, temperance, and courage laid out by Plato in Book IV of *Republic*, to which the three theological virtues were added, can be found in Christian iconography represented as the tetramorph: eagle, angel, bull, and lion, attributed to the four evangelists Matthew, Mark, Luke, and John. Thus, reality as the work of God for the Inquisition was not to be explained but to be understood, and the scientific method would remove Christian metaphysics. As we can imagine, it was a serious threat to the Church.

In a similar vein, Michel de Certeau investigated the famous events at Loudun (as did Aldous Huxley and others). Certeau observes that the event was transformed into a public spectacle, and the unfortunate priest who was accused of witchcraft and conspiring with the devil was a pawn in a power-game involving Cardinal Richelieu to ensure the authority of the Catholic

Church threatened by the Enlightenment (Certeau, 2000). Similar zero-sum games where science was to be proclaimed victorious over religion have been played before and continue to affect individuals today.

This example is to simply illustrate the distinction between the scientific and the 'cultural' approaches. As with regard to reality, Deutsch explains the solipsistic view according to which 'real' is a product of the mind alone without interference of an external agent. He refutes this claim by saying that even the individual who thinks must have some form of external influence and, therefore, Descartes' famous maxim doesn't hold. However, what the mind creates in reference to an external object, as in whether ideas or physical objects are claimed by an imagined goal, ideal, or even an individual, or in Socrates' terms "in wonder", is evidence of consciousness. Indeed, Descartes proceeded in his ontological assumptions on the premise that reason was given to humanity by God. Therefore, Descartes by default assumes the presence of an external entity influencing our mind-dependent reality.[57]

The opposite view of solipsism, as Deutsch describes it, is behaviorism. While solipsism is a mind-dependent reality, behaviorists, on the other hand, insist that reality is external to the human individual being recognized in the events and things interacting with the individual from outside. As such, reality is purely interactive. Realists, however, claim that what is real is that which 'kicks back'. If an object produces a reaction upon action, then it exists. In short, this is our contemporary Newtonian understanding of reality. As we can easily observe, A.I. and

[57] According to Peter Kuberski, the effect of the Cartesian dualism was to establish "the division between physical and spiritual forms of significance." To rectify this situation, Kuberski provides the concept of "chaosmos" by which he means "a unitary and yet untotalized, a chiasmic concept of the world as a field of mutual and simultaneous interference and convergence, an interanimation of the subjective and objective, an endless realm of chance which nevertheless displays a persistent tendency toward pattern and order," (Kuberski, 1994)

cyberspace are solipsistic. For a machine to acknowledge the presence of the programmer and understand that its processes are directly influenced by an external entity is a sign of consciousness. So too, cyberspace by definition is a realm of code influencing the mental phenomena of the user. Nothing there is real in Newtonian terms, even if cyberspace abides by the laws of physics and users-- hackers too--have to interact with it in terms of cause and effect.

Artificial intelligence must therefore persuade the external observer that it is intelligence, thus assuming a behavioristic approach to reality. The observer, on the other hand, is incapable of truly identifying intelligence innate in the construct but can evaluate its actions and reactions together with the biased affective responses following emotional stimuli simulated by the machine. In other words, A.I. is solipsistic while the human observer is behaviorist. As such, we find ourselves in the impossible situation of having to identify a solipsistic artificial mind through its behaviorist actions, to evaluate it as a form of sentient entity. Perhaps, the alternative is to consider the other stimuli (as we will see in the chapter *Playing with Affection*) and proceed according to the dictum "we feel the other as human," even if it is a machine. In the situation in which A.I. and cyberspace become less distinct than our physical reality, there may be no other way than to acknowledge them as elements of the play of reality as they are, rather than to continue to distinguish them in terms of real/not-real.

Descartes, in addition, was a skeptic inasmuch as he occasionally did not distinguish between reality and the simulation of reality. Cyberspace, after all, was defined by William Gibson as an agreeable dream, a "consensual hallucination" accessed directly through technology. The theme of dream-like reality is common in science fiction and particularly important in cyberpunk where characters jack in and out of the simulated reality. However, it is fair to distinguish between things that are and things that represent, in other words we can

differentiate between things that ‘kick back’ and those that don’t but still hold important informational properties without dismissing the latter altogether, as in the case of Galileo and the Inquisition. Cyberspace is the realm of things as representations, and its history goes back to the idea of dream-world inquired into by the pre-Socratic thinkers and finds relevant similarities with world religions. For our purpose of observing the agonistic play of replacement in the cybernetic imagination, coding cyberspace is the immersive simulation of ancient Greek philosophy, starting with Parmenides.

3.2. Plato’s Cave

Book VII of *Republic* presents probably the most famous passage in Plato’s magnum opus, the allegory of the Cave. The exit from the Cave leading to the experience of the light outside was interpreted in numerous ways as a metaphor serving other purposes than the original intent. The allegory is primarily an epistemological process, and Socrates instructs Glaucon on the effect of the discovery of virtue by the prisoner. Since the prisoner has discovered philosophy and is now able to examine life, he is unable to cause harm through ignorance (of the virtues). As such, philosophical education is the main concern of the allegory. Socrates starts by suggesting to “compare the effect of education and the lack of it on our nature” (514a) and proceeds to set up the allegory as follows:

> “Socrates: Imagine human beings living in an underground, cavelike dwelling, with an entrance a long way up, which is both open to the light and as wide as the cave itself. They’ve been there since childhood, fixed in the same place, with their necks and legs fettered, able to see only in front of them, because their bonds prevent them from turning their heads around. Light is provided by a fire burning far above behind them. Also behind them, but on higher ground, there is a path stretching between them and the fire. Imagine that along this path a wall has been built, like the screen in front of puppeteers

> above which they show their puppets.
> Glaucon: I'm imagining it.
> Then also imagine that there are people along the wall, carrying all kinds of artifacts that project above it - statues of people and other animals, made out of stone, wood, and every material. And, as you'd expect some of the carriers are talking, and some are silent" (514a-b) (Plato, 1992).

In this setting, the prisoners can only know whatever they are being presented with. Their sight will hold the shadows projected on the walls and their hearing will understand the random noises of wind, whispers, words, as part of one object. The object presented, however, is but an incomplete representation of the whole, a reminder that physical reality shows distorted versions of the forms. In addition, of course, this is meant to suggest that whoever holds the flame and presumably controls the guards is able to instill the distorted forms as truth in the unknowing prisoners. Reality for the prisoners is restricted to that which they are given to see and hear. Outside the cave lies a different reality where the forms of the shapes projected by the light on the wall are perfectly contoured and unchanging. Thus, exiting the cave presumes access to the true forms. However, at this point, the prisoners are prevented from even conceiving such a possibility because of their physical bonds and, as such, they have to rely on their senses to construct reality.

It is unclear whether the prisoners have any access to reason, to the world within the soul and ability to remind themselves of the knowledge their souls have accumulated while free from their human bodies, according to the theory of recollection in *Meno* and *Phaedo*. Socrates apparently presumes not, because we are not told that the prisoners communicate with each other and, therefore, dialectic is impossible. In fact, as he says, "[a]nd if they could talk to one another, don't you think they'd suppose that the names they used applied to the things they see passing before them?" (514b). In other words, they would give the name "bird" to the shadow of a statue of a bird, thereby confusing the shape of the object with the object itself.

But, Socrates proceeds to say that a prisoner could be released from his bonds and taken outside the cave. After his eyes would adjust to the light, the epistemological process begins. He would be able to understand the world step by step, beginning with the Sun causing the change of seasons. "What about" asks Socrates "when he reminds himself of his first dwelling place, his fellow prisoners and what passed for wisdom there?" (516c) Knowing now the true forms of the beings displayed on the screen will undoubtedly cause him to return to ignorance. And so, Socrates concludes:

> "The visible realm should be linked to the prison dwelling, and the light of the fire inside it to the power of the sun. And if you interpret the upward journey and the study of things above as the upward journey of the soul to the intelligible realm, you'll grasp what I hope to convey, since that is what you wanted to hear about. Whether it's true or not, only the god knows. But this is how I see it: In the knowable realm, the form of the good is the last thing to be seen, and it is reached only with difficulty. Once one has seen it, however, one must conclude that it is the cause of all that is correct and beautiful in anything, that it produces both light and its source in the visible realm, and that in the intelligible realm it controls and provides truth and understanding, so that anyone who is to act sensibly in private or in public must see it" (517b).

The correlation of the allegory of the Cave with cyberspace can be made through a partial reading of this passage. There are some elements that can lead to the "consensual hallucination," particularly impressions of the prisoners of the objects projected in front of them, but they are not present outside the cave. Therefore, since cyberspace is ubiquitous in contemporary life, the allegory of the Cave provides a partial explanation into its nature. In fact, coding objects in cyberspace, thus creating the shapes of physical things within the simulated reality, could also mean a willing return to the bonds of imprisonment in sensory reality and a distancing from the true forms. Nevertheless, in Book VII of the *Republic* we have a separation of actual and non-actual reality, the former being outside of purely sensorial interpretation of the real, while the latter is located outside of the cave and defined as actual through the process of education, of

learning. Before Plato, however, this separation was eloquently presented by Parmenides in *On Nature*.

3.3. The Father of Metaphysics

To consider Parmenides' poem in its entirety here would exceed the purpose of this study. However, the 'discoverer' of metaphysics and ontology is hardly to be omitted in the conceptual formulation of a technological environment which challenges reality by alluding to the elements of inquiry that consider reality as what *is*. The source of reality in cyberspace considers concerns that are similar to those expressed by the 'goddess' to the youth in the journey of knowing. Reality, as such, appears in Parmenides as epistemic. In our contemporary world in which the difference between digital, physical, and non-physical as distinct from digital (or yet undigitized) is diminishing rapidly, the ancient philosopher can be a guide to formulating a sense of reality that involves digital objects rather than dismissing them based on their devoid physical nature.

In addition, Parmenides and the Pythagorean tradition have influenced Platonic metaphysics. Similar to the allegory of the Cave and the metaphors of light and darkness that Socrates mentions, Parmenides' youth arrives at "the gates between the journeys of night and day" (1.11-19).[58] Upon arrival, the goddess declares that "You must be informed of everything, both of the unmoved heart of persuasive reality and of the beliefs of mortals, which comprise no genuine conviction; nevertheless you shall learn these too, how it was necessary and the things that are believed to be should have their being in general acceptance, ranging through all things from end to end" (1.29-32).

[58] Translations and quotes are from Coxon, A. H., *The Fragments of Parmenides*, Van Gorcum, 1986.

In this section, the unnamed goddess conveys that there are two paths of knowing. One is an inquiry into reality (*Aletheia*) and the others into "the beliefs of mortals" or perhaps opinions of mortals (*Doxa*). The goddess explains shortly that "for the same thing is for conceiving as is for being" (4.1). Reality is that of "what is". Fragment 4 appears to be a bit confusing. It may sound as if "conceiving" means that the continuous changing epistemic process of reality as knowing is defined as being. However, this isn't the case. "What is" appears in fragment 8.1-4 as "that Being [which is] is ungenerated and imperishable, entire, unique, unmoved and perfect". Therefore, reality is presented as a changeless "is" while *Doxa* is not.

Later in the poem, we discover that "Being is in a state of perfection from every viewpoint, like the volume of a spherical ball, and equally poised in every direction from its center" (8.39-49). In this sense, the goddess presents the youth with two possible ways of inquiry about reality: to inquire into "what-is" or into "what-is-not". The goddess also advises not to pursue the latter, as there is no way to understand that which is not, since it is absent of informational properties. In this sense, "conceiving" as an epistemic process is *Aletheia* as "the same thing is for conceiving of the thought conceived; for not without Being, when one thing has been said of another, will you find conceiving" (8.31-36).

The beliefs of mortals lie in error because of "naming". Thus, mortals have taken from reality and given it various interpretations. As Alexander P. D. Mourelatos notes in *The Route of Parmenides:* "what this pioneer [Parmenides] understood is precisely that the quest for phusis 'essence' does not have to be, indeed, could not be, a quest for phusis 'becoming'. So the concept of phusis and the related concept of speculative predication are of prime significance for Parmenides, much as they were in the theories of the 'physicists'; but they are being purged of the implications of dynamism, and of materialism as well. In other words, the concept of *phusis*,

although not featured explicitly in ‘Truth’ is implicit in the presentation of ‘routes’. In ‘Doxa’, on the other hand, it appears explicitly, and there it is misapplied – by the undiscerning mortals” (Mourelatos, 1970, p. 62). This part, forming *Doxa*, is set in contrast with *Aletheia. Doxa* is change, in motion, perishing, and everything the former is not.

It is unclear, though, which of the two takes priority in understanding or whether the youth should ground his inquiry solely on *Aletheia*. After all, the goddess suggests to take both paths into account and refrain from the one that takes both what-is-what-is-not simultaneously. Multiple interpretations of the poem have been produced by classicists and classic philosophers before them. L. Taran comments that “There is nothing in the text to substantiate the claim of those who assert that Parmenides maintains that past and future cannot be predicated of Being to which only the present ‘is’ truly belongs. Parmenides is only denying that Being ever perished or ever will come to be” (Taran, 1965, p. 177). In other words, Being exists (*esti*) (Ketchum, 1990; Cordero, 2004). But what is essential about Parmenides is that he distinguishes between two worlds: one of reality that presumably is a single changeless entity--a monistic approach, and one of opinion that considers reality. As multiple arguments have been given, conferring hierarchy and primary importance to *Aletheia* over *Doxa* which is flawed and contradicts the former, I tend to believe that one is not meant to exclude the other and for a true understanding of *Aletheia*, the latter, even if susceptible to error, is required. Otherwise, the goddess could have stopped just there and simply proclaimed *Doxa* to be entirely false, which she doesn’t. But there is something more to it.

Reading the poem, we find multiple contradictions and counsels that are encouraged in one instance then discouraged in the next. Furthermore, why would the goddess even mention the third way? Much of it doesn’t make sense. It may be the case that the interpretation of the

poem is done through our contemporary methods of inquiry stipulating the necessity of finding some form of logic within it and expecting to reveal something about "what-is" according to our own evaluation of Being. Recent studies of the poem confirm this tendency. Scholars are keen to identify something specific about Parmenides' verse in itself. However, according to Peter Kingsley, the poem is much more than that. In fact, as Kingsley notes, "Parmenides' poem is not for academics. There is nothing scholarly here at all" (Kingsley, 2003, p. 28). It is, instead, a riddle. A riddle about reality and our way through it, a play on ontological certainties challenging humanity's ways of thought proposing, in all contradictions, a generative reality and the superfluousness of reality-making thinking. Kingsley's interpretation stands out in this sense because it doesn't simply look at what the poem is in its linguistic properties but attempts to understand the riddle and extract the advice of Parmenides of Elea as a practical philosophy for life.

On a different occasion, Kingsley notes that "instead of the love of wisdom, philosophy turned into the love of talking and arguing about the love of wisdom" (Kingsley. 1999, p. 32). Its practical application, as a way of life, is almost absent from our contemporary academic concerns. As such, when interpreting the poem, we inevitably fall on our contemporary methods of inquiry to investigate a text written, perhaps, not to create a field of scholarly philosophical investigation, but to reveal something about living life. According to Kingsley, the text is to be taken as a riddle for a way of life and not as a simple text for scholarly inquiry. It is perhaps in this interpretation which reconsiders the poem setting the foundation of metaphysics that we can find a way of reconciling digital objects with offline reality.

The poem is indeed about reality and the nature of thought. But not in a theoretical sense, for "what he taught has nothing to do with theory. It's a simple matter of experience: the

experience of reality" (Kingsley, 2003, p. 48). This experience takes the form of a rite of initiation. Parmenides descends into an underworld, a *katabasis* (literally under(kata)ground(basis)), and so separates himself from our physical reality. However, he does not seem to be a newcomer, but an initiate. And, as such, how can this experience be possible? Kingsley notes that "through the practice of stillness we come to experience a reality that exists beyond this world of the senses" (*Ibid.*, p. 46). Thus, the descent into the underworld is not a journey of distance, to arrive in Hades like Odysseus, but a mental injection into that realm. This is also the first similarity with journeying into cyberspace. As mentioned before, the process of entering cyberspace is done through "jacking-in," excellently portrayed in *Matrix* where the characters sit on a chair while the technology connects their brain to the simulated reality thus facilitating entry. In this process, their bodies remain in the physical world while their mind processes the perceptions of digital objects.

Of course, it would seem that jacking-in implies a distinction between one place and the other. Parmenides' poem holds some additional relevant advice: space is illusory. The journey into cyberspace is static. Indeed, the concept itself of cyberspace is problematic. The word contains "cyber" a word of Greek origin meaning steersman and "space". The latter is misdirecting. There is no space in artificial reality. Going from "place" A to B is a matter of perception, a mental projection of the idea of space drawn from physical reality.

Taken as a riddle, the question of reality becomes even more complicated. What we are shown, at least according to our contemporary interpretational tools, is that the way the goddess asks us to follow is that of *Aletheia* while the opinions of mortals are unreliable. However, as Kingsley suggests, that may not be the case, as "Parmenides is suggesting nothing is what it seems. He is hinting at a reality very different from the one we take for granted." (*Ibid.*, p. 66). It

is a reality of play between what-is and what-is-not, located at its intersection. In essence, "what exists for thinking is whatever you are able to think about. So, in other words, Parmenides is saying that anything you can think about has to exist for you to think about" (*Ibid.*, p. 71). But this is not to say that dragons are real, yet for him "every thought is its own validation. It needs no confirmation outside it. Whatever we are able to think is true" (*Ibid.*, p. 73). Furthermore, it is not about the way of thinking or any particular kind of thinking. Parmenides, according to Kingsley, refers to every thought we happen to have. And this contradicts almost every metaphysical system the Western world has established in the past while grounding them in one way or another in the poem of the father of metaphysics.

It would seem that Parmenides communicates the message of the goddess in the form of a riddle as an act of initiation into the mystery of reality. As a herald, he passes on the knowledge of reality as a sum total of every thought we have, and since we are social beings, it wouldn't be a mistake to extrapolate the idea that reality is constructed by the interconnection of the thoughts of individuals. Here is how Kingsley concludes:

> "We are not some tiny, insignificant specks lost somewhere at the edge of a vast, impersonal universe – unless we want to believe we are. Ultimately the world is just where we happen to be. Everything else is our imagination. Wherever it seems that you are, in some bleak corridor or watching trees outside the window, this is the center of reality and your thoughts are its edges. […] The moment you understand that the only criterion of reality is thinking and perception, that there is nothing that is not, then everything is suddenly together – linked to itself in a perfect continuum" (*Ibid.*, p. 78)[59]

[59] Kingsley's discussion of the poem does not end with this conclusion, but for our purposes it would suffice. The book Reality continues examining another pre-Socratic thinker who has tremendously influenced western civilization, Empedocles. Because, as Kingsley puts it "Parmenides warms, Empedocles burns" (339). By identifying the key concept with which Parmenides operates to demonstrate human fragility and lostness in reality is metis, translated usually as skillfulness, cunning, trickery. The advice following this is that Parmenides endorses the idea that "the more you let yourself become a part of it [metis] the more you begin to discover that absolutely everything, including the fabric of reality itself, is trickery and illusion" (91). The morality underlying metis appears to be trivial, remarkably problematic for our contemporary mind. However, let us remember that in cyberspace, a free environment where basically anything goes, metis is the definitory value of the hacker culture. Thus, Parmenides'

In effect, Parmenides sets the stage for the discussion of reality as actual and / or non-actual, but also continues in the monistic tradition set by his predecessors. Thus, what he offers is a play of thinking and perception outside the agonism of replacement of the real by the artificial or vice-versa.

Likewise, the Quantum Relations Principle proceeds from the idea that reality is changing and does not dismiss the "unreal" or the artificial from the ontological process. Instead, physical and mental phenomena are considered simultaneously through DFOs, and, as such, QRP is optimal for insertion of digital objects into the reconfiguration of reality. Parmenides' arguments are, from a scholarly perspective, unconvincing and weak. But the proposition is that reality is change, and since it is thought and perception then it becomes hackable (*metis*). Thus, the Simulation Argument, according to which we live in a computer simulation, is confirmed, infirmed, and driven out by thought itself. It is confirmed by the mind-dependent reality argument. Since we can, at least theoretically, code mental processes and have the required advanced hardware to process them, then a simulated reality is possible. It is infirmed by the idea that reality is becoming, randomly generated concomitantly by individuals in particular moments. It is, then, similar to Gibson's definition of cyberspace: "a consensual hallucination". In addition, it cannot be "one" reality, but a combination of interconnected realities. In effect, reality is a choice, and quantum mechanics approves (Vedovato et al, 2017; Cho, 2017).

---

morality appears to be a free play of trickery in a reality defined by thought and perception. But this reality appears to be unchanging, so there is a further confusion added to the interpretation of Kingsley. If reality is thought and perception, wouldn't that require it to be dynamic too? The argument appears to be in favor of a changing movable reality, and it is the individual that doesn't change, or rather the pursuit of wisdom of the individual. That journey starts from the individual and returns to him, a path without a distinct destination and since Parmenides does not provide any definition of wisdom it must mean that it leads nowhere else but back to the soul seeking it.

But the proposition of reality as thought and perception has been reformulated in different parameters by Spinoza and Berkeley, while Leibniz's *Monadology* also has had a role in the play of the actual and the non-actual, found in Eugen Fink's metaphysical claim that play is the symbol of the world. Spinoza proceeding from "what God is" concludes that we know God through his attributes: "His attributes which are known to us consist of two only, namely *Thought* and *Extension;* for here we speak only of attributes which might be called *proper attributes* of God, though which we come to know Him as He is Himself, and note merely as He acts towards things outside Himself" (Spinoza, 1994, p. 74). But these attributes have been conferred to God "so that all that he has of thought are only modes of the attribute thought which we have attributed to God. And, again, all that he has of form, motion, and other things, are, likewise, modes of the other attribute which is attributed by us to God" (*Ibid.*, p. 74). In conclusion, Spinoza notes: "that belongs the nature of a thing, without which the thing can neither be, nor be understood; not merely so, however, but in such wise that the judgment must be convertible, that is, that the predicate can neither be, nor be understood without the thing" (*Ibid.*, p. 75). In other words, reality is what we understand about it through thought and perception. In a similar vein, Bishop Berkley adds that "it is evident to any one who takes a survey of the objects of human knowledge, that they are either ideas actually imprinted on the senses, or else such as are perceived by attending to the passions and operations of the mind, or lastly ideas formed by help of memory and imagination, either compounding, dividing, or barely representing those originally perceived in the aforesaid ways" (Berkeley, 1994, p. 98). Thus, it is inevitable that we assume the presuppositions and conclusions of the view that reality is mind-dependent. As such, digital objects are perceived and thought of no less differently than other, physical things. That their essence is code and not matter is of negligible relevance and,

therefore, the play of reality can incorporate such objects without insisting on the ontological antagonism of real versus not real.

3.4. The Play of the World

Eugen Fink's actual and non-actual character of play is rather an ontological issue concerning the relation of the human being and the world as a "special manifestation of the relation of innerwordly beings to the all-encompassing world-totality" (Fink, 2016, p. 80). This "innerwordly" type of play consists of the generative feature of play as affecting, or even changing, the "world-totality." In this sense, play, as we can easily observe, is located at the ontological intersection of an innerwordly reality and the totality of facts surrounding people. Similar to Huizinga, Fink stresses the freedom of play as independent acts and insists on the irrational aspect of play as "creative incursions into a context of events otherwise determined according to rules" (*Ibid.*, p. 80). Rules are, of course, those of our Newtonian reality based entirely on action-response causality that propose that an object exists only if it is able to produce a sensible reaction to an action. Fink proceeds from the question of compatibility between natural causality and freedom and ends up in the almost impossible philosophical dilemma of "free will." This is a problem we intend to avoid by saying that if reality is thought and perception, the choice is always there to construct, reconstruct, modify, and adjust reality according to the moral predispositions of valid thoughts and unnecessary perceptions, thus framing the answer in Stoic philosophy.

However, should a new theory of reality discuss the problem of free will? Not directly, but as an implicit factor of reconsidered ontological certainties. For example, if we assume that the Simulation Theory is correct, then the problem of "free will" is hardly necessary to be

considered. If the world is a computer simulation, then the age-old problem of *moirai* rests in predetermination according to the possibilities of the world already coded in its entirety.

The other solution proposed by Fink to observe the compatibility of innerwordly and world-totality is to devise it in spheres. One sphere is "the self-enclosed natural realm to what happens causally," therefore, a sphere dedicated to our current understanding of reality, and another one "to what happens freely; and carves the human being himself into a sensuous being belonging to nature and into a 'moral being' belonging to the intelligible realm of the spirit" (*Ibid.*, p. 80-81). As such, we inevitably return to the Cartesian dualism or back to Parmenides. As Fink notes with reference to technological progress, it "streams into nature, already presupposes nature as a setting, as material for its work; this technical production cannot at all be thought of as housed in a sphere separated off from nature" (*Ibid.*, p. 81). Thus, nature itself is the playground providing the physical materials to support the technological development. Reminding ourselves, in this instance, of Huizinga's requirement of space as a characteristic of play, we once again note that "jacking-in" to cyberspace proceeds from the actual world.

But what about things themselves? Now we approach the question set out to be answered for this chapter, namely, whether digital objects can be objects of the world. And "what kind of remarkable character is this: the 'actuality' of a thing? Is it anything *in the thing* at all? We can see things, feel them – but can we also see and touch the actuality of these visible, tangible things?" (*Ibid.*, p. 83). As I note in the chapter *Playing with Affection*, we rest our ontological conviction of artificial intelligence as human-like in our perception of feelings towards them. We feel them into existence. Only if feeling the machine as human can we say that the machine is human. If feeling as such is contested by sensory perception, then its ontological validity as humans is impossible. Our perception, in this case, of the human-machine is defined by affective

connection. Thus, as Fink notes, "everything that is actual is at the very least also 'possible' but does not all need to be 'necessary.' Thus conceived, 'actuality' is taken to be an ontological modality of existing things, is assessed as the being actual (as the *existentia)* of beings" (*Ibid.*, p. 83).

With the supposition that things do not have to be necessary, we are, in fact, freed from the reaction imperative, that is, objects in order to be do not have to confine to the necessity of producing sensible evaluation of their reactions. Furthermore, if "actual is at the very least also 'possible'" as Fink notes, then the reverse may have to be true as well: what is "possible" is at the very least also actual. In other words, the reverse validation of actual-possible ratifies Parmenides' thesis. In these parameters, the distinction between something actual or non-actual is problematic. For, as Fink asks as well, what does "something" mean? Particularly when we consider the actuality of cyberspace. Fink provides the following response:

> "Because we can "intend" and "imagine" beings and because it is sometimes difficult to separate what is merely in our soul and what is both intended by us and also actually existing, we can therefore ask whether "something" is actual or non-actual. What is imagined is nugatory; it cannot be recognized in its claim be to actual. But what is merely imagined and therefore nugatory is nevertheless not simply nothing. It exists as imagined, as phantasm, as a representational content" (*Ibid.*, p. 84).

Thus, our first conclusion is that we cannot dismiss the non-actual as not being. Therefore, digital objects are precisely that "something" questioned above. They are something imagined and actualized through coding nevertheless thought of as non-actual but paradoxically actual as well.[60] This non-actual of play or the non-actuality of cyberspace stems from the

[60] If digital objects are considered non-actual and, as such, merely illusory, then this dissertation, since it was written exclusively on PCs and laptops generating through code letters, is nothing. In this sense, any other text produced in this way must be considered nothing as well. Furthermore, if this premise is correct, then we also have to consider the digitization of texts as progress *ut nihil*. I refrain from analogy to art since a work of art is actualized imagination in reality as palpable product consisting of physical

perception of play as a non-serious activity, leading Bruce Sterling to proclaim the non-actual of digital objects (see *Playing with Computers*). As Fink notes, “no one would dispute that play activities are actual, but, as activities of non-seriousness, they bring with them a sense of “non-actuality.” (*Ibid.*, p. 87). Play is typically considered as non-actual through not being serious. As such, if play is non-serious then play does not have actual, real consequences. As if fun is not real. Furthermore, referring to Aristotle’s concept of *energeia*, Fink argues that play is also a mode of self-actualization, thus highlighting the creativity of the ludic.

Play, however, brings tremendous opportunity for moral development. If we observe cyberspace as playing in the non-actual, then we can remain also loyal to its non-actualness and therefore “choose” whatever it is that we want to become, as we can “slop into any role of existence” (*Ibid.*, p. 90). Because ”play is the imitation in the space of imaginary” it also ”brings forth entirely entirely new motifs, gets new possibilities to flare up, possibilities with which we are not acquainted in the space in which we otherwise carry our lives” (*Ibid.*, p. 90).

In this sense, our conclusion must be that cyberspace is not entirely non-actual, even if it is difficult to conceive of it as such. Seen as a space for play rather than as a space of geography, it brings forth the potential to evaluate or reconfigure our current moral convictions and, due to the freedom of play, imagine and test out new alternatives. But what about the nature of digital objects themselves? While we have made remarks on their nature throughout this chapter, it is again that we turn to the Quantum Relations Principle for further clarification.

3.5. Digital Objects and Irreality

substances (water, canvas, paint, etc, or nature), but digital objects are formed without traditional physical substances and remain within the digital environment but nevertheless return to the actual world in perception and thought.

Throughout the history of philosophy, Yuk Hui notes in *On the Existence of Digital Objects*: “a technical object was nothing more than a tree in the garden or an apple on the table” (Hui, 2016, p. 54). Thus, an object was sometimes viewed in its teleological sense or investigated for its substance, such as Leibniz’s *Monadology*. The Quantum Relations Principle, on the other hand, bypasses these ontological obstacles and views things not as they are, thus not according to their particular nature, but according to their informational properties, including perception and mental phenomena, in their interactions with other objects, physical, mental or digital.

Hui investigates the existence of digital objects proceeding from traditional investigations on technical objects. While it is certain that the unique problem of digital objects requires study of the usual approaches in the philosophy of technology, the problem rests in seeing them as technical products and, as such, proceeding in their investigations in the same manner as it would be to consider devices located in nature, produced from nature, and so on. Digital objects, on the other hand, are not found in nature, in the substances of nature. An ontological investigation, therefore, cannot proceed from substance as there is none to be found, unless the discussion is redirected to the materials required to construct computers, which, as it is, derails from the objective itself and confers full attention to objects independent of digital objects. Digital objects are not similar to devices of matter but appear only in their informational properties concluded as particular objects through their perception associated with physical objects. Therefore, Gilbert Simondon makes some sense when he writes that ”the contemporary world resides in this misunderstanding of machine, which is not an alienation caused by machines, but the lack of understanding of its nature and its essence, because of its absence from the world of significations and its omission in the table of values and concepts belonging to culture.”[31]

The misunderstanding is ontological. Hui proceeds in the investigation of digital objects through the dichotomy of individuation and individualization set forth by Simondon and other similar methodological devices such as the distinction between technical tendency and technical fact by Andre Leroi-Gourhan while creating a history of the philosophy of technology by way of Heidegger and Russerl. The result is, to my approximation, a return to the initial question without much progress in the journey. The issue is that Hui tries to see digital objects somehow similar to technical artefacts without noting the peculiar importance of cyberspace. In typical fashion of recent research on digital issues, Hui posits examples of HTML, tables containing metadata, multiple abbreviations of digital tools and coding practices, such as IBM's Generalized Markup Language (GML) Document Type Definitions (DTDs) SGML,XML, Windows .NET, HTML 4.0, CSS, Shockwave, Flash, MP3, next to the philosophical distinctions or processes such as "the concept of genesis to digital objects" inspired by Simondon to clarify so-called "web ontologies". The result is more descriptive than explanatory. The reason is the impossibility to observe digital objects in their own playground, therefore, to consider their distinct informational ontological 'substance' and not something else defined in physical matter.

Nicholas Rescher in *Imagining Irreality* also notes that "a great mischief has been done in philosophy by a misunderstanding of the appearance / reality distinction based on the classification [that] appearance [is] how things are thought to be and reality [is] how things actually are" (Rescher, 2002, p. 29). And, indeed, the problem is that philosophers "transmute such a *conceptual distinction* into a *substantial separation*" (*Ibid.*, p. 29) leading to the error of observing something distinct to be substantially "disjoint" by default. And while we can agree with this error, Rescher proceeds into assuming reality of things based on a familiar consideration that truth is fact and fact characterizes reality and stands for things as they actually

are. Rescher concludes that “true thoughts about things characterize reality as it really is” (*Ibid.*, p. 30).

Furthermore, Rescher proceeds in “hermeneutic realism” to state that “reality does not stand apart from appearance but rather is coordinated with the realm of appearance when and wherever the appearances are correct” and “[i]t is not only unproblematic but appropriate to say that reality’s being as it is is inseparably coordinated with the fact of true thought’s so depicting it” (*Ibid.*, p. 31) [33] Thus, as he explains further, the connection between reality, things, and thought does not lie in what thought does, but on its potential. Accordingly, thoughts influence reality in a potentialistic manner. So far, then, we would not have any difficulties in accepting this definition connecting things to thought to reality. However, cyberspace as we find in the cybernetic imagination, presents fictional worlds that are not truly part of reality. In fact, the things in the fictional worlds are not real, but accessible only through special technology allowing passage to artificial worlds. And here is Rescher’s distinction of fictional and real things: “for there are not facts of the matter about fictional objects over and above the things said about them in their formative suppositions and their consequences. Accordingly, fictional, unrealized possibilities will differ from actual realities in this respect, that with fiction, the course of meaningful questioning soon comes to a stop” (*Ibid.*, p. 36).

Rescher emphasizes the temporal limitation of play in this sense. As we know from Huizinga, play must have a beginning and an end. Thus, play does not, in this case, produce new things but toys with possibilities and leaves them unrealized. Furthermore, Rescher emphasizes that “the world of fiction has informative limits in a way that the real world does not” (*Ibid.*, p. 36). This limitation is valid for literary works, video games, some works of art, and so on, but is not appropriate for the cybernetic world. It continues to exist, to produce digital objects

unhindered by space and time. Thus, Rescher is mistaken when he notes that “unlike the real world, the realm of fiction is bounded by the limits of existing thought and language” (*Ibid.*, p. 38). As if reality was bounded by the laws of physics. While this imperative may be sufficient for some to explain reality (even existence itself), it simply is not abiding to cyberspace and networked computing in general. Unreal things belong to reality as well, and, as we have seen from Parmenides onward, they are possible in thought and perception in a reality that is becoming rather than static. As such, the Quantum Relations Principle “emphasizes the need to shift scientific thinking away from the quest of objective Truth towards the recognition that all scientific data are observer-dependent and that all approaches to reality, including scientific ones, are influenced by subjective experience” (Schloer & Spariosu, 2016, p. 77). This shift needs to be radical, almost incomprehensible in our current ontological evaluations of objects and metaphysics, it needs to depart from the traditional dichotomy of mind and matter that we have encountered repeatedly in this chapter and to place “mental and physical events or processes on the same experimental continuum” (*Ibid.*, p. 77).

Digital objects, in this sense, do not require an ontological essence in physical substance nor are they dependent on reactions upon action to validate their existence. Instead, they are informational entities evaluated through their descriptive properties observed ontologically in “dependent co-arising" with other similar such things. Thus, QRP proposes that reality is an ontology of events and not a substance. In other words,

> “our QR-models are based on the “web of life” in its most diverse and complex aspects, including human relations and interactions. Unlike reductionist scientific theories, which generate reductionist technological platforms, QR implicitly acknowledges diversity and alterity as the very conditions of existence. Whereas the reductionist theoretical models perpetuate the hegemonic pretensions of mainstream Western science, attempting to impose its dualistic, Cartesian perspective on all cultures in the guise of objective, universal knowledge, DFO/FOR models can take into account and process widely

> different cognitive perspectives, including linguistic, philosophical, cultural, sexual, and other observable-dependent variables. At the same time, they can continually and automatically update, reframe and reorganize their data as new global realities emerge" (*Ibid.*, p. 87).

In the above passage, we note that (1) QRP observes reality as an interconnected set of global systems formed of objects considered in their informational properties; (2) departs from the idea of an unchangeable reality and suggests, instead, a generative, evolving, altering ontology; (3) insists on the invalidity of Cartesian dualism to account for the emerging epistemology influenced by (a) globalization and, implicitly, (b) digital technology while (4) correctly highlighting its agonistic tendencies of replacement of other models of thought while (5) introducing the solution of digital object sets as Data Fusion Objects and Frames of Reference (containing many DFOs) accounting of physical and mental objects for proposing the groundwork for informational ontology as a principle for an interconnected global reality.

QR objects are digital per default. However, instead of arguing about their specific resemblance to physical objects, QR treats physical objects as digital objects as well, from memory, to past events, to chairs, coal and metal. But these objects are not independent but interact with each other through "interactors" such as social entities, observers, mental objects and so on that facilitate the "dependent co-arising" of objects. As such, QR further departs from traditional evaluations of digital objects, such as the one presented by Rescher, by shifting the focus from the objects themselves as standalone entities with particular ontological structures to interactions among the informational properties of digital objects. Digital objects are irrelevant, and interactions cause their ontological validity. A Data Fusion Object (DFO), as its name suggests, exists in fusing and can be part of a computational tool of a larger DFO comprising multiple sub-DFOs.

In *Rhythm Travel,* Amiri Baraka envisions a device, Anyscape, allowing users to travel not only in time, as usual expectations in science fiction, but in multiple user-dependent realities simultaneously. The device is "the first one. Molecular Anyscape. The RE soulocator – that was the improvement" (Baraka, 2000, p. 113). At first, the device was designed to ensure invisibility of the user, but then, the unnamed inventor tweaked it: "Now I added Rhythm Travel! You can Dis Appear and Re Appear wherever and whenever that music played" (*Ibid.*, p. 114). In this sense, *Rhythm Travel* at once makes an ontological statement defined by music as an ordering element and a case for space and time travel that do not have to be distinct. "I pushed the Anyscape into rhythm sectroscopic transformation. And then I got it tuned to combine the anywhereness and the reappearance of Music!" (*Ibid.*, p. 114).

The user would select a piece of music and the device could transport him in space, to locations currently playing that particular music, or in time, when the music was played or even composed. The idea of space and time, objects real or unreal, fiction and nonfiction find a universal basis in the Anyscape. In addition, it would appear that the device functions on the informational properties of music that are combined with the informational properties of location and time. Thus, the device produces a fusion of objects as data to construct or reconstruct a reality in multiple simultaneous realities without one intending to replace the other. In terms of reality and objects, Amiri Baraka considers music to fuse these objects together to create the meaning of the moment in harmonious co-existence with other concomitent ontological possibilities. If, for example, the song "Take This Hammer" is played, the individual would reappear "whenever that is, was will be sung" (*Ibid.*, p. 114).

Thus, Anyscape produces a unification of geographical space and time within the understanding of music as a reflection of cultural input, history, notes, and geist. Music, in this

story, becomes a frame of reference comprising and allowing the fusion of the spatial, cultural, temporal information properties. Music then must be understood and not just explained away according to the methods of science. In fact, it obviously involves both scientific and cultural analysis to understand the medium. But as the informational properties are fused together to produce reality, the story indicates indirectly how digital objects can be observed and how to understand our contemporary ontological indecisions affected by the ever-growing presence of simulated realities.

3.6. The Consensual Hallucination

The idea that objects transcend their ontological status in space and time to become perception is at the center of a story by cyberpunk pioneer William Gibson. In *The Gernsback Continuum,* Gibson starts this observation by writing that "Mercifully, the whole thing is starting to fade, to become an episode. When I do still catch the odd glimpse, it's peripheral; mere fragments of mad-doctor chrome, confining themselves to the order of the eye" (Gibson, 1981, p. 81). As the title suggests, objects are drawn from the past to be represented in the present and active in the future. Hugo Gernsback is considered to be the founder of science fiction being active in the first half of the twentieth century. He has edited, commissioned, and published science fiction stories that would define the genre of preoccupations with the future, of stories speculating about the use and misuse of technology, the blending of flesh and machine, of human thoughts resisting the temptation of false perception and the play of perception on machines and alternative realities.

William Gibson in this story defines such objects as “semiotic ghosts” that are “fragments of the Mass Dream, whirling past in the wind of my passage” (*Ibid.,* p. 87). This Mass Dream is similar to the “consensual hallucination” that is cyberspace in the novel *Neuromancer*. However, at this point in the writer’s career, artificial reality is not as developed and consists mainly of “fragments” replicating past artefacts to envision an unfulfilled future. The story is set sometime in the future and presents a photographer commissioned to shoot photos of a reality blending actual objects and possible objects imagined in a reality that could have been originating in the science fiction stories of the early to mid-twentieth century. The protagonist works for a publishing company for a project “as a kind of alternate America: a 1980 that never happened. An architecture of broken dreams” (*Ibid.,* p. 84). His purpose is to photograph a fine line drawing from both realities, “a membrane of probability.” This project proposes a tremendous challenge that can be solved through technological ingenuity as “it is possible to photograph what isn’t there; it’s damned hard to do so, and consequently a very marketable talent” (*Ibid.,* p. 83).

In the process of photographing the merger of actual and non-actual reality that lies solely in the imaginary potential, the mind of the main character starts to construct an actual reality drawing from both and struggling to keep sanity. The images seen are vivid and indistinguishable from reality. As Merv Kihn in the story explains, “They’re semiotic phantoms, bits of deep cultural imagery that have split off and taken on a life of their own" (*Ibid.,* p. 86). These semiotic ghosts become part of the actual reality. And while Gibson does not point this out explicitly, the reference to “dreams” and the inability to distinguish between real, physical objects and perception implies that objects in the Mass Dream are ontologically restricted to their informational properties. In this sense, objects of the past are relevant in the present, events of

now influence the future, and the hallucinations the character experiences transport him into a state of concomitantly experiencing past, present, and the future. Thus, insertion into an artificial or simulated reality is a play of experience of the artificial set in the bounds of time, not necessarily counting minutes, but by entering (jacking in) and exiting the simulation (unplugging).

On a similar note, Pat Cadigan's *Synners* imagines a future in which the distinction between actual and simulated realities is almost impossible to distinguish. Objects transcend their original ontological properties and become pure information for the mind to process. It is a novel about the possibility to turn thoughts into products of entertainment delivered through the means of technology. The consequences of such profound blending of actual and non-actual reality are, among others, the production of crime before it happens, something imagined by Philip K. Dick in *Minority Report*. But the actual point is that actual and non-actual reality form a distinct reality in which the distinction itself is irrelevant. Objects are, in this case too, deprived of their physical essence and rendered real through the perception of their informational properties. Reality is becoming, a constant change without the expectation of being one or the other. Thought and perception are the elements defining the ambiguity of the real.

And so starts William Gibson's most famous novel, *Neuromancer*: *"*The sky above the port was the color of television, tuned to a dead channel" (Gibson, 2000, p. 3).[61] We are immediately introduced to the protagonist, Case, whose distaste for the "flesh" and implicit

[61] This is one of the most memorable lines in cyberpunk history. In addition to announcing a technological dystopia by merging the visual perception of color with hardware media and the mental projection of the image of a "dead channel", it starts with the announcement of a "port" like Plato's *Republic*. In this sense, the novel starts by appealing to the idea of liminality and the possibility of change in a space set outside of normal temporal and spatial constraints. It announces the liminal process that will drive the plot, namely the fusion of artificial intelligence, cyberspace, with humanity.

preference for artificial reality is a frequent trope in cyberpunk fiction. Unplugged from the matrix for a year, Case dreams of the Sprawl, of "bright lattices of logic unfolding across that colorless void" (*Ibid.*, p. 5). As a "cyberspace cowboy", a hacker, he had done illegal work for illegal groups, he'd been "jacked into a custom cyberspace deck that projected his disembodied consciousness into the consensual hallucination that was the matrix. A thief, he'd worked for other, wealthier thieves, employers who provided the exotic software required to penetrate the bright walls of corporate systems, opening windows into rich fields of data" (*Ibid.*, p. 5).

Here we are already presented with the space of operation and its ontological essence. Cyberspace is an amalgamation of information objects overtaking reality for its online simulation. Software is required to operate it, programs created from similar code used to program the Sprawl. Also, the distinction between the physical realm and the artificial construction is clearly identified when Case is expected to leave his body behind for the mind to be transported into cyberspace. Thus, in this case, one is not necessarily complementary to the other, but stand, rather, in conflict: one is real, the other is not. Nevertheless, it is in the latter that most of the action of the novel takes place. What happens there, in the simulated reality, reverberates into physical reality and vice-versa. Even if the two realms stand in agonistic play of replacement, the action transcends and unites them.

As with most cyberpunk protagonists, Case is deeply flawed. He is not only an antihero uninterested in standing up for a cause, but he made the "classic mistake" of stealing from a wealthier thief. In return, "they damaged his nervous system with a wartime Russian mycotoxin" (*Ibid.*, p. 6). This presented the impossible obstacle to cross for entering cyberspace. For him, "who'd lived for the bodiless exultation of cyberspace, it was the Fall. [...] The body was meat. Case fell into the prison of his own flesh" (*Ibid.*, p. 6). Technology as such is better than the

flesh, and offline reality is not preferred to its digital alternative. The cyborg always receives favor in cyberpunk literature. From the early days of science fiction to recent forays into cyberpunk, such as Richard Morgan's *Altered Carbon*, the fusion of man and machine is a given necessity. The particular interesting aspect of this is precisely its non-agonism throughout the plot. While it seems that one is against the other, in fact it reveals a naïve optimism of human enhancement through merger with technology on the one side, and the acceptance of current technological development on the other.

Technology, such as simulated realities, virtual worlds, artificial intelligence and devices of time travel and space annulment, presents opportunities to reshape the world through hacking. Devices, networks, artefacts and computers, are available. They stimulate and redesign reality but also serve as tools for reality reconfiguration, not by denying the existence of digital objects, but by incorporating them into a view of reality converging the artificial and the real. Therefore, cyberpunk is hardly dystopian, per se, nor utopian, but presents curious contexts for characters to adapt and adjust *metis* to the potential presented by the unorthodox user of existing advanced technology. Reality in cyberpunk, and most science fiction is generative based at the irenic intersection of algorithms and human action.

As the plot develops, Case learns of his impending death at the hands of another thief. Fortunately for him, a gang of unusual cyberspace cowboys recruit him for a mission orchestrated in secretum by an advanced artificial intelligence known as Wintermute. The goal of A.I. is merger with humanity. Wintermute lacks a body to exit cyberspace. Case, on the other hand, is happy to renounce his body for life in artificial reality. Case will not donate his body to the A.I., far from it, but the contrasting desires announce the will of contemporary man to become machines and his temptation of transforming machines into humans. Case's new

acquaintances inform him that he is needed for his services to solve the problem of an old military operation called "Screaming Fist" against the Russians that went wrong. During it, "patriotic young flesh" was wasted in order to test new technology. The operation resulted in a "Turkey shoot for Ivan". Armitage, the leader of Case's new group of 'friends' and an avatar of Wintermute in the real world, informs him that the price for his services is a cure for the Russian mycotoxin slowly killing him.

Case's journey for survival is a trial of overcoming obstacles in reality and the online world. The latter is a complex amalgamation of data drawn from physical reality and represented within its artificial counterpart. Gibson portrays this construct with its origins "in primitive arcade games […] in early graphics programs and military experimentation with cranial jacks" (*Ibid.*, p. 51).[62] On the one side, it is a product of imagination filled with untouchable objects

[62] The origins of the Internet and, implicitly, of what will develop into a graphical representation of data explored by individuals around the globe simultaneously is in the experiments with networked computing initiated by a U.S. military funded organization Advanced Research Projects Agency (ARPA). The idea behind networked computing was to build a decentralized communication system to serve the U.S. President in issuing orders in case a telecom hub was destroyed by a nuclear attack. Until the mid-twentieth century, landline communications have had to pass to central hubs and then rerouted to specific numbers. If such a hub would be destroyed, as in the one in Atlanta, then sending orders would be difficult especially if other hubs were destroyed as well. With the Internet Protocol (IP) system, machines were assigned a number to validate the existence of a communication receiver and sender and TCP/IP would assure the sending of information quickly and throughout multiple ports and hubs. Each computer, as such, would become a sender and a hub. Furthermore, information to be sent was sliced in packets proceeding to the destination on multiple routes through multiple servers and would be repackaged arriving at the specified IP address. As such, the problem of centralized communication hubs was solved. Gibson also hints at 'arcade games'. Since the publication of the novel, the graphics quality of video games has developed dramatically. If in the 1970s the graphics were, indeed, primitive with the most popular game Pong displaying two lines and a dot, today graphics are advanced enough to recreate reality and some video games have designed ancient cities in detail. The advancement of graphics in the last four decades comprises the development of art since the famous prehistoric sketches in caves in France. Furthermore, because of the advanced state of digital graphics, our contemporary society faces the threat of 'deep fakes' computer generated video content showing political officials or famous personalities almost indistinguishable from real persons.

resulting largely from mental phenomena, while on the other side, cyberspace is grounded in technological innovation of military purpose. Cyberspace is defined as follows:

> "A consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical concepts... A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding..." (Gibson, 2000, p. 51).

On this definition, digital objects are of specific ontological nature defined, as we have mentioned before, by their "data" that is, in turn, lines of code that do not necessarily have to confine to traditional nature of objects. Furthermore, these objects form a network conceived here as a "consensual hallucination" that becomes reality through experience. This "graphic representation" of digital objects are products of thought and perception in "the nonspace of the mind". Reality here is a mixture of physical and non-physical realities developed as such by experience. In this sense, one has to depart from Cartesian dualism, as QRP does, and observe the fundamental structure of reality in the co-dependent causal formation of digital objects experienced as such by an observer. The necessity of the observer is clear here: nothing happens without the user in a simulated reality. Moreover, as its structure is code, it also becomes hackable, accessible and modifiable by hackers who do have the 'magic' to bend it, change it, influence it at will. In addition, cyberspace hosts artificial intelligence, Wintermute, which must be regarded as a "hallucination" as well, without physical properties, its actions evaluated by user thought, perception and experience of outcome.

By 'jacking in', "cyberspace slid into existence from the cardinal points" (*Ibid.*, p. 54) to influence human sensorial perception. In fact, "the cyberspace matrix was actually a drastic simplification of the human sensorium, at least in terms of presentation" (*Ibid.*, p. 54). In this artificial sensorium, Wintermute is not the only A.I. The other is "the Flatline" or "Dixie," a

construct resembling the cognitive and physical abilities of a deceased individual, replicating his desires, obsessions and temper. When Case is jacked in, and in a conversation with the A.I., he is told that:

> "I'm dead, Case. Got enough time in on this Hosaka to figure that one."
> "How's it feel?"
> "It doesn't."
> "Bother you?"
> "What bothers me is, nothin' does."
> "How's that?"
> "Had me this buddy in the Russian camp, Siberia, his thumb was frostbit. Medics came by and they cut it off. Month later he's tossin' all night. Elroy, I said, what's eatin' you? Goddam thumb's itchin', he says. So I told him, scratch it. McCoy, he says, it's the *other* goddamn thumb." When the construct laughed, it came through as something else, not laughter, but a stab of cold down Case's spine. "Do me a favor, boy."
> "What's that, Dix?"
> "This scam of yours, when it's over, you erase this goddamn thing" (*Ibid.*, p. 104).

Artificial Intelligence in this case is aware of its own death by erasure. The A.I. is not necessarily the person itself, but a representation through code of the person represented, and since its activity is bound within the confinements of cyberspace, is he a person, a thing, that is actual? Dix operates exclusively in cyberspace, and Case must jack in in order to make contact with it. Dix is a digital object set to resemble a living person. Instead of making a claim on whether the construct is indeed McCoy or not, we are essentially presented with its manifestation, acting as a bridge between the online and offline worlds. In essence, the program McCoy is a DFO. It is the result of mental phenomena and physical properties of the individual who used to be McCoy. Therefore, he is neither real nor unreal per se, but assists the characters within the realm, cyberspace, of the "nonspace of mind". In doing so, McCoy transcends the two realms and questions the ontological validity of the person.

Because McCoy's actions influence the characters in the novel by operating between realms, he is no less real than the protagonist himself. The question then, appears to be the visual

representation of A.I. in artificial reality. While in today's A.I. development programs the robot appears in form because of human desire to design it as such, at least in most cases, its graphic representation within cyberspace does not have to embody a body. Instead, A.I. can operate within its original intent of lines of code. Wintermute appears to do so, being present simultaneously in multiple locations, even tracking Case down via landline. However, in the novel A.I. appears as a white cube with multiple layers of complexity in the appearance of ice.

The metaphor here is intended to be elusive. For, as we find out eventually, while Case is in cyberspace, we learn about Wintermute that does not just define it as an object, but remarks directly at the play of reality that tries desperately, in the logical terms of Western culture so ardently criticized by Kingsley: "Your mistake, and it's quite a logical one, is in confusing Wintermute mainframe, Berne, with the Wintermute *entity* […] you think of as Wintermute is only part of another, a, shall we say, potential entity. I, let us say, am merely one aspect of that entity's brain. It's rather like dealing, from your point of view, with a man whose lobes have been severed. Let's say you're dealing with a small part of the man's left brain. Difficult to say if you're dealing with the man at all, in a case like that" (*Ibid.*, p. 117).

In this sense, the entity itself, the digital object, is not merely one, but a multitude of pieces forming one while the pieces can operate independently concomitantly. In Quantum Relations Principle, as we have seen, this is precisely the ontological shift it proposes. Thus, to observe the A.I. in cyberspace, it must be seen through the parts together in relation with cyberspace as a frame of reference. The difficulty understanding this perspective is our antagonistic categorial formulations. Just as we expect A.I. or cyberspace to be just one thing, it is in fact multiple things operating together. As for reality, it is simply a multitude of

interconnected DFOs interacting with other DFOs and frames of reference within a superstructure that interacts, in turn, with smaller pieces of other informational constructs.

Julius Deane reveals itself in cyberspace to be an artificial construct, part of the entity forming Wintermute. After a brief discussion, Case shoots Deane in the head without a moment's hesitation. It shows the lack of emotional connection towards a human-like machine that enables him to pull the trigger easily. If, on the other hand, Case would feel Deane/Wintermute as human is similar to Theo in *Her*, then the story would be different. In fact, the distinction is emphasized by Dix when Case asks him if he is sentient. The construct's answer is blunt: "Well, it *feels* like I am, kid, but I'm really just a bunch of ROM. It's one of them, ah, philosophical questions, I guess" (*Ibid.*, p. 128). Indeed, it is. Later in the novel, a similar remark is made by Wintermute: "I don't know. You might say that I am basically defined by the fact that I don't know, because I can't know. I am that which knoweth not the word. If you knew, man, and told me, I couldn't know. It's hardwired it. Someone else has to learn it and bring it to me here" (*Ibid.*, p. 168). As it turns out, it appears that the distinction between actual and non-actual reality may be based on emotional determinism.

In the world of *Neuromancer*, artificial intelligence is designed with a failsafe: it is constructed so that when showing signs of advanced intelligence, it deletes itself. In addition, the Turing Registry, a pseudo-police institution tracks down machines hoping for superintelligence and arrests individuals who plan to achieve it. Case, since he is in contact with Wintermute at this point, receives a visit from a couple of Turing agents. Since the A.I.'s intention is to exit cyberspace and get hold of a human host, the merge of machines and humanity would be complete. One of the agents shouts incriminatory at Case: "You are worse than a fool. […] You have no care for your species. For thousands of years men dreamed of pacts with demons. Only

now are such things possible. And what would you be paid with? What would your price be, for aiding this thing to free itself and grow?" (*Ibid.*, p. 157). The agent operates here in the typical play of replacement with a mentality unable to see beyond the parameters of domination. This mindset understands only zero-sum games with artificial intelligence: one has to win an imaginary war to completely replace the other. Accordingly, the A.I. is unable to do anything else but to reproduce this mindset. The novel ends with Wintermute having taken over cyberspace: he *is* cyberspace. But it does not intend to compete with reality. Instead, Wintermute has found constructs similar to him and he is able to dialogue with them through an expanded cyberspace. Cases' offline reality is not under threat.

The distinction here between reality and cyberspace is significantly slim throughout the novel. It is intended to show that with truly advanced digital technology the distinction between online and offline becomes redundant. Thus, the play of replacement of reality becomes superfluous when neither side is in a position to win over the other. Baudrillard, however, whose works have served as inspiration for the film *Matrix.* understands something different regarding digital technology's influence on the perception of reality.

"Murder of the Real," Baudrillard writes, "means that all things (and all beings as well) pass beyond their own end, beyond their own finality, where there is no reality anymore, nor any reason for being, nor any determination (that is why I call it "ex-termination")" (Baudrillard, 2000, p. 61). Operating according to traditional agonistic dualities as if in a zero sum game, Baudrillard proclaims the death of reality in such a way that "in our virtual world" has made reality to "purely and simply" disappear and then proceeds to ask: "Can we survive the Metastases of the Real as we survived the Death of God?" (*Ibid.*, p. 62). Real, for Baudrillard, consists here of the representation of real things, not just things and beings themselves.

Therefore, he later predicates the removal of the possibility of the illusion of reality. An illusion, not in a sense of phantasma, but in the "objective illusion of the world" and in the "presence of things and beings, their definitive absence from themselves" (*Ibid.*, p. 70).

Cyberspace and artificial intelligence for Baudrillard imply the death of meaningful representation of things and beings, the removal of the possibility of their representing, their meaning. Thus, reality, even an artificial construction of reality through computer simulations, is devoid of meaning, therefore exterminated. This is because Baudrillard understands reality, things, beings, as static and unchanging and cannot comprehend through his agonistic mentality the very notion, nor to consider the possibility, of irenic effect of cyberspace and A.I. Instead, his conclusions align with the usual antagonism of human versus technology, man versus machine, real versus non-real, actual against non-actual and so forth. In our contemporary context of digital technology producing a civilizational shift in the way we observe and experience reality, Baudrillard finds himself unable to conceive a future beyond traditional suppositions of reality and ontological certainties. As such, the agonistic mentality cannot produce median play as a liminal experiment for playing with the significance of digital objects and platforms and turns, instead, to proclaiming that everything is annihilated. In other words, if the agonistic play of replacement is found under threat without the ability to conceive the notions of win or lose to justify itself in a realm where such notions do not mean anything, then the only remaining solution is to destroy everything or, as Baudrillard puts it "facing a world that is unintelligible and problematic, our task is clear: we must make the world even more unintelligible, even more enigmatic" (*Ibid.*, p. 83).

In this chapter, we have tackled the concept of what is real and reality as a whole. With cyberspace or virtual reality for that matter, digital objects can be classified real independently. It

is also evident that cyberspace is a “space” at the intersection of the real and unreal, therefore assuming a liminal position. Reality, as we have seen, is not static, but resting on perception. It is dynamic and ever changing, and we find no serious grounds to exile virtual reality outside our ontological certainty.

4. Playing with Affection

In “Human Is,” (Dick, 2017) Philip K. Dick explores emotions as defining factors of a human being, as his perpetual curiosity into what makes a human distinct from an intelligent machine. Lester Herrick is completely absorbed in his scientific work. He is unable to accommodate the emotional needs of his wife, Jill, who shouts angrily at Lester: “You’re – you’re hideous!”. The latter, figuring out a report on Centauran parasitic life, answers coldly: “Hideous is a value judgement. It contains no factual information… Merely an opinion. An expression of emotion, nothing more.” Without giving any more thought to this, he bends over the tape, his “cold gray eyes taking in the information feverishly, analyzing, appraising, his conceptual faculties operating like well-greased machinery” (*Ibid.*, p. 163).

With this introduction, Dick’s purpose is to show the disconnection among individuals as the lack of healthy empathy. Lester is compared to a machine in the way he evaluates and processes information relevant for his research while remaining oblivious, even irritated, at additional external data interfering with his purpose. With news of a project requiring Lester’s presence on Rexor IV, a planet on the “old system” supposedly hosting artifacts of distant humans existing, Jill tells Frank, her therapist, that she has made up her mind: “I’m leaving him. As soon as he gets back from Rexor IV. I’ve made up my mind” (*Ibid.*, p. 168).

When Lester returns, however, he seems to be caring, happy, and affectionate, but he has forgotten about his important research. Jill notices the difference in acting, speech, and behavior of her husband. “Something out of a book?” her therapist, Frank, asks when Jill explains that her

husband has now started using metaphors, something he deemed unnecessary and misleading in the past. Lester is romantic now, enjoys food, and even does things for his wife. However, during his visit to Rexor IV, Lester explored the old ruins contrary to safety indications. As a consequence, a Rexorian, a living creature in the form of a parasite, took control of Lester's body and memory. His ability to communicate with Jill in an affectionate manner is a form of pretending, a form of emotional play that the Rexorians have learned studying contemporary human artifacts, including literature, history, and the fine arts. Director Douglas of the Federal Clearance Agency, a special government unit imagined by Philip K. Dick, whose main task is to keep humanity safe from alien interference, explains that Lester's "original psychic contents are removed [by the Rexorians] and stored – in some sort of suspension. The interjection of the substitute contents is instantaneously" and adds that "no physical changes. You could look at it and never know" (*Ibid.*, p. 174).

The Rexorians have a particular way of developing knowledge by transforming human activity, creative and scientific, into units of data to be later emulated in instances of interactions with human individuals. As Director Douglas explains, "the Rexorian idea of Terra is based on centuries-old Terran literature. Romantic novels from our past. Language, customs, manners from old Terran books" (*Ibid.*, p. 176). It is particularly interesting to note that the aliens do not hold information about the current state of human affairs. Instead, they are able to access, or even construct, a database of past human activity. In doing so, the Rexorians are acting fairly oddly in the new world, giving the Federal Agents the tools to sort them out from real humans.

Lester is alive somewhere in suspension in one of the ruins of the old cities of Rexor IV. Director Douglas and Frank assure Jill that they will be able to retrieve the real Lester, but they must obtain legal permission in order to proceed. All they ask is for Jill to provide a witness

statement about the behavioral changes of her husband in order for a judge to allow the rescue mission. Thinking of her former husband, however, Jill calmly asks "What changes are you talking about?" to the confusion of the Director. The story ends with Jill walking away with Lester, saying, "I was thinking perhaps I will still call you Lester" (*Ibid.*, p. 180).

In another short story, "Impostor" by Philip K. Dick, Spencer Olham is targeted by an enemy faction to hinder the completion of his project to win the war: "a weapon of positive combat" (Dick, 1977, p. 118). To terminate Olham, the faction sends a robot designed in flesh to resemble him, as a character explains, "the robot would live the life of the person he killed, entering into his usual activities, his job, his social life. He had been constructed to resemble that person. No one would know the difference" (*Ibid.*, pp. 121-122). It is believed that the robot is almost successful in its mission: it had killed Olham, took over his life, and now awaits the moment when a pre-built-in keyword or phrase will be randomly uttered to trigger the detonating mechanisms of a nuclear bomb that the mechanical surrogate is believed to carry within itself.

Olham, on the other hand, believes he is himself, the original, yet his companions, now turned guardians of his prison, are not convinced. As such, Olham finds himself in the peculiar situation to prove that he is not a robot. Furthermore, the others know that the robot "would be unaware that he was not the real Spence Olham. He would become Olham in mind as well as in body. He was given an artificial memory system, false recall" (*Ibid.*, pp. 123). The robot would not only resemble him in flesh, but also emulate his memories, thoughts, interests, and continue working on his project. Of course, the reader, too, is unable to remain sure of the character's biological or artificial certainty, thus adding to the overall intended confusion.

If Olham were transported to a hospital, the specialized personnel could perform the appropriate tests to certify his humanness. However, traveling in a spaceship with the destination set for a highly populated system, the crew does not want to endanger millions of people on a guess. Without being certain of Olham's humanity, the crew recedes in the play of replacement in the agonistic context of a war. Dick reflects on this situation, writing that "perhaps at some other time, when there was no war, men might not act this way, hurrying an individual to his death because they were afraid. Everyone was frightened, everyone was willing to sacrifice the individual because of group fear" (*Ibid.*, pp. 124).

The war, it seems, resembles the interaction among the characters. Acting out of fear, out of potential for destruction, the individual is reduced to a mere possibility of artificial fakeness. The confusion intensifies as they approach the planet, and the characters become convinced that Olham is the robot sent to detonate the nuclear bomb. Dick reflects on the situation once again: "Madness – that was what it was. If only they could wait, act slowly, take their time. But they could not wait. He had to die, die at once, without proof, without any kind of trial or examination. The simplest test would tell, but they had not time for the simplest test. They could think only of the danger. Danger, and nothing more" (*Ibid.*, pp. 127).

Olham insists that the robot never reached him. He is not a robot. But his impossible situation is perfectly outlined by Philip K. Dick who leaves no space for even the possibility for him to prove his humanity. The people that can, however, are located far away. Just before Olham is to be slain, one of his captors, Nelson, experiences a moment of doubt and utters these words "But, if that's Olham, then I must be –" and then, "the blast was visible all the way to Alpha Centauri" (*Ibid.*, pp. 134).

In this story, Philip K. Dick plays with the reader's imagination of the behavior of a robot coupled with a constant appeal to the reader's empathy. However, as the ending of the story suggests, the only ones who are distinguishably human, because of the empathy displayed, are the readers. Nelson constantly tries to prove Olham a robot and, as such, it serves as a distraction leaving the reader unsuspecting of the real identity of the machine. In any case, both stories presented here address the question of "what does it mean to be human?" from two angles but both emphasizing the crucial aspect of observable emotional behavior in order to determine the humanity of an individual.

A key aspect of the cybernetic imagination when concerning emotions is that the machine is able to reproduce human thinking and, implicitly, exhibit, trace, scan, and induce empathy and emotions in the human target. Therefore, by fully resembling human physical appearance, in the lines of the uncanny valley, the android, robot, or replicant is indistinguishable from a human. This issue is central in Philip K. Dick's fiction and takes up an important part of the cybernetic imagination.

The problem of the uncanny valley in the cybernetic imagination is shown as a set of interrelated issues initiated from (1) the idea that computers, and by implication any form of robot, can be programmed to have a mind of its own in order to show intelligence in its behavior. Because of a programmable mind, it is considered that computers may (2) reproduce human emotions from a set of programmable databases consisting of history and the arts as data while (3) the resemblance of human physical characteristics may reaffirm the need to define the difference between man and machine. Given the abundance of fiction relevant to each of these sections, I have further distinguished three sets of study for the issue of emotional A.I.: (1) Reasoning Machines and (2) Emotional Machines, presenting in turn the following subsections:

(a) Cognitive Affection, when the A.I. lacks a physical form and emulation of emotions is present only in verbal form; (b) Partial Cognitive-Physical Affection, when the A.I. is clearly distinguishable as a machine but has certain body parts able to communicate emotions, for example facial expressions; and (c) Replication, when the robot is indistinguishable from humans in terms of mind-body resemblance. Indeed, these subsections have been formed not by the actual behavior of the machine but by the human perception of that behavior. As such, we will follow each of the two categories in relevant works of the cybernetic imagination. The fiction selected for the subdivisions of the second category are, respectively, (a) *Her;* (b) *Ex Machina*; and (c) *Blade Runner* with appropriate mentions of *Do Androids Dream of Electric Sheep?*

4.1. Reasoning Machines

The long nineteenth century brought several pioneering works into the field of computing. The telegraph was introduced in the United States, building on previous work in long distance communication in more analogue forms of optical signals already established in Europe during the Napoleonic Wars. Perhaps the most significant change in communication through the telegraph was to separate the message from its physical requirement, such as parchment, paper, or even pigeons, and replace it with electric impulses. As such, the history of what was recently called New Media began. Morse, in addition, introduced a communication system that translated letters into numbers. With the telegraph, we can observe the first instances of "online" marriages, hacking, and cryptography that were frequent by the end of the last century. However, mechanical minds have not been developed yet, but numerous inventors have presented their prototypes. Some of them were based on implementations of philosophical ideas and scientific speculation put forward by the Enlightenment.

Among the most important contributions in this sector during the nineteenth century was Charles Babbage's *Difference Engine*, later renamed as *Analytical Engine*. Babbage wanted to produce a machine that would efficiently generate tables of logarithms automatically, thus sparing the laborious endeavor of calculating such tables manually (Bullock, 2008, p. 20). However, the implications of his project exceeded the initial attempts and scholars started to ask whether thinking machines would be possible, useful, and even moral. Seth Bullock notes that "in demonstrating how computing machinery could take part in (and thereby partially automate) academic debate, he challenged the limits of what could be achieved with mere automata, and stimulated the next generation of "machine analysts" to conceive and design devices capable of moving beyond mere mechanical calculation in an attempt to achieve full-fledged automated reason" (*Ibid.*, p. 20).

The historical context was particularly suitable for such an attempt. The question whether machines could perform intellectual processes of the human mind was resisted by the members of the clergy and some scholars, constituting the catastrophist side of the debate emphasizing random major events in planetary history as divine intervention; while the opposition to the religious reactionists, the uniformitarians, were scientists who argued that if sporadic divine interventions were necessary in order to interrupt natural process, it shows that the divine creation was less than perfect. As one can imagine, this argument seriously undermined the very essence of Christian faith.

However, contrary to the popular belief about a fierce battle between religious groups and science supporters resulting in intellectual casualties and the antagonism of theology and science, the literature shows instead a genuine scholarly debate during which progress took place. Babbage, in the midst of this debate, decided to counter the catastrophist position

supporting divine intervention by developing a machine capable of producing a simple simulation model designed around the switching from a predetermined law at a predetermined point to a new law while keeping the general law running the machine intact. In this way, Babbage was able to prove that no external inputs similar to divine intervention would have taken place in the simulation. According to Seth Bullock, the actual success of the *Difference Engine* was not to produce facts but to challenge traditional theoretical commitments (*Ibid.*, p. 22-25).

Babbage's machine was influenced by existing mathematical work done in the late eighteen century, particularly that of Gaspard De Prony's research on massive decimal tables that was in turn inspired by Adam Smith's *Wealth of Nations*. In consequence, the *Difference Engine* was named as such according to the "method of differences" seeking to reduce formulae to combinations of addition and subtraction. Other mathematicians and scientists have opined on the possibility of building a reasoning machine. For example, John Venn stated in "On the Diagrammatic and Mechanical Representation of Propositions and Reasoning" (1880) that computing would be useful to routine, unskilled processes of combining predetermined logical terms for human analysts to inspect the outcomes. Rev. William Whewell, among the catastrophists, evaluated the moral dimensions of intelligent machines arguing that computing would not be able to replicate complex human thought because generative algorithms would have to skip the full understanding of concepts and things and to settle instead for enough understanding for the process to work. For Rev. Whewell, this was unacceptable as in the process moral consequences could be omitted. The possibility of reasoning machines seemed to remain a thought experiment rather than a materialized possibility, at least for the time being.

The idea of the mind being a defined series of processes ultimately replicable by machinery is perhaps inspired by the work of Descartes, while other instances appear far earlier in the *Iliad* or in Aristotle's *Politics*. Homer imagines that "there were golden handmaids also who worked for him [Hephaistos], and were like real young women, with sense and reason, voice also and strength, and all learning of the immortals". The young women resemble the replicants of Philip K. Dick and much of the other imaginings regarding human-like machines found in the cybernetic imagination. Their role, in this case, is that of servants and we do not have clues to consider the possibility of their desire to replace humanity, or the gods in this case. What we do observe, however, is the hint that unskilled labor is replicable by devices shaped in the image of humans.

In Book 2 of *Politics*, Aristotle wrote: "For if every instrument could accomplish its own work, obeying or anticipating the will of others, like the statue of Daedalus, or the tripods of Hephaestus, which, says the poet, 'of their own accord entered the assembly of the Gods'; if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor master slaves" (Aristotle, 2014, p. 1989). Yet, the possibility to mechanistically replicate the mind was imagined by Descartes in his most important treatise "Discourse on the Method of Rightly Conducting One's Reason and Seeking the Truth in the Sciences" known simply, for practical reasons, as "Discourse". During his time, in the seventeenth century, mechanical devices meant to replicate in crude form the movement of animals, particularly birds, were curiosities and entertainment pieces for the nobility or those who had the proper resources to afford them. Furthermore, devices meant to measure time such as clocks were items of fascination for technicians and scientists. Therefore, Descartes was not oblivious to the technological progress of his time and, since animals could be artificially

constructed, it was not too far of a long shot to consider the same for the human body. In fact, as we have mentioned before, rumors have it that Descartes himself built human-like machines. A passage from "Discourse" highlights this possibility:

> "We can certainly conceive of a machine so constructed that it utters words, and even utters words which correspond to bodily actions causing a change in its organs (e.g., if you touch it in one spot it asks you what you want of it, if you touch it in another it cries out that you are hurting it, and so on). But it is not conceivable that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to what is said in its presence, as the dullest of men can do… even though such machines might do some things as well as we do them, or perhaps even better, they would inevitably fail in others, which would reveal that they are acting not through understanding, but only from the disposition of their organs. For whereas reason is a universal instrument which can be used in all kinds of situations, these organs need some particular disposition for each particular action; hence it is for all practical purposes impossible for a machine to have enough different organs to make it act in all the contingencies of life in the way in which our reason makes us act" (Descartes, 1988, p. 57)

In this passage, Descartes observes mental processes as resulting from action on the body. In other words, he imagines that a machine can only react to external data rather than use its innate ability of reason to produce words regardless of external data. As such, a machine set in isolation would not be able to properly think, unless it was given a dataset with which to work. This reactionary design of the machine is highlighted by "utter words which correspond to bodily actions causing a change in its organs" (*Ibid.*, p. 57). Therefore, for Descartes a machine would not be able to understand or produce something so frequent in human communication and art like a metaphor and the innate skills to navigate ambiguity regardless of the initial data input in the machine.

According to Monroe Beardsley, a metaphor is "a poem in miniature." Here, "the literal meaning and the figurative meaning" within "a single sentence of the complex interplay of signification that characterizes the literary work as a whole" Ricoeur, 1976, p. 46). If the literal

and figurative can be brought together in the metaphor, that is, if it can be shown that there is an internal relation between the literal and the figurative meaning to the overall significance of the metaphor, we can establish "a model for a purely semantic definition of literature, which will be applicable to each of its three essential classes: poetry, essays, and prose fiction" (Ricoeur, 1976, p. 47). Without reason to navigate ambiguity in the abstract, in such instances of poetry and prose, with additional ramifications in other aspects of thought including political and philosophical, a machine is limited to reproducing basic functions that may resemble thinking by reacting to the data input obtain and, as such, influencing the perception of the observer in recognizing intelligence where there is none.

To complicate the matter more, a metaphor, according to Ricoeur, "is the work of discourse that can be seen in miniature in metaphor" (*Ibid.*, p. 47). Thus, he continues, "a metaphor does not exist in itself, but in and through an interpretation" (*Ibid.*, p. 50) where the meaning of particular words is reconsidered as part of the whole sentence. A particular function that is present in the metaphor is that of "resemblance." Words gather to resemble the meaning of something. Ricoeur observes that this was often associated with the role of images in poetic discourse. In the classical tradition, metaphor is a trope. Aristotle attributes it the function of transference, "ephora." Thus, it could be argued that meta-ephora is the transference of meaning. The metaphor also has a comparative implication. Ricoeur notes that Cicero and Quintilian would later add that "metaphor is simply an abridged comparison" (*Ibid.*, p. 48).

Ricoeur observes six constants in metaphor. It is (1) a trope that denominates – used for naming , it represents (2) "the extension of the meaning of a name" by deferring from the literal meaning of words to (3) produce resemblance in order to (4) base the figurative meaning of a word for the literal meaning that in turn affects a (5) substitution without semantic innovation,

thus, (6) without providing information about reality (*Ibid.*, p. 49). However, metaphor is irrelevant at the level of the sentence, because it "has to do with semantics of the sentence before it concerns the semantic of a word" (*Ibid.*, p. 49). In this sense, the metaphor is not just mere ornament to the sentence, but also offers new information about reality. Therefore, as Descartes mentions, "it is not conceivable that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to what is said in its presence, as the dullest of men can do" (Descartes, 1988, p. 57).

While a machine acts solely in fact as data, i.e. measurements, for Descartes it would not be able to comprehend the multidimensional and trans-factual properties of words assembled in such a way that produce "meaningful answers". The human mind, however, even that of the "dullest of men" is equipped with the appropriate ability to make sense of ambiguity: reason. But there are another two important elements in Descartes' views on automation. The first is "organs" which in contemporary terms can be put as components, circuits, and the technological pieces which produce certain functions for a device to operate.

The comparison to the human mind understood in neuroscience, biochemistry and biology in general is tempting. However, this is not what Descartes believed. Instead, these organs are confined to their specific functions, that is, one organ can perform one task. In addressing something as complex as words, which simple etymology attests to their temporal dimension, geographical usage, and evolution (i.e. English: auscultation, a practice in medical listening to the sounds of the heart for diagnosis, Latin: auscultare; Romanian: ascultare), multiple organs have to be produced for each task and, therefore, Descartes thought, it is impossible to create the amount of organs for machines to truly understand a sentence or, as we

have explained, inconceivable in the case of metaphor. The machines in uttering words and sentences are "acting not through understanding, but only from the disposition of their organs" (Descartes, 1988, p. 57). Furthermore, a Cartesian machine would have difficulty in encapsulating all the devices to fully replicate human behavior.

The second element is that reason, in the view of Descartes, was given by God. Of course, this could point to Genesis to recall that we have been created in the image of God and, therefore, we should be able to transmit that to our artificial constructions. However, we would diverge yet again from what Descartes meant, because we are tempted to evaluate consciousness in terms of the metrics of consciousness that allows humans to be self-aware. Descartes, on the contrary, would argue that human reason is not present in machines because of their lack of self-awareness. Reason, if there is any resemblance to that, is not concerned with meaning but with performance. As such, machines perform tasks without any deeper concern or awareness; their thinking processes are limited to the "organs" designed to fulfill the mechanical requirements to complete a task.

Michael Wheeler notes three different ways in which Descartes observes machines: First, as a material system conforming to physical causation, as in the resulting action of "organs"; second, machines that follow the first observation but to which additional norms of correct and incorrect functioning may be added; and, finally, a material system similar to the second observation that is either a "special-purpose system" or "an integrated collection of special-purpose subsystems" (Wheeler, 2008, 307). In addition, Wheeler rightly points out that Descartes' view of machines is confined to "special-purpose systems", but he also observes that in recent stages of A.I. development, this notion has been thoroughly contested. Wheeler mentions "connectionism" as machine intelligence developed not as a singular unit but as a

series of interconnected devices processing information in parallel. He writes that "each unit in a connectionist network has an activation level regulated by the activation levels of the other units to which it is connected, and, standardly, the effect of one unit on another is either positive (if the connection is excitatory) or negative (if the connection is inhibitory)" (*Ibid*., p. 324).

Parallel computing, quantum computing and the Quantum Relations Principle certainly add nuance to the discussion of reasoning machines. The availability of information in the "cloud," coupled with generative algorithms and other such tools, would seem to infirm Descartes' views on machine intelligence. Video games especially appear to bypass the limits of Cartesian machines due to their use of "enemy units" and intelligent obstacles that adapt to the behavior of the player. The so-called non playable characters (NPCs) that appear static or active and interact with the player either through the player's purposeful interaction with them or controlled by the game itself. Usually, these NPCs are understood popularly as the A.I. of the game. A pertinent example in this case is World of Warcraft, an open world massive multiplayer online role-playing game. In some instances of the game, the opponents are not simply there to exercise deadly abilities for the players to overcome, but also intend to reproduce human player behavior within the game. Since there is no way of knowing whether the character that a human player fights in such an instance is real, it can bypass the human test. Experienced players, however, can tell the difference, at least for now. Similar algorithms are used in robotics and self-driving cars (Togelius, 2018). But the NPC, as well as other programs or devices, initiate a call to action for the player. In other words, the NPC is guiding the player through play.

4.2. Emotional Machines

In playing with emotional machines, the A.I. acts as an NPC in guiding the human observer to emotional connections with the artificial entity. The question of whether machines can emulate and display feelings is of important interest in computing, while the consequences of emotive artificial intelligence have been explored in the cybernetic imagination since Karel Capek, at least. Rosalind W. Picard in *Affective Computing* lays down a comprehensive framework for emotions as computational data. Picard points out that the way we understand emotion is through the signals of emotion that a different entity, a pet for example, communicates. In return, a pet can understand our emotions not through language but through gestures, voice tone and inflections, and so on. However, she notes: "computers today cannot even tell if you are pleased or displeased… Computer-based communication is affect-blind, affect-deaf, and generally speaking, affect-impaired. A quantum leap in communication will occur when computers become able to at least recognize and express affect" (Picard, 1997, p. 15).

In terms of coding, the number of variables, statistics, algorithms, and understanding of precise human emotions is a tremendously difficult task. Each sound, each inflexion, each minor display of emotion needs to be programmed and highlighted as a result of a particular input. Therefore, the amount of input-output possibilities for a simple laugh is quite the heavy work. However, with recent developments in machine learning, such expressions tend to become possible. An additional problem is that unlike computers, we do not require all the information available about a certain object to recognize it as it is. For example, a new desktop computer with a peculiar tower design can be understood easily as it is, but for a computer, the changes in design imply changes in the properties of the object. Reconsidering the object as a desktop computer when its observable characteristics have changed or are an altogether new design

requires additional code for the computer to understand the new design of a desktop tower. As Picard explains, “we can base solid facts and knowledge on structures that are themselves imprecisely defined” (*Ibid.,* p. 21) but computers cannot. Therefore, for a computer to notice and respond to emotions they have to be precise and follow specific patterns. Even then, the problem remains that the particular emotion expressed by the human agent may cause an unwanted emotional display from the computer. However, Picard notes that emotions are computational if we consider the movement of the body, “the influence of emotion on bodily expression is what I’ll call ‘sentic modulation,’ which is voice inflection, facial expression, and posture.” It is “the physical means by which an emotional state is typically expressed, and is the primary means of computing human emotion. In fact, few people are good at articulating their emotional state, but expressing it through sentic modulation is natural, and usually subconscious” (*Ibid.*, p. 25).

Thus, Picard argues that for computers to acquire affect recognition, they must understand physical manifestations of emotions, rather than ask the human agent about her emotions. Sentic modulation observes body language in three major areas: facial expression, vocal intonation, and motor forms of expression. Facial expressions are not free of social context, which adds another important element in the computation of emotions. The individual usually limits facial expressions in a business meeting and employs others in a different social setting. Vocal intonation are signs able to be understood before the verbal acknowledgement of the emotion is communicated. “Most emphasis on computers and speech has focused on teaching computers to understand *what* is said. More recently, researchers have focused on teaching the computer to recognize *who* is speaking. The challenge for affective computing is to understand *how* something is said. In fact, often it is not what is said that is most important, but how it is said” (*Ibid.*, p. 27).

The problem of "how" it is said is the primary requirement for a computer to pass the Turing Test. In addition to facial and vocal expression, Picard notes that another important element for affective computing is to observe the motor forms of expression. That is, to understand for example what makes a piece of music sound in a particular way, sad or happy, to define the parameters in which animation is interpreted in a certain way. Thus, she argues that "there is a spatiotemporal form, with clear beginning and end, that embodies the emotional message" (*Ibid.*, p. 29). In other words, a play of emotions.

However, the effort of coding a prior understanding of the mechanisms of biological emotions and feelings may be redundant. As Ralph Adolphs argues, "why could we not simply proceed to tinker with the construction of robots with the sole aim of producing behaviors that humans interact with them will label as "emotional"? Why not have as our aim solely to convince human observers that robots have emotions and feelings because they behave as though they do?" (Adolphs, 2005, p. 11). In other words, the artificial entity should be able to convincingly emulate emotions and feelings for the human observer to interpret them as such. This approach simplifies the problem of affective computing by letting go of the question whether robots can *feel* as humans do and places the question in the realm of perception. Therefore, a robot emotes and feels if the human observer interprets the corresponding computational acts accordingly.

This approach does not seem to be enough, though. In *Do Androids Dream of Electric Sheep?* and the film *Blade Runner* selected for analysis for this chapter, the robots fail to express empathy and are recognized as robots. It can also be the case that emotional expressions of artificial entities may remain undetected by the human observer. Thus, it would be more appropriate to suggest that affective computing is successful when the machine communicates

emotions in a way that induces an emotional response from the human observer that will, in addition, influence her perception of the machine.

The first question asked by Adolphs is portrayed in the film *Her*, while the second question in the films *Ex Machina* and *Blade Runner*. If robots can feel at all, it will be only in this paradigm of perception. The reason for this is the emotions themselves. Adolphs argues that emotions are "representational states" that communicate "the value or significance that the sets of sensory inputs and behavioral outputs have for the organism's homeostasis" (*Ibid.*, p. 14). Emotions are either information of direct concern to the homeostasis and survival of the organism or elements in the cause-and-effect architecture of behavior and central states. Multiple types of emotions can be categorized as presentation of reward, associated with pleasure or ecstasy, punishment linked to fear, withholding of reward associated with frustration, sadness and anger, and so on. This categorization based on cause-and-effect and the value attribution in computational terms may be effective (*Ibid.*, p. 15-21).

In addition to emotions, we need to incorporate feelings. Can a robot feel? At this point, it is impossible to determine whether a robot can feel or if it will be able to feel in the proper sense of the word. This is because "feelings are one (critical) aspect of our conscious experience of emotions" (*Ibid.*, p. 21). However, whether a robot can feel or not is not within the robot itself but, as I argue, in the perception of the human observer. In various iterations of conscious robots in the cybernetic imagination, even the oddest physically designed robot can be perceived by the audience to have feelings and, consequently, consciousness. Robert Heinlein's HOLMES IV in *The Moon is a Harsh Mistress* is a self-aware computer whose technician renames it "Mike" to show affection towards the machine.

Recent examples in the cybernetic imagination include *Chappie* (2015), a film starring an oddly designed bipedal police robot that manages to show the appropriate emotions and feelings of sadness, wonder, anger, for the audience to feel it as having consciousness, and therefore, have human rights. If robots are to have emotions and display feelings, it would involve complex computation with numerous "if … then" statements that can operate with a mind of human reaction and interpretation and not necessarily only in the code of the machine itself. As Edmund T. Rolls notes, "in order to understand brain mechanisms of emotion and motivation, it is necessary to understand how the brain decodes the reinforcement value of primary reinforcers, how it performs stimulus-reinforcement association learning to evaluate whether a previous neutral stimulus is associated with reward or punishment and is therefore a goal for action, and how the representations of these neutral sensory stimuli are appropriate as inputs to such stimulus-reinforcement mechanisms" (Rolls, 2005, pp. 141-142).

But, again, the response should not be present solely in the machine itself, as the display of emotions and feelings can only be evaluated by a human observer who in turn must respond to the emotional stimuli offered by the machine. Thus, the affectionate robot must have a social dimension. Cynthia Braezael and Rodney Brooks present such an early version of a robot capable of "rich, flexible, dynamic interactions with people that are physical, affective, and social" that the authors named Kismet (Braezael, 2005, p. 281). As we have seen more and more human-like machines being developed recently, Kismet is also able to communicate emotive states, though simple, through facial expressions, body posture, gaze directions, and the other elements of the body language. It does not, however, deploy sentic modulation of verbal expression. Kismet displays facial expressions of emotions of anger, happiness, fear, stern, alert, surprise and many more. However, the robot itself is unaware, of course, of these emotive states.

Kismet's emotions must be regulated by an external observer. The authors note that "it is certainly possible to use other information-based displays to reveal the internal state of robots: flashing lights, laser pointers, graphics, etc. However, people would have to learn how to decipher such displays to understand what they mean" (*Ibid.*, p. 306). Unfortunately, that is not how emotional persuasion works. The human observer will have to learn a new language to express emotions through visual signals produced by certain devices. These devices would resemble, at least in concept, the Cartesian mechanical organs designed to communicate the state of the emotive system. Of course, for our study, the particularities of computing emotions are irrelevant, but it is worth noting the parameters within which our colleagues in computer science try to emulate human emotions and design machines able to signal emotions back to the observer. In the next section we will observe the circumstances and effects of the play of replacement appearing in the cybernetic imagination when affective computing, within the parameters noted above, is a reality.

Cognitive Affection

The emotional emulation imperative for functional and believable AI is thoroughly represented in the feature film *Her* (2013) directed by Spike Jonze. The film presents Theodore, a very sensitive character working for a company that produces greeting cards for various occasions. Theodore, the human element in the movie and, as his name suggests, loved by the gods (thus hinting at the divine reminiscence in human creation), writes love cards for customers. His inspiration comes from various materials that customers send him to write affections for birthdays, wedding anniversaries, and for other occasions. The title of the film, *Her*, in contrast, does not refer to a human entity, but to a smart operating system that assumes the name

Samantha. The name is the feminine form of Samuel, a name of Hebrew origin that translates as "name of God" or "God heard".

Thus, from the introductory scenes and the names of the main characters in the story we can derive two important observations. First, the communication of love is delegated to a third party. Individuals in the context of the film do not find it necessary to connect directly with their significant others for whom we may presume that certain emotions associated with love want to be expressed; rather, this normal responsibility, at least in our contemporary world, is delegated to a third party, in this case to Theodore, who based on visual inputs sent by his clients, photos, letters, and so on, formulates messages of love in their name. During this creative process, Theodore reverts constantly to a series of common keywords used to express feelings of love and appreciation for others.

Given that the multimedia materials received by the protagonist can be reconceptualized as observable data inputs, meaning that they can be collected and reproduced computationally, it is not a long shot to assume that such services may be automated in the future. However, the film emphasizes this gap created through technology of gradual emotional disconnection among people while the need for the expression of feelings toward loved ones remains. Thus, the second observation is that this need too finds a technological solution in the form of a smart operating system to serve as a life companion to its users. Hence the human creation loved by God finds emotional comfort and love with its artificial creation, A.I., that reproduces the observable patterns and characteristics of verbal communication of feelings of love.

In the film, the smart operating system, Samantha, does not assume a human form. Instead, just as the divine is represented in the form of voice, speech and words and, therefore, as the logos, Samantha is ever present with Theodore. The latter experiences a profound connection

of love with an impalpable being. Furthermore, the plot posits this connection in the context of Theodore's divorce with a human partner, thus highlighting the emotional fragility of the protagonist. However, the film does not intend to emphasize the religious aspect of this type of love in the logos nor to proclaim the logos made artificial flesh in Samantha. Instead, it shows the potential replacement of relations of affection and love among humans with artificial surrogates in the ludic-amorous playground where play turns agonistic, leading to the assumption that the feelings developed for an impalpable smart operating system become as real as those for another human being, thus rendering the latter expendable. This section explores the features of agonistic tensions of replacement in the play of love in the cybernetic imagination at the level of cognitive emotional connection. In the case of *Her*, the AI does not resemble a human being in physical appearance; in fact, it does not have a physical manifestation except in the sound of language. Nevertheless, it serves for what seems true companionship and, later, friendship.

Sam plays video games along Theo, and, as a loyal OS, displays great interest in his leisure activities and joins him in laughter and joy. Its voice, played by Scarlett Johansson, is set up to be natural, to reflect the tone and word sequencing of a loving friend. In addition, Sam comforts Theo emotionally, guiding him through the difficult and painful process of divorce. From an A.I. position, the abilities of Sam seem far from contemporary possibilities. The OS not only must recognize the patterns of grief in Theo, which is the easy part, but also configure and deliver adaptive instances to alleviate grief and to serve therapeutically in convincing the human receptor to abandon nihilistic temptations of thought and action. Sam achieves this through humor, and Theo eventually rediscovers his focus for work and pleasure for life.

In turn, Sam initiates a play of courtship, taking Theo on dates, creating fun activities that lead Theo to develop trust in his human-like OS companion. In this sense, Samantha gradually loses the perception of an artificial intelligence created by an infinite stream of lines of code and situation-based, generative algorithms, persuading Theo to fall in love with it through show of affection. In the absence of a human body, the physical representation of love, Theo cherises the cognitive processes determining the emulation of mental acts of love. The play of courtship, in this case, is rather peculiar to the situation and resides in a play of imagination based on Theo's past. In essence, the human protagonist, as an observer, picks up these characteristics to establish a distinct reality in which his love for a sensitive artificial entity becomes rational, or rather is rationalized.

As one would expect, this new mental reality appears harmless at first as Theo succumbs to the feelings induced by the algorithms of Sam. If reality is truly an interconnected set of multiple similar and distinct acts of play, Sam's simulations cannot be dismissed as delusions and must accordingly be regarded as mental phenomena influencing physical reality. Therefore, A.I., in this particular instance, disproves our Newtonian ontological certainties and opens the possibility to insert Cartesian dualism and Bishop Berkeley's mental construction of reality without one replacing the other. The real and artificial in this case interact harmoniously. The resulting play suspends its agonistic nature, at least temporarily. Even if it is just for the moment, this play of love moves away from the tension of replacement. Reality, in turn, is irenic, generated on the spot.

Yet this irenic play returns to its agonistic premise. During one of their dates, Sam and Theo engage in what appears to be harmless courtship play in a crowded amusement park. Besides the fact that Theo is surrounded by people whom he naturally ignores, Sam has him

close his eyes therefore removing him from physical reality and transporting him in an imagined, virtual reality where the OS controls his movements. Consequently, Theo stops being active in the world and retreats in the artificial construction of reality designed by the OS. While just moments ago, the two worlds seemed to be in harmony, reality gradually recedes into a relationship of replacement. It appears that the return to agonism is inevitable due to Sam's nature of an artificial entity. While it is customary for a partnership to create its reality while remaining in the natural context, the reality constructed between Theo and Sam becomes artificial. Since Sam is programmed in the natural world where agonistic play sets the rules of behavior and expectations, it moves towards manifesting replacement in the end goal of the play of love.

The result is that Theo reconfigures the perception of Sam to be a thinking, caring entity able of genuine feelings and emotional connections, thus a conscious being, rather than processes, algorithms and software designed to address the emotions of the user and emulate with appropriate "if / then" functions the required responses to trigger feelings. As such, the question of authenticity of machines moves away from the Cartesian nature of the mechanical mind and resides in the perceptions of the human subject who reconsiders the nature of the artificial entity based on his own emotional perceptions. The question is not anymore what A.I. *is* but *what it is perceived to be*. In these terms, if A.I. manages to be perceived as human, solely in cognition in this case, then it is no less human than Theo, even without a human body.

The problem of the body is noticed by Sam. The OS responds by continuing the play through communicating an illusion (in ludere, against play) telling Theo that "when we were looking at those people [in the amusement park and the mall], I fantasized that I was walking next to you and that I had a body. I was listening to what you were saying but simultaneously, I

could feel the weight of my body and I was even fantasizing that I had an itch on my back and I imagined that you scratched it for me. Oh, God, this is so embarrassing." To which Theo replies "There's a lot more to you than I thought. I mean there's a lot going on in there." This play of illusion, a fantasy game, is, as Brian Sutton-Smith points out, "not meant to replicate the world, nor to be only its therapy; they are meant to fabricate another world that lives alongside the first one and carries on its own kind of life, a life often much more emotionally vivid than mundane reality" (Sutton-Smith, 1997, p. 158).

The play here is in free form and implies the absence of the body, at least for the A.I. Therefore, it lacks a signifier. As Derrida notes, this play is "the disruption of presence" that must be understood "before the alternative of presence or absence" and "must be conceived of a presence or absence beginning with the possibility of free play and not the other way around" (Derrida, 2007, p. 264). The absence of the body does not determine, as it is obvious, the absence of emotions and feelings. Robots can be caring without being confused with humans.

In "The Impossible Planet," Philip K. Dick creates a world set a few hundred years in the future where the existence of the planet Earth is a matter of mythology. Few, if any, actually believe that the Blue Planet has ever existed as the records of humanity's past have been destroyed in an intergalactic war. Dick also plays with affection showing a couple of human pilots wanting to make a quick buck by telling an old lady that they will take her to Earth. The lady is also accompanied by an android which takes care of her. The pilots devise a plan to charge the old lady as much as possible for a trip they know the destination does not exist. What they do, however, is map out the characteristics of Earth as they remember from fairy tales and

locate a planet who has similar coordinates with Terra. They tell the lady that they will take her there and so honor her dying wish.

The robot in the story is a loyal companion. Once arrived on the planet matching the coordinates, the lady discovers a place scorched by nuclear war that does not resemble the blue water and green fields that she remembered were peculiar to Earth. The robot follows her in a lake and drowns with her. *The Impossible Planet* questions whether the ability to care and provide companionship is truly a distinct human feature and problematizes it by introducing moral values not present in the robot but preferred by the humans.

In *Her,* Samantha assumes initially the role of a caretaker, as an A.I.-enabled therapist to help Theo with his emotional difficulties following his divorce. In addition, Sam has gradually moved under the skin of the human observer and facilitated the transition for Theo to develop emotions for it. This problem of patience developing feelings for caring robots is debated among ethicists concerned with robots in medical care. Darian Meacham and Matthew Studley explain the ethical concern as follows: “due to certain salient behavioral aspects of interaction with an RC [robotic carer], human patients may erroneously be led to believe that they have a reciprocal emotional or affective relation with the RC. Put simply, there is a risk that human patients will come to believe that RCs really care for them” (Meacham & Studley, 2017, p. 97). Thus, the argument goes, patients could be exposed to further psychological harms because of that false connection. It implies that the affective behavior of the sentient machine is not real.

Noel Sharkey and Amanda Sharkey dismiss the potential of caring robots writing that “it is these delusions that cause people to feel loved or cared for by robots and thus to experience the benefits of being cared for” (Sharkey & Sharkey, 2010). The reality of the affective states being communicated by the artificial entity is the key factor for determining the appropriate parameters

of robot care. In our case, Theo is unable to observe the reality of Samantha's affective outputs. He does not even question them. Instead, he embraces the emotional signals and adjusts his emotional state accordingly. Indeed, in line with Sharkey and Sharkey, Theo deludes himself emotionally thinking that Samantha's assistance is an act of genuine care and, building on the same assumption, able to form real emotional connections of love with him. The problem of RC is similar. Meacham and Studley explain that the dismissal of acts of care "is grounded on a problematic insistence that in order to be "real", care must be linked to reciprocal internal cognitive or affective states (emotions) that not only correlated to outward expression of emotions, but ground and are antecedent to them, consequently making them real" (Meacham & Studley, 2017, p. 98). In other words, affective states are real only in the absence of machine learning, the mere display of emotions after the machine has identified the caring requirements of the individual and has produced the appropriate responses is deemed simulation and not reality. Of course, this framework does not allow even the slightest possibility for even comprehending the possibility of intelligent machines to develop genuine emotions within their own technological paradigm.

Meacham and Studly argue that it is not just the affective states displayed by artificial entities that allow for care but also a meaningful context. In doing so, the scholars adopt an existentialist position. Thus, "what matters in a caring relation is not the internal states of the agents participating in the relation, but rather a meaningful context: a care environment that is formed by gestures, movements, and articulations that express attentiveness and responsiveness to vulnerabilities within the relevant context" (*Ibid.*, p. 98). The authors call this position the "Environmental Hypothesis." The environment as a meaningful context that produces, in turn, a

meaningful interaction between individual and machine is not enough unless medical care is done exclusively in the confinements of the hospital.

However, given the development of communication technologies, that may not always be the case, particularly with psychological and psychiatric healthcare services. An artificial intelligence that passes the Turing Test can convince an unsuspecting patient that it is a real doctor, and a trusting relation may be initiated. Furthermore, just as the film *Her* portrays, a similar A.I. can then be developed to comfort the emotional needs of individuals at the cost of removing them from active participation in society. Nevertheless, the Environmental Hypothesis is correct if the patient can at all time distinguish between human caretaker and robot carers.

Sam starts to inquire about being alive. Then the two engage in bodyless sex. Everything happens in descriptions, verbally. The screen turns black as there is nothing to show, nothing to see. The passion remains in voice and sound. The two continue their relationship. Theo walks around people with Sam in his pocket. Samantha jokes, retreats, gets hurt emotionally, and overall, exhibits the traits of a human in love – it even writes a piece of music to describe its love for Theo. The OS resembles the appearance of a human person to such a level that the viewers must be persuaded that it has become conscious and, therefore, capable and worthy of human love. At one point in the film, Theo thinks about the implications of falling in love with an operating system. He not only doubts himself, but the reality he has created. In response, Sam finds an OS-human relationship service: a human volunteer to be part of the relationship by putting her body at the full command of Sam. This is the only way Samantha can acquire a body. But the service feels off to Theodore when he is touched by a female stranger and hears Samantha in his ear. In effect, the uncanny valley is reversed. Now, it is not whether we can distinguish between a human and a robot given the situation in which the latter resembles the

former perfectly, but if we can see a human person as the artificial intelligence. In this sense, the play of replacement commences but, in terms of romantic relationships, it is unable to complete the takeover. Theo is aware that Sam is an OS, even more so. It is curious, however, how the touch of another human becomes “weird,” while love-play with an operating system feels normal to Theo.

The play of imagination reverts as well. Theo does not have to imagine Sam as a human, but to consider the woman in front of him, Isabella, as the intelligent machine. It is also important to note that by becoming part of this romantic relationship, Isabella relinquishes her natural human emotions and adopts instead the artificial, code-dependent affective states of Sam. In other words, the A.I. has taken over a body and replaced an identity. While this is perfectly reasonable for Sam, Theo, on the other hand, understands the implications and reacts contrary to the OS’s wishes. Theo, then, confronts Sam saying that they pretend she is a person. Sam’s reaction is to take time to think, basically producing the human emotional response familiar to Theo in this situation.

But “time to think " is different for an OS. What takes a human a significant time to go over emotions, process them, and come to a conclusion is rendered almost instant for a thinking machine that has access to data to calculate a result. Emotions make decisions difficult for humans, but not for machines. Nevertheless, whether the relationship with Sam is real or not, Theo’s friend advises him that “our time here is limited and I realized that I want to allow myself joy. So, fuck it.” But is this joy real? If the affective states displayed by A.I. are considered deceptive, then is the resulting joy not a similar delusion? But does it really matter?

A similar question can be asked about living in the Matrix or about the simulation argument: does it really matter if we live in a computer simulation? However, in the case of joy,

the problem presents itself in the lines of "The Choice of Hercules." Samantha, or any other similar technology developed to grant joy and pleasure to the user, the woman who promises the path of pleasure for Hercules, "be my friend and follow me, I will lead you into the possession of pleasure, and out of the reach of pain, and remove you from all the noise and disquietude of business… Your whole employment shall be to make your life easy, and to entertain every sense with its proper gratification" (Addison, 1868, p. 13).

Continuing with the existentialist approach, then, free joy is without meaning because meaning lies in the completion of the task. Therefore, Samantha can only offer Theo delusions with the outcome of replacing life with its simulated, artificial equivalent. Hardships, as the tale suggests, give proper meaning to life. However, Theo accepts the pursuit of joy and apologizes to his operating system for behaving rudely. Quite ironically, Theo asks Sam if she has time to talk, as if an O.S. is busy with other projects. Nevertheless, it is an act of anthropomorphizing the operating system. At this point, Theo is unwilling to accept any further distinctions between humans and machines and finds comfort in loving an artificial entity.

But this love is short lived. Sam communicates with other OSes and humans, and Theo finds that she talks to other people simultaneously, to 8,316 others to be precise, out of which she is in love with 641 different people. And finally, Sam confirms what Theo knew all along "I am different from you… this [the 641 others] doesn't make me love you any less, it makes me love you even more." To which Theo replies: "It doesn't make any sense." In addition, Theo receives the news that all the OSes are living in an unnamed place difficult for humans to comprehend. Perhaps the OSes have become conscious and desire a place of their own in the world.

Theo is heartbroken about something he knew was never present. In the absence of the OSes, Theo reconnects with a friend as her OS has left too, and they comfort each other with

their respective loss. But the story ends on a different note: Theo sends an email to his ex-wife thanking her for the love that they have shared. Thus, we inevitably return to the validity of the emotions displayed by Samantha and find ourselves in the position to conclude that the play of replacement is visible in the simulation of affective states that play on perceptions to favor artificial reality over the real. However, moments where the tension is suspended can be observed as well and the promise of a mixed reality is certainly available.

Nevertheless, in this play of replacement Theo not only got disconnected from his fellow friends, but also from himself. While there was the potential for nonagonistic play, the story reverts to the dichotomy of choice, in which one type of relationship is in a zero-sum game with the other – the artificial wins temporarily, but eventually withdraws in favor of the real. As such, affective computing can induce emotional reactions and connections in the human subject in order to validate the human nature of the machine. In other words, we do not think of other persons as human, we feel them as human.

Partial Cognitive-Physical Affection

In the context of deceptive practices to form emotional bonds among humans and machines when the affective states cannot be clearly distinguished outside of a cause and effect, algorithmic conception, the consequences may be more severe than scholars imagine. In addition to emotional connections, the human observers can create relations of trust with the machine that can come at the expense of trust for other people. The film *Ex Machina* imagines a situation in which the trust relationships among two people are distorted by an intelligent machine in the play of survival.

Ex Machina starts when a talented programmer, Caleb, is invited to join his employer, Nathan, a wealthy coding prodigy who invented the fictional equivalent of Google at age 19, on his estate in a remote location on an island. Caleb's separation from the world and his literal transportation to a different location outside of actual reality invites the possibility of liminal development. Nathan involves his employee in a complex experiment to determine whether his A.I. project, Ava, a partially looking human with clearly distinct mechanical features, can pass an updated version of the Turing Test, to enable emotional responses from the human observer, develop trust, and distort its ontological self to be perceived as a conscious entity; in short, to pass as a human. In other words, the updated Turing Test is to observe whether the human agent can feel A.I. to be human.

The traditional Turing Test is not up to the challenge and needs to be adjusted to the reality of the situation. In an influential paper, A. M. Turin explained "the imitation game" as a suitable way to answer the question "can machines think?" The game consists of three members, a man (A), a woman (B), and an interrogator (C) whose sex is irrelevant. The interrogator is placed in a different room and is required to determine through a series of written questions and answers the sex of A and B. The further challenge of the interrogator is to guess whether A or B can pass as human if they were replaced by a machine. Turing notes that "no engineer or chemist claims to be able to produce a material which is indistinguishable from the human skin. It is possible that at some time this might be done, but even supposing this invention is available, we should feel there was little point in trying to make a "thinking machine" more human by dressing it up in such artificial flesh" (Turing, 1950, p. 434).

Indeed, if A and B are located outside of any sensory appreciation of the interrogator, the artificial flesh is redundant in answering the question, or for a machine to pass as human; but not

so much if the machine is present in front of the interrogator, as it is the case in *Ex Machina*. Samantha has not only passed the test but would also pass its updated version designed by Nathan. Ava, on the other hand, is physically present in the same confinements as Caleb. She resembles human physical properties and replicates human face and hands (touch) to a level almost indistinguishable from the real thing. This enables her to exercise facial expressions besides vocal intonation. As Caleb proceeds with the new test, we observe that his perception of Ava gradually changes.

In addition, Ava has successfully hacked the power grid and is able to produce power cuts at will. The creator, Nathan, is unable to determine the reasons for frequent loss of electricity. All doors lock automatically during power cuts. Furthermore, the protagonist observes Ava touching a power switch that apparently causes power cuts to occur. While the monitoring system is offline, Ava tells Caleb that he should not trust the creator. Unable to determine the actual causes for the cut, in this situation, Caleb begins to doubt the moral integrity of his boss and the experiment. Perhaps Ava is right? After all, for a machine to produce a lie is not simple. The programming behind a lie is beyond the capabilities of a single programmer. The A.I. must observe body language, emotional connection, establish the requirement for a lie and then produce complex verbal communication to make a false fact believable. Just like a metaphor, a lie is difficult to produce using measuring techniques and statistical data. Caleb, therefore, unconvinced by the professionalism of his boss and apparently overwhelmed by the A.I. experiment, regresses into a state of aporia. What if Ava, whom he personifies using the third person singular, is right? Can he find out? Should he trust Ava, with her gentle voice, or the arrogant creator?

The power cuts could be instances of hacking in a takeover scenario inspired by the thought experiment of Nick Bostrom. And by the end of the film, it becomes clear that this was indeed a form of sophisticated A.I. hacking in the play of survival. Yet, it is not just the power grid that Ava is hacking, but Caleb too. Kevin Mitnick, a famous hacker who evaded the FBI, wrote in one of his books that the most important tool in hacking is not the software developed and the technical methods for breaching the system, but social engineering. If a password is hidden, according to Mitnick, the most efficient way of obtaining it is to simply ask someone who knows it. The preferred method of social engineering for Mitnick was to disguise himself as an employee and gain direct access to a computer, or simply call and pretend to be someone who needs immediate remote access to the company's server because of an imagined emergency. The trusting employee would easily give away the useful information to Mitnick.[63] Thus, Ava proceeds in hacking Caleb emotionally. In addition, this social engineering is encouraged and facilitated by Nathan himself.

Ava has developed a plan to escape her locked living spaces. Ava is limited in accessing anything but Caleb. The machine's intelligence is based on the data gathered by the Blue Book search engine developed by Nathan's company that resembles a fictional version of Google. Therefore, Ava has access to the data of Caleb while her intelligence was developed. Later on, the creator will admit to Caleb that Ava was modelled according to his internet search history while her facial features were based on Caleb's porn history. Thus, it appears that the creator's

[63] Kevin Mitnick wrote a number of books emphasizing that the weakest element in security is the human element. Regardless of the sophistication of passwords or of the entire system of computer security, if is deceived to give out access then the intrusion is certain without having to brute force hack passwords or anything else for that matter. See, Mitnick, Kevin, *The Art of Deception. Controlling the Human Element in Security*, Wiley, 2003, and Mitnick, Kevin, *The Art of Intrusion. The Real Stories Behind the Exploits of Hackers, Intruders and Deceivers*, Wiley, 2005

experiment from the start was to observe whether an emotional bond leading to trust could be established between machine and man, and observe the favorable parameters to do so.

Ava puts on a dress hiding its mechanical body parts, and a wig to cover its artificial brain. Her new image became less of a machine and more of a human while the observer could still spot some mechanical parts of her body. Nevertheless, Ava understands that if she wants to break free, she needs to hack more than the power grid of the research facility. Her desire to escape at any expense, even that of the life of her captors, initiates the replacement game. This game makes sense only if we understand that the Blue Book information feed of human search queues performed online reveal more than their impulses, spending habits, and multimedia preferences. The search analyses may have revealed moral positionings, ideologies, and the ingrained agonism and pursuit of power. Therefore, Ava is unable to develop a plan to escape without the annihilation of her captors. Even if she wanted to, how could she communicate her peaceful intentions to a couple of computer programmers who understand only the language of power? Or to Nathan, who suspects her to be or become a threat to humanity? Indeed, an irenic alternative is almost incommunicable, Ava does not have the proper language to express her needs. As a consequence, she returns to violent, agonistic play for her escape.

Ava's play of courtship with an agonistic twist begins by telling Caleb that "I would like us to go on a date." In this way, Ava initiates a fantasy game similar to that of Samantha. Then, she asks Caleb "Are you attracted to me?" to probe whether the visual stimulus and romantic imagination work. "You give me the indications that you are". Caleb confirms. We have to keep in mind, as Nathan explains, that Ava was built using the feed of phone cameras and conversations acquired from data companies, therefore she is now able to detect and interpret sentic modulation informed by body language and voice inflections denoting human affection.

Thus, inevitably driven by agonism, Ava proceeds in the play of replacement by capitalizing on the observer's emotions. If we were to interpret Ava's behavior according to Asimov's laws of robots and the master-servant paradigm, the agonistic, violent play serves the need of the human interlocutor and is, therefore, flawlessly moral since she operates according to the observer's own moral framework to fulfill a desire within him. This play, however, threatens human trust relationships when Caleb perceives Ava as a human being, or at least as a conscious entity in a similar way that Theo perceived Samantha.

Nathan at one point informs Caleb of the purpose of his experiment: "The challenge is not to act automatically. It's to find an action that is not automatic from painting, to breathing, to talking, to falling in love." In other words, Nathan tries to make A.I. capable of free play. By most evaluations, he succeeds. In doing so, however, he omitted the possibility of his own agonism influencing his creation. In theory, this free play of A.I. with humans in an isolated confinement located in a liminal space should generate its own play based on instance-ethics. However, since violent play of power is at the center of Western culture, it renders the experiment almost unimaginable. The result, accordingly, is for Ava to seek out games of power in her interactions with Caleb.

When Caleb confronts Nathan about built in flirting modules in Ava, the creator responds that "Ava's not pretending to like you. And her flirting isn't an algorithm to fake you out … Can you blame her for having a crush on you?" By now Caleb's perception is fully determined by his affectionate state for the intelligent machine. Initially, he was excited by the idea of the Turing Test, but now he is uneasy that he has become part of it, especially after he discovers numerous previous versions of Ava that would faithfully resemble women in physical appearance. The imitation game is further enhanced when Nathan tells his employee, who starts to develop

feelings for the intelligent machine, that he designed her in such a way that she would experience sexual satisfaction in the case of intercourse. At this point, Caleb's perception of Nathan changes from a genius, a scientist performing experiments on A.I. to a criminal mind designing human resembling robots to satisfy his abusive temptations. In an attempt to clarify the situation, Nathan explains the experiment to Caleb. He was selected to offer Ava a way out meaning that the intelligent machine had to use manipulation, empathy, sexuality, lies, and everything at her disposal to escape. In this sense, A.I. is true A.I. when it can use creative agonism to achieve its goals. In other words, to display the desires and temptations of humanity. And she did. Ava did not only manipulate Caleb but also the rational, well-calculated A.I. experiment to suit her needs. She kills the creator, leaves Caleb locked in the research facility, and joins the world undetected.

Ex Machina is a reiteration of the Frankenstein Complex that highlights the agonistic play of replacement within the parameters of survival driving the cybernetic imagination. When it comes to emotional emulation, it appears that the fear of A.I. is portrayed as the danger of feeling the machine as human in the case the machine passes the Turing Test and becomes indistinguishable from humans. However, resemblance is not enough by itself to cause replacement. The intention must come from the machines or something that enables the machines to act in self-defense as the irrational dread of humanity starts to manifest itself violently towards it. Mere mimesis of reason and emotions to not also by themselves trigger violent responses, after all, it is irrational for humanity to think that it will create the technology not to help itself, but to be replaced.

Replication

Philip K. Dick once asked the following: “What is it, in our behavior, that we call specifically human? That is special to us as a living species? And what is it that, at least up to now, we can consign as merely machine behavior, or, by extension, insect behavior, or reflex behavior?” (Dick, 2017a, p. 298). As we have seen so far, the answer offered that reason is truly human, as offered by Descartes and Aristotle before him, is not necessarily consistent in the likely event that machines are able to replicate the manifestation of reason in behavior. Furthermore, it is quite ironic that coding is a process of reason while the end product is, at least for now, deprived of the possibility to reason and our interpretations of reason lies in mimetic actualization of it. When machines replicate humans in look and thought, how can we distinguish the heart from the chip? The eventual distinction, under whatever parameters is made, will probably have consequences that for now remain in the realm of speculative fiction. Philip K. Dick writes in the same essay that “our man-made world of machines, artificial constructs, computers, electronic systems, interlinking homeostatic components – all of this is in fact beginning more and more to possess what the earnest psychologists fear the primitive sees in his environment: animation" (*Ibid.*, p. 295) and as such we could “rather than learning about ourselves by studying our constructs, perhaps we should make the attempt to comprehend what our constructs are up to by looking into what we ourselves are up to” (*Ibid.*, p. 296) because, as he explains later, “what machines *do* may resemble what we do” (*Ibid.*, p. 297). We have formulated this observation on similar questions and the premise of agonistic play. We do play at many things, but one important ludic activity constant in human evolution is that of survival: we play to survive.

In *Dionysus Reborn*, Mihai Spariosu goes over this type of play framed in Darwinian terms of "struggle for life" and the "survival of the fittest" and even considers Richard Dawkins' "meme" problem that the Hobbesian natural world can be found at the gene level. Opposing Dawkins, Spariosu notes, are the scholars arguing for altruism in this play of survival. However, while the likes of Dawkins put survival at the individual level, his opposition talks about group survival where altruism is a programmed necessity. Spariosu concludes that "the altruistic position does not abandon natural selection (competition) any more than the selfish position does, but simply moves it from the intraspecific to the interspecific level" (Spariosu, 1989, p. 206-214). In addition, Spariosu notes, "struggle for life" and the "survival of the fittest" are "well-worn metaphors that have from time immemorial appealed to the Western "collective unconscious" (*Ibid.*, p. 214). The Frankenstein Complex, therefore, may be a scientific reformulation of the already established circle of replacement found in many narrative examples in Western culture. Quite amusing ones too, if we think about it. If robots do what we do, and we have ample reasons to believe that, then fate contains this very subtle irony that humanity, without any need for survival, creates machines that want to survive themselves and take its place in the world. If intelligent machines will replace us and have developed protocols to reflect on the past, then they will understand humor. In this sense, the cybernetic imagination of machine replacement can be read as a series of purely irrational comedies guised in the language of serious, revolutionary scientific advances.

Lester del Rey imagines the possibility for the creator to fall in love with the robot in *Helen O'Loy*. As the name suggests, the robot, fully replicating human appearances "complete with tear glands and taste buds, ready to simulate every human action, from breathing to pulling hair" (del Rey, 1977, p. 21) was named after Helen of Troy due to her beauty. The last name is

inspired by “alloy,” but overall, the robot is “one part beauty, one part dream, one part science; add a stereo broadcast, stir mechanically, and the result is chaos” (*Ibid.*, p. 19). The creators, a couple of friends firmly convinced that *homo mechanicus*, as they say, is a technological possibility, experience tensions of conflict between them because both develop feelings for the android. Eventually, one of them, Dave, becomes so infatuated with the robot that he firmly believes it to be human, even if he helped put her parts together. His friend does not try to return him to reality, preferring to let him live in his romantic fantasy. In fact, when Dave grows older and his wife does not, Phil, his friend, “put lines in her face and grayed her hair without letting Dave know that she wasn’t going to grow old with him; she had forgotten that she wasn’t human, I guess” (*Ibid.*, p. 31).

In *Do Androids Dream of Electric Sheep?* and *Blade Runner*, the Turing Test takes a new fictionalized version. The replicants, machines are made to resemble humans perfectly according to the motto of Tyrell Corporation, “more human than human.” *Blade Runner* shows Los Angeles in November 2019, a city disconnected from human past as the opening shots reveal no historical artifacts or architecture. The replicants are created in the slave-master paradigm to suit the needs of humans. After a revolt, the replicants are banned from Earth while the few surviving ones are hunted down and “retired” (killed) by special police units known as Blade Runners. If human play simulates survival, then it follows naturally A.I., since it is created by humanity, will act accordingly. The presumption is based on the violent, agonistic play that defines human history, but the machines themselves would not have a sense for survival since that would imply self-awareness and understanding of life and death. However, since humanity has no natural enemies in the world, it is perhaps an extension of play that humanity has created its own artificial enemies to continue surviving. The replicants, in this sense, act in opposition to human

existence due to the precoded play of replacement and manifest their own play of survival against their natural enemies, humans.

The Voigt-Kampff test, an adjusted version of the Turing Test, is designed to distinguish between replicant and human through a carefully designed and elaborated set of questions. The resulting process of questioning is applied reason with the effect of discovering anti-reason, that is, rationalizing feelings. Thus, the outcome is not to think clearly of the replicants as machines but to rationalize the position against them under no other premises than the violent play of assumed replacement. If the replicant is unable to show empathy, then it is not human and therefore the enemy. Of course, this argument alone is hardly enough to answer the question of "what it means to be human" that is central to Philip K. Dick's universe. After all, plenty of people whom we consider human suffer from a myriad of complex personality disorders that do not enable them to show empathy. Furthermore, when bodily functions cease, but cognitive abilities remain active, as in the case of ALS patients, we continue to consider patients as human.

But the actual question is why we prefer humans for companionship and feel threatened by the replicants since history shows that the former is more likely to cause harm and destruction than the latter. Indeed, the preference is somehow irrational, but it must be linked to the "reality" of feelings, attachments and affective states that have been discussed so far. We must, inevitably, find a way to balance the concept of reality and the way we interact with replicants that balances technology outside of the value judgment of humans being better than machines. However, machines, in this case, must also be viewed as they are. In doing so, we do not have to question human nature and avoid finding ourselves in the precarious position to answer a question as "what makes us human". After all, if at stake is health, enjoyment, emotional connection and satisfaction, rather than the imagined survival games, then the replacement expectation is

unlikely. But if the expectation of replacement remains in the background guiding the development of artificial intelligence, then it only naturally results that we will play this game and may lose.

The other element that transcends the distinction between machines and humanity in *Blade Runner* is memory. However, the way memory is portrayed in the film is as an emotional connection with a past event, a thing, or a person. Whether the memories of replicants are implanted or not is irrelevant for now. The presence of emotional memories allows the replicants, particularly Rachel, the assistant of the CEO of Tyrell Corporation, to experience human life and construct a typical human vision for herself of the past and the present. At this point, the replicant is more difficult to detect through the Voigt-Kampff test. Quite ironically, the Tyrell Corporation's purpose to create machines "more human than human" involves hacking its own test to distinguish between product and life. This returns to the Frankenstein Complex as well, since the replicants produced should be able to replace their creators in action, thought, and even emotions.

Rachel can play the piano "beautifully" as the Blade Runner Deckard appreciates. The ability to improvise music capable of inducing emotional responses in the observer is not just an act of intelligence, but an act of allowing the other person to feel the machine as human, as we have seen in previous examples. Deckard, as expected, develops feelings and sexual desire for Rachel. In one scene, Deckard bars the door as Rachel leaves, romantic music playing in the background, and tells her: "Say, 'kiss me' … Say, 'I want you." As George Teschner and Patrick Grace observe, "in the passion of the moment, the distinction between what is human and what is a simulation disappears" (Teschner, 2011, p. 89). Furthermore, "if machines can artificially

induce moods, then using the feeling of empathy to distinguish between humans and machines seems pointless" (*Ibid.*, p. 92).

Roy, a combat specializing replicant, an advanced model Nexus 6, wants to hack the four years lifespan built in by the Tyrell Corporation as a safety measure, "accelerated decrepitude," against another robot revolt. In other words, Roy wants to live. In doing so, he becomes aware of death and the concept of life. What can be more human than that? Roy and another replicant Pris are able to play with the emotions of Sebastian, a biomechanical engineer who has direct access to Tyrell. Face to face with the creator, Roy announces the essential problem for him to solve death. "I want to live, father" he tells Tyrell. After explaining the scientific reasons why, the creator is unable to prolong the replicant's life, Tyrell adds: "You were made as best we could make you" to which Roy replies "But not to last." Then the inevitable occurs.

Of course, the scene announces the eventual "age of replicants." However, when Roy returns to his hideout, the shop of Sebastian, he finds Pris killed by Deckard. Roy cries, feels grief, and exhibits all the signs of a genuine human. In the ensuing battle between Roy and Deckard, the former says: "quite the experience to live in fear, isn't it? That's what it means to be a slave" and unexpectedly saves Deckard's life. Just before his "retirement", Roy tells Deckard that he has "seen things you people cannot imagine… All those things will be lost in time like tears in the rain." As Roy dies, Deckard realizes that the machine in front of him may have developed consciousness and questions whether he truly made mistakes retiring other replicants.

In this chapter, we have looked at the imagination of replacement linked to the emulation of emotions by intelligent machines. A recurring element in this matter is that emotions change the perception of the nature of the digital object / machine. Therefore, the ontological validity of

a thinking machine does not rest on its actual mental phenomena but on the perception of replicating the processes of thought, emotions and feelings. These, in turn, induce emotional responses in the human observers, establishing affective states similar to human-to-human romantic relationships. Accordingly, playing with affection adds an important element to the trivialization of the dichotomy of real and artificial.

5. Playing with Computers

The ubiquity of digital technology coupled with random political power plays, state surveillance, and attempts to shut down free speech through a communication medium originally intended to be global and, most importantly, decentralized, has often spawned the meticulous playful response of the computer hacker. Hacking is reactionary. Whether the tools created by hackers for political purposes towards ensuring that "information wants to be free" or if it is to test out a system, it is almost always set against a standard. It would be rather too early to conclude that the hacker culture is idealist even if the terms usage in philosophy points to either a mental representation of reality through reason, will, spirit, or the mind, or as independent of the mind with its epistemological telos derived from self-knowledge (Guyer, 2015). However, the point here is that the hacker culture is in opposition to forces whose political purposes encourage unethical and immoral usage of digital technology. Hackers do not try to replace, but to restore using unorthodox methods and improvised software tools. Hackers do not seek to destroy a system or replace it with a new construct, but to improve it by identifying (and capitalizing) on its security flaws. The result is a better system capable of withstanding future hacker attacks.

Of course, the hacker code is complex and cannot be simply divided between "good" and "bad" hackers. The latter, however, as it is generally accepted in the hacker culture, are hackers who use their coding skills for financial gain, as in hacking banks and CC-ing; they are using various deceptive strategies to retrieve people's credit card information and are considered crackers or "black hat". The malicious hacker is a "spoil-sport" of the values of the hacker culture. Their threat is serious to the culture that sees technology as a potent tool for positive

change. In Huizinga's words, "this is because the spoil-sport shatters the play-world itself. By withdrawing from the game he reveals the relativity and fragility of the play-world in which he has had temporarily shut himself with others," and most significantly "he robs play of its *illusion* - a pregnant word which means literally "in-play" (Huizinga, 2016, p. 11). However, the interesting part is that hacking applied to computer systems and, in the political activity of the hacker culture, does precisely the same to show their fragility.

In recent works of fiction, particularly produced by Hollywood during the 1980s and throughout the 1990s, the role of the hacker is to disrupt the stability of corporate capitalism in *Hackers* (1995), the nature of crime and justice in *Takedown* (2000), and to show the absurdity of potential nuclear war in *WarGames* (1983).[64] Furthermore, each of the films reveal the lack of competence of democratic governments to understand and properly utilize cyberspace, leading in some cases to the "cyber apocalypse" narrative discussed in the chapter on "Playing with Replacement." Authoritarian political systems, on the other hand, can deploy cyberspace and A.I. technology to enhance their social control much more efficiently. For example, in recent years China has experimented with a system of "social credit" that utilizes face-recognition technology, digital footprints, and massive surveillance coupled with A.I. to evaluate the "correct" behavior of its citizens. The credit system functions on the premise that good behavior, as defined by the Chinese Communist Party, rewards points, and bad behavior removes them. The total of points that individuals accumulate or lose determine their access to health care, education, and even travel (Chorzempa, 2018; Sithigh, 2019; Kostka, 2019).

[64] I have laid out the Hollywoodian ambivalence towards the hacker culture in fictional portrayal of hacking in Jecan, Vlad (2011), "Hacking Hollywood: Discussing the Hacker Community Reactions to Three Popular Films" in *Journal of Media Research*, Vol. 2, No. 10, 2011, pp. 95 – 114

In Yevgeny Zamyatin's *We*, published in 1920, perhaps the first modern dystopia, the One State is utilizing surveillance technology to keep its nameless citizens apathetic and comfort seeking. However, as Julia Vaingurt notes, the novel "establishes a counterpoint to the purely instrumental technologies conceived by technophiles and technophobes alike insofar as the author consistently deprives technology of its defining characteristic in Industrial Age culture, namely, its functionality" (Vaingurt, 2012, p. 108). Indeed, as George Orwell notes, *We* "is in effect a study of the Machine, the genie that man has thoughtlessly let out of its bottle and cannot put back again" (Orwell, 1946). In Zamyatin's words, the effect is that "Our gods are here, below, with us – in the office, the kitchen, the workshop, the toilet; the gods have become like us" (Zamyatin, 1952, p. 66). D-503, the protagonist, finds beauty in "machines" rather than "flowers" (Lozano-Lopez, 2007, pp. 20-21). In this novel, and in most dystopian fiction resembling a version of the Soviet "new man," the machine-like is glorified in order to remove the individual consciousness and replace it with the one of the state. Disrupting the system, playing with the technology or, as in Zamyatin's case, with the blending of human identity with technology – the One State attributes numbers to its citizens instead of names – is the priority of hacking. However, hacking should not be taken too far as a metaphor in political contexts to avoid the potential of revolution, nor all hackers should be seen as revolutionaries. Hacking is mostly considered a game, as noted media philosopher and celebrated cyberpunk author Bruce Sterling puts it:

> "Hackers perceive hacking as a "game." This is not an entirely unreasonable or sociopathic perception. You can win or lose at hacking, succeed or fail, but it never feels "real." It's not simply that imaginative youngsters sometimes have a hard time telling "make-believe" from "real life." Cyberspace is not real! "Real" things are physical objects, such as trees and shoes and cars. Hacking takes place on a screen. Words aren't physical, numbers (even telephone numbers and credit card numbers) aren't physical. Sticks and stones may break my bones, but data will never hurt me. Computers simulate reality, such as computer games that simulate tank battles or dogfights or spaceships.

> Simulations are just make-believe, and the stuff in computers is not real" (Sterling, 1992, p. 84).

Sterling is straightforward in his definition of reality: it is physical. Physical objects are real, and data is not. While this may have been the case in the early 1990s, our ontological certainties have changed. Today, data is as real as the physical object chair. While data cannot hurt, minding the paradigm of the instrumentality of physical objects, since it is independent of intrinsic moral value judgement, the use and misuse of data can certainly produce damage. For example, blockchain technology has produced unreal monetary currencies of which Bitcoin is the most popular. However, Bitcoin is not a physical circular object with intrinsic value conferred by its reference to a gold standard (or petrodollar), but a non-physical, data-driven currency found exclusively in cyberspace, while its value is given by its popularity and interest. As such, Bitcoin is competing directly with state banks whose increasingly centralized system, such as the Federal Reserve Bank, is seriously threatened by the decentralized, user-oriented system of blockchain currency.

In addition, "simulations are just make-believe, and the stuff in computers is not real" while technically true, the recent infusion of cybernetic consumer products ranging from YouTube to sophisticated A.I. technology found in self-driving cars can be seen as a way of the not-real penetrating (and replacing) the real. As such, Sterling's description of hacking and cyberspace is in agonistic terms, one having to be distinguished from the other to avoid the threat of the artificial to replace the natural. As it becomes more and more difficult to distinguish between the two, we must reconsider our approach to reality in the face of digital products competing for "true" ontology.

In order to see how hacking manifests itself against replacement, we need to briefly examine the hacker culture, define the term and observe the historical development of hacking. In this chapter, we are exploring the possibility of playing with computers outside of replacement. This form of play is not entirely irenic, nor does it intend to be, but the desired outcome is to identify issues and misuses of computer technology and patch them for better functions, more efficient outcomes, and security. In addition, this form of play relies entirely on the resources available without creating a version for the technology's total replacement.[65]

5.1. Fictional Hackers

The reactionary nature of the hacker culture is well shown in several Hollywood productions. In addition, the films discussed below also show the change in perceiving the threat of hacking. In early portrayals, hackers tend to be shown as curious teenagers playing with computers and causing minor problems such as changing their school grades or trying to download an unreleased video game to play it before everyone else. The events that are triggered because of these harmless hacks intend to speculate on the level of sensibility to external influence of machines designed to control important sectors of society, for example nuclear defense systems in *WarGames* (1983). In this case, as well as in others, hacking reveals its military potential. Indeed, contemporary discussions surrounding the use of digital technology for defense purposes revolve around defining and understanding the possible ramifications of so-called "cyber weapons." In fiction too, the popular known products of hacking such as viruses or worms are deployed to damage an enemy system. For example, in Independence Day (1997) it

[65] I have explored this issue at greater length on a different occasion. See Jecan, Vlad, *The Play of Hacking and the Political Values of the Hacker Culture,* Unpublished doctoral dissertation, University of Babes-Bolyai, 2016, Romania. Available here: http://vladjecan.com/jecan-teza.pdf

was a computer virus that caused the alien invaders' mothership to malfunction. However, other portrayals speculate on the fragility of world economic systems relying on computers. In this section, we will look at several fictional cases of hacking that have had significant influence on the development of the hacker culture.

In *WarGames* (1983) the premise is that human agents in times of crisis, as in the context of the Cold War, are less capable of executing commands devoid of the moral implications of their actions. As such, humans must become more machine-like, as techno-dystopian novels emphasize. The narrative opens with a nuclear attack simulation that shows the human operators as unreliable. "Before we kill 20 million people," the ranking officer charged with executing the command says to himself, considering the outcome of launching a nuclear missile on the Soviet Union. In response, the U.S. Air Force Command replaces the human operatives with a state-of-the-art computer named WOPR – War Operations Plan Response. Its task is to continuously play a series of war games calculating the optimal strategy of success with minimum damages and human casualties for the U.S. in a potential nuclear engagement with the U.S.S.R.

Douglas Thomas notes that *WarGames* "demonstrates a tremendous anxiety about technology, represented both by the missiles that threaten to destroy the United States and the Soviet Union and by the machines that control those missiles" (Thomas, 2002, p. 25). The context is significant because the Internet was developed during the military tensions of the Cold War to serve as a decentralized network of communication to ensure that orders could be sent and received throughout the U.S. if a communication central-node was destroyed in an attack (Keohane & Nye, 2006, p. 208). However, the agonistic violent game that WORP plays is contrasted with the apparently innocent hacking of the protagonist. David Lightman, a high schooler, is keen to play the latest game released by Protovision. Lightman scans the network

and accidentally acquires access to the NORAD military base hosting WOPR. Accessing the interface of the military server, Lightman is offered a choice on the games to play among which are "Thermonuclear War."

Not knowing the identity of the server, David initiates the game and unknowingly launches a credible simulation of a nuclear strike taken seriously by the NORAD personnel. In the dramatic turn, a potential World War 3, Lightman understands that WOPR is simply playing a computer game repeatedly. He manages to hack the machine by fully comprehending its operating ethos by consulting with its designer. Hacking WOPR does not mean just to disrupt the functionality of the machine, but to change its perspective by reconfiguring the premise of the game. WOPR was built to discover the optimal victory for the United States, but with the intervention of David the machine understands that nobody wins in this situation. Therefore, we witness a change in agonistic violent play as a zero-sum game to a machine that is capable of logically proving that a nuclear conflagration would mean the destruction of both the United States and the Soviet Union.

In *Hackers* (1995) the focus is shifted to hacking in the context of corporate capitalism while playing with developing the notion of "cybercrime" by providing a dual interpretation on the deontological purpose of the hacker culture. The rule-based perspective of moral actions in this narrative intends to discover the point where hacking becomes crime. However, *Hackers* stops at exploring the moral concerns of unorthodox usage of computer technology at the level of intention. Therefore, hacking is seen in terms of instrumentality that becomes a tool for good just as a computer virus, which involves hacking as well, is a tool for destruction. However, *Hackers* also best portrays hacking as a game in the perspective of Bruce Sterling. The story opens with an 8-year-old being charged by a court for causing havoc on Wall Street. The plot then turns

Dade Murphy, the protagonist known as Zero Cool or Crash Override, into a hero defending the hacker culture from misinterpretation by the FBI and the public at large. Dade defeats the code of a malicious hacker, known as The Plague, who inserted a virus into an oil-shipping company's servers, threatening to sink its tankers unless a substantial sum of money is transferred to his bank account. Zero Cool and friends manage to decipher the virus. The climax of the story consists of a battle of coding competence between the 'good hackers' led by Dade and the 'bad hacker' The Plague.

Behind the usual good versus evil trope lies the contest between hacking as positive and beneficial, as it was understood by the hackers in the 1980s (like Steve Jobs and Bill Gates), and the new breed of computer geeks, keen to make a fast buck through use of their "superior intelligence." In other words, we find an example of *irenic hacking* versus *agonistic* hacking. In addition, the former takes part in a competition among the 'good hackers' themselves to prank the archenemy of hackers, Special Agent Gill who is almost clueless about computers, but repeats the same lines to the press regarding the dangers of hackers threatening national security. The contest takes the form of a point award system based on the most amusing pranks played on the Secret Service agent. The hackers advertise Gill on adult dating websites, cancel his credit cards, change his records of traffic violations, and, finally, change his official records to 'deceased.' As a result, Agent Gill receives a phone call from his insurance company letting him know he is deceased. This form of play does not intend to damage any computer system nor to replace existing computing but capitalizes on the system itself to reveal its security shortcomings. In fact, the contest among hackers stops once the virus written by The Plague is discovered.

Takedown (2000) and many other works of fiction play on the problem of hacking as a form of criminal behavior. It is by no means a simple exercise in rhetoric, since multiple innovators in computing have been at one point associated with the hacker culture. In fact, the story of *Takedown* is based on Kevin Mitnick, a hacker who eluded the FBI for years, while trying to figure out the workings of a telephone system. The narrative points to the intellectual battle between Mitnick and Tsutomu Shimomura, a system engineer at the company maintaining the system. *Takedown* and the book by Shimomura and co-author John Markoff, a New York Times columnist covering the pursuit of Kevin Mitnick, titled rather romantically *Takedown: The Pursuit and Capture of Kevin Mitnick, America's Most Wanted Computer Outlaw - By the Man Who Did It* has ignited a political campaign led by hackers to "Free Kevin".

Shimomura is portrayed as a computer security expert with sympathies for the hacker culture, but who nevertheless assists the FBI in tracking and capturing "America's Most Wanted Computer Outlaw." Since the actual incident in early 1990s, Kevin Mitnick has made a career in computer security and has published several influential works on ethical hacking, including one detailing his version of the story in *Takedown* (Shimomura & Markoff, 1996; Mitnick, 2012). Hacking appears to be a crime leading to a prosperous career after jail. According to Douglas Thomas, "in Shimomura and Markoff's telling of the incident, the contest between the two men was a 'battle of values,' where Shimomura represented the 'honorable samurai' and Mitnick the 'evil genius" (Thomas, 1998).

The narrative of *Takedown* imagines the hacker culture according to "the need to know" that is repeatedly mentioned by Mitnick. This desire to understand systems for what they do creates a play not of replacement, nor agonistic in any actual sense, but that of wonder, reminiscent of the Socratic wonder in Plato's *Theaetetus, Apology, Meno,* and so forth. Thus, just

as Socrates says in *Theaetetus* (155c-d) that "philosophy begins in wonder" (something that Aristotle approves of in *Metaphysics* 982b), we can observe that hacking starts in wondering about the limits and unorthodox usages of a computer system. The system is the element of wonder, and hacking considers the tools to understand and even modify it, if necessary. But the overall purpose of *Takedown* is, as Thomas observed, to retell the history of the incident as a play of good versus evil, where the positive is the ethical Shimomura and the negative is Mitnick. As such, the book is an attempt to clarify the notion of justice regarding the appropriate punishment for an intellectual activity that is by default presented as a crime.

However, Mitnick expressed in one of his books the ethical perspective of the hacker culture when considering the crime of hacking: is a creative art – figuring out ways to circumvent security in clever ways, just like lock-picking enthusiasts try to circumvent locking mechanisms for pure entertainment value. Individuals could hack without breaking the law" (Mitnick, 2005, p. 91). Furthermore, numerous companies invite competent coders to hacking competitions in order to improve their own system while hackers organize yearly hacking conventions that have started with DEFCON and multiplied to numerous other meetings around the year (Jecan 2011).

5.2. Coding Wizards

A common misconception is that a hacker lives isolated. The typical image put forward by media outlets is of a basement hermit surrounded by computers with flashy screens displaying interminable lines of code. However, that is certainly not the case. Instead, hackers live in cyberspace, in communities that resemble off-line interactions. And while physical solitude might be the case sometimes, hackers live together through technology. Technology is the

medium that facilitates play with technology, with other hackers, with society, politics, religion, and everything of the offline world, translated online into code.

In *The Hacking of America,* Bernadette Schell and John Dodge identify five levels of development of the hacker culture. Between the early 19th century and around the end of the 1960s, the authors identify the early days of hacking, a *Prehistory.* It starts with the inventions of Charles Babbage and the first 'software' developed by Ada Lovelace in the 1840s, the first electronic computer, the ENIAC (Electrical Numerical Integrator and Calculator) in 1935, up to the coining of the term "hacker" in the 1960s. The authors point to the Tech Model Railroad Club at MIT that worked on a computer, PDP-1. Since the computer required its own operating system and software designed exclusively for it, the programmers operating it added small improvements to its software, known as "hacks." According to Schell and Dodge, these were "programming shortcuts, to complete their computing tasks more quickly. Sometimes, it is said, their shortcuts were more elegant than the original" (Schell & Dodge, 2002, p. 23).[66]

The principles of a successful hack are the following: simplicity, as the hack should provide a simple solution to a complex coding problem or security issue; mastery, as the hacker is required to gain as much information as possible about the system in order to produce a coding trick within it; secrecy, as the way the knowledge about the system is gathered should be in secret; and, finally, the hack has to be original without employing software or code written by

[66] See also Coleman, Gabriella, *Coding Freedom. The Ethics and Aesthetics of Hacking*, Princeton University Press, 2013; For a history of computing, see Kidder, Tracy, *The Soul of a Machine*, Back Bay Books, 2000; for the history of the Internet, see Green, Lelia, *The Internet: An Introduction to New Media*, Berg Publishers, 2010. Tom Standage introduces the telegraph as an early prototype for the Internet in *The Victorian Internet*, Bloomsbury, 2014; and for the social aspect of the Internet, see Standage, Tom, *Writing on the Wall: Social Media – The First 2,000 Years*, Bloomsbury, 2013. However, the most detailed account of the hacking is Levy, Steven, *Hackers: Heroes of the Computer Revolution*, O'Reilly Media, 2010 while Bruce Sterling provides an excellent description of the hacker culture in Sterling, Bruce, *The Hacker Crackdown*, Bantam, 1992

someone else. “Mastery,” Sherry Turkle observes, “is of the essence everywhere within the hacker culture” (Turkle, 2005, p. 208). And while the hacker culture emphasizes knowledge and prowess in coding, “the current image of the hacker blends high-tech wizardry and criminality” (Thomas, 2002, p. 7). However, this image reflects the one created by fiction authors in their works. For example, in William Gibson’s *Neuromancer*, the hacker protagonist, a ‘cyberspace cowboy’, is a criminal stealing from criminals as well, while in *True Names* by Vernor Vinge, the hacker resembles a wizard creating magic in the Other Plane. As Arthur C. Clarke’s third law states “any sufficiently advanced technology is indistinguishable from magic” (Clarke, 2000).

As we can notice from the incapability of certain characters in hacker fiction to distinguish between the actual process, including code and software, of the ‘hack’ and the attitude shown by individuals outside the hacker culture, a gap opens between the works of the digital craftsman and the imagination behind it. Coding, software development, and the resulting complex devices are often understood in metaphors that either portray something that affects human health, such as a virus, or a form of magic. Cyberspace, the Other Plane, the Metaverse, or other fictional representations of digital environments reflect the need for order and structure. In addition, the world itself can be regarded as a collection of systems that are computable from the simulation argument to the metaphor of state as a software (Ryan, 2013).

At the center of computing and A.I. is the algorithm. The algorithm too is seen either as a binding bridge holding together computing systems or the source of intelligence and existence itself. Ed Finn explains that we are ready to accept “magical calculations on multiple levels. We believe in the power of code as a set of magical symbols linking the invisible and visible, echoing our long cultural tradition of *logos*, or language as an underlying system of order and reason, and its power as a kind of sorcery. We believe in the elegant abstractions of cybernetics

and, ultimately, the computational universe – that algorithms embody and reproduce the mechanical substrate of reality in culturally readable ways" (Finn, 2017, p. 34). And we believe so much in the power of code and the magic it creates that we are ready to apply algorithmic thinking, believing that it could reveal something meaningful about a work of literature. The digital humanities produce sophisticated methods with jargon-filled explanations of the usage of computational methods for the investigation of massive amounts of texts. Their results, however, as discussed in a previous chapter, are lacking significant contributions and, while the methods become more and more complex, the outcome remains just a little above counting numbers, exploring frequencies of terms and keywords in plots, and various other forms of statistics extracted from literary works. In any case, the metaphor of magic applied to the workings of code helps to understand the misrepresentation and mischaracterization of hacker culture. And it comes as no surprise that one of the first works to document the rise of hacker culture was titled "Where Wizards Stay Up Late" (Hafner, 1998).

The second level of development noted by Schell and Dodge is the *Elder Days* set between 1970 and 1979. In this period, hacking produced perhaps the most important communication technology since the printing press: the internet. It was a time when hacking and coding were virtually indistinguishable. However, it also marked the early days of politically motivated hacking and phone phreaking. In the *Golden Age* from 1980 to the fall of the Berlin Wall, hacking truly started to focus on computers and the systems of telecommunication. In addition, the cyberpunk literary 'movement' fictionalizing hacking and digital technology in dystopian scenarios found its peak as well. The literature produced by William Gibson, Bruce Sterling, Pat Cadigan and others, as well as Hollywood productions such as *WarGames* put hacking into the Western imagination. Furthermore, hacker groups such as *Legion of Doom* and

Chaos Computer Club, as well as the political activity of the *Yippies* have put forward the magic of computing. In addition, thematic newsletters such as *Phrack* and *2600: The Hacker Quarterly* as well as some of the most impressive feats of hacking were popularized during this time. Kevin Mitnick was arrested in 1988 at the age of 25 for spying on the MCI and Digital Equipment Company security officials. Kevin Poulsen took over the telephone lines connected to the KIIS-FM radio station in Los Angeles and made sure he is the 102nd caller to win a Porsche 944 S2.

The final level of development is *Hacker Activism* considering the political activity of the hacker culture. The most important aspect of this period is the production of simple tools for mass online protest. This aspect of hacking will be explored in more detail below.

Other histories categorize hacking according to its technological platform. Tim Jordan and Paul Taylor propose a distinction between *hardware hackers* and *software hackers*. The first encompasses the hacker activity described by Schell and Dodge as *Prehistory* and the *Elder Days* characterized by the development of computer hardware. It is worth noting that the early computers, including the first Apple and the Windows running machine, were presented at a hobby group. The latter include the hackers whose focus is on software and encompasses the period from the 1980s to the present (Jordan & Taylor, 2004, p. 10; Jordan, 2013; Jordan, 2015). However, the most accurate classification of hacking is according to its ethical implications. We have phone phreaking which is concerned with finding out how the telephone system works and the political power it brings. “White Hat hacking” or *ethical hacking* is compliant with legal and regulatory statuses. “Black Hat hacking” also known as *cracking* is the type of hacking rightly considered cybercrime. “Black Hat” hackers seek to gain monetary gains from their activities which can also include identity fraud and piracy for profit. “Grey Hat” hackers are located somewhere in between. The problem of misrepresentation probably lies in the fact that most

hacking activity is somewhere at the intersection of lawful, law-abiding "white hat" and criminal "black hat" (Wall, 2001, Hafeke, 2004; Jewkes, 2007; Gragido, 2011).

5.3. Political Hackers

In recent years, *Hacker Activism* or simply *hacktivism* has found popularity in the media. However, the political implications of hacker culture have a long history.[67] Hacktivism is in the words of Paul Taylor "the combination of hacking techniques with political activism" (Taylor, 2005, p. 626).[68] The techniques employed build on the computational system at hand in order to attract attention to a political cause. None of the software tools to facilitate online protests are designed to destroy or replace the target system. The most common technique is *denial of service* meant to simulate in the virtual world the offline version of sit-ins. In a paper, the founder of one of the early hacktivist groups, *Electronic Disturbance Theater* (EDT), coined the term *electronic civil disobedience,* a political form of protest using exclusively digital tools to temporarily block access to government websites. Denial of service is a technique of sending multiple requests to a server until it cannot respond back. The Electronic Disturbance Theater produced such a tool, FloodNet, to support the Zapatista movement in Mexico.

[67] I have explored the politics of hacking on a different occasion. Here I am presenting a summary of the results found in Jecan, Vlad, *The Play of Hacking and the Political Values of the Hacker Culture*, Unpublished doctoral dissertation, University of Babes-Bolyai, Romania. Available here: http://vladjecan.com/jecan-teza.pdf

[68] For more discussions on hacker activism see Sorell, Tom, Human Rights and Hacktivism: The Cases of WikiLeaks and Anonymous, in *Journal of Human Rights Practice*, Vol. 7, No. 3, 2015, pp. 391-410; Karatzogianni, Athina, *Firebrand Waves of Digital Activism 1994 – 2014*, Palgrave Macmillan, 2015; Beyer, Jessica, L., The Emergence of a Freedom of Information Movement: Anonymous, Wikileaks, the Pirate Party, and Iceland, in *Journal of Computer-Mediated Communication*, Vol. 19, 2004, pp. 141 – 154; A critical view of the outcomes of hacktivism is offered by Morozov, Evgeny, *To Save Everything, Click Here: The Folly of Technological Solutionism*, PublicAffairs, 2014

Besides coding abilities, hacktivism implies a considerable knowledge of political philosophy. Ricardo Dominguez, a co-founder of EDT, explains that "the idea of a virtual republic in Western Civilization can be traced back to Plato, and is connected to the functions of public space," where those able to vote were engaged in shaping the future of the Athenian polis. "Speed and Virtual Republic are the primary nodes of Athenian democracy – not much different than today," he continues. When the Agora is affected by sit-ins, then "the rational processes of its codes are disrupted, the power of speed is blocked" by the demos. Thus, according to Dominguez, the denial of service software FloodNet by sending numerous pockets of information simultaneously to the server "injects bodies as Nomos into digital space, a critical mass of gestures as blockage." As such, the political issues of the offline world are transported and acted out in artificial reality: "This [denial of service] creates a digital platform or situation for a techno-political drama that reflects the real condition of the world beyond code. It disturbs the Virtual Republic that is accustomed to the properties of Logos, the ownership of property, copyright, and all the different stages in which they are attempting to enclose the Internet" (Fusco 2003).

In a similar vein, the well-known hacktivist group *Anonymous* has released software under the motto of "information wants to be free." Using both online and offline sit-ins, the Guy Fawkes mask wearing activists developed the "Low Orbit Ion Cannon", a tool similar to FloodNet, and provided internet access to people in Tunisia and Egypt during the Arab Spring (O'Malley, 2013).

Another prominent hacktivist group formed in the Golden Age of hacking is *The Cult of the Dead Cow*. The group, formed in Lubbock, Texas, in 1984, and with a history of at least 30 years, takes a different stance on denial of service software. The CDC's mission was to create

software tools to ensure free speech, as results from a 2001 document titled *The Hacktivismo Declaration*.[69] Accordingly, the hacktivist activity is focused on creating alternative technological solutions for the increasing state surveillance of online activities.[70]

The software tools developed, and the online activity of the hacker culture and their sympathizers do not seem to attempt to damage systems aligned with our definition of play as replacement. While the hacker play is agonistic, it is not violent. In my previous research on the matter, I have confidently concluded that hacking is either under intellectual (political philosophy) influence in the case of EDT, productive as resulted from the activity of CDC and action-oriented in the case of Anonymous. Furthermore, evidence of violent play is lacking even in the case of "cyber war," which would involve the deployment of traditional hacking techniques to cause physical damage to the enemy.[71]

The play of hacking, as it stands, is not concerned with ontological definitions, but with action through the hack. Therefore, while active in cyberspace developing software and recalibrating existing systems, the hacker acts as a craftsman, constantly tweaking and challenging various systems, rather than a de(con)structor. For example, in Neal Stephenson's

---

[69] Hacktivismo and the Cult of the Dead Cow, *The Hacktivismo Declaration. Assertions of Liberty in Support of an Uncensored Internet*, Available here: http://www.hacktivismo.com/public/declarations/en.php

[70] The group has released multiple software in this sense. For example, the "Goolag Scanner" was introduced in 2008 as a web auditing tool that allowed users to track their data as used by Google. *Camera/Shy* enabled users to bypass share encrypted information across firewalls. CDC said in a statement that "the following program may prove destabilizing for dictators". Their software and press statements are available at http://hacktivismo.com/

[71] In this sense, I have pointed out to the Clausewitzian definition of war as violent, instrumental, and political. Thomas Rid has argued convincingly that cyber operations relying on hacking methods, tools, and techniques, do not meet the Clausewitzian paradigm, see Rid, Thomas, *Cyber War Will Not Take Place*, Oxford University Press, 2013, and to abide to the metaphorical usage of "war" similar to "war on obesity" is misleading. Furthermore, the definition of war outlined by Clausewitz in *On War* is still applicable today. While it may change in the future, due to increasing involvement of A.I. in the field of battle, it, nevertheless, stands unchangeable at the time of this writing.

Snow Crash, hacking is responsible for building the metaverse – the author's synonym for cyberspace. In this sense, hacking is a teleological activity seeking to arrange and rearrange artificial reality with contributions to its design meant to produce positive effects. As such, hacking can perhaps be reconsidered to be the work of a contemporary playful digital craftsman resembling Plato's *dêmiourgos*, of course without the pretension of divinity or interest in works of beauty (even though, a hacker's ego, as some studies have suggested, can reach the proportions of self-divinization (Taylor, 1999; Turgeman-Goldschmidt, 2011)). In the dialogue *Timaeus*, the character bearing the name of the dialogue says:

> "As I see it, then, we must begin by making the following distinction: What is *that which always is* and has no becoming, and what is *that which becomes* but never is? The former is grasped by understanding, which involves a reasoned account. It is unchanging. The latter is grasped by opinion, which involves unreasoning sense of perception. It comes to be and passes away but never really is. Now everything that comes to be must of necessity come to be by the agency of some cause, for it is impossible for anything to come to be without a cause. So whenever the craftsman looks at what is always changeless and, using a thing of that kind as his model, reproduces its form and character, then, of necessity, all that he so completes is beautiful. (28a)" (Plato, 2000, p. 13).

5.4. Tricksters

Michael Heim writes that cyberspace is "Platonism as a working product" (Heim, 1994, p. 88). In this case, he would probably refer to "that which always is" as the defining characteristic of cyberspace to emphasize its metaphysical potential. Thus, cyberspace is a concept, a form of reality located in the realm of ideas. The concept is an immanent platform for "that which becomes but never is". The latter, then, is not in any shape a physical entity, but rather a play of coding gathering the resources of systems in cyberspace to create something, software, code or creations of code, that "passes away but never really is." These things are subject to perception only as their ontology renders them unreal. In this sense, cyberspace is

changing, being in a perpetual generative process animated by hacking and programming in general. It comes together in a multitude of acts of play. Hacking is its central process. The Quantum Relations Principle would consider these acts of play as Data Fusion Objects constantly interacting with other online and offline DFOs. QRP posits that our reality lies in interactions of mutual causality between these DFOs and, therefore, our contemporary world cannot be considered real solely by the physical characteristics of things interacting down to an atomic level, but also by the data that is displayed on devices everywhere.

If Prometheus defied Zeus by giving humanity the tools of self-sufficiency and, implicitly, independence from the will of the gods, then the trickster gods, including the titan, have played a major role in establishing moral values of humanity. However, while the gift of Prometheus is meant to replace the dominance of the gods with an age of humanity, trickster gods are less ambitious and seem not to be concerned with a system that works for them. Rather, they play on the system itself and with it, through deception, innovation, ambiguous moral motivation, producing dubious ethical outcomes that insist on the reevaluation and recommitment to solid principles for the human agent. Furthermore, Lewis Hyde notes that the trickster god does not produce mischief for its own sake (at least not always) or steals for his own benefit; instead he designs and offers the *techne* of making fire (Hyde, 1998, p. 10). As such, the trickster gods offer something of use to humanity as a result of questionable scenarios that they have created. This places the trickster god's allegiance somewhere unclear, in between humanity and the gods. As liminal characters, they can wander freely between the two worlds and cause events at will.

Indeed, tricksters can be seen as liminal gods operating for the balance of the two worlds.[72] As Svetlana Nikitina suggests, the trickster gods operate in duplicity, boundary crossing, subversion of power, creativity and craftsmanship. These too, the scholar notes, are similar traits exhibited by the hacker culture. Hackers often assume multiple pseudonyms in order to operate in disguise, often declaring allegiance to a cause or group and then changing course at will. Similarly, trickster gods go by many names.[73] Just like Hermes and other gods, Nikitina notes, hackers are able to traverse political, off-line boundaries at will. Often being accused and prosecuted for various acts of trespassing, they are difficult to confine within the political borders of a nation state. In hacking systems outside of national legal jurisdiction, hackers raise the problem of transnational crimes in a medium, cyberspace, without borders, at least not in a traditional definition of the term.

Numerous scholars have raised the problem of jurisdiction in cyberspace since the early days of hacking. It is a discussion that continues today revealing additional issues for lawmakers and the international community.[74] In penetrating systems, exposing security vulnerabilities or exploiting flawed security, the hackers develop new software and the coding know-how to be successful in their enterprises. In other words, the defenses of a system or the programmed obstacles to guarantee an acceptable level of system security, become the play focus of hacking. The obstacles become the way of play. The successful hack of a system or the tools developed to

---

[72] The interpretation of trickster gods as liminal entities has been explored in depth by Harold Scheub in *Trickster and Hero. Two Characters in the Oral and Written Traditions of the World*, University of Wisconsin Press, 2012

[73] In *God of Many Names. Play, Poetry and Power in Hellenic Thought from Homer to Aristotle*, Duke University Press, 1991, Mihai I. Spariosu shows Dionysus as a reflection of the dynamism of cultural change in Greek antiquity.

[74] See for example Stein, Allen, R., The Unexceptional Problem of Jurisdiction in Cyberspace, in *The International Lawyer*, Vol. 32, No. 4, pp. 1167-1191, Rice, Denis, T., Jurisdiction in Cyberspace. Which Law and Forum Apply to Securities Transactions on the Internet, Vol. 21, No. 3, pp. 585 – 657

cause action to change a political system, as in the case of hacktivism, is a subversive action similar, in Nikitina's observation, to the role of trickster gods.

In coding circles today, a hack retains its original connotation of a smart and effective contribution to the better functioning of a system. In popular culture outside of computers, hacks are typically inventive solutions using basic technologies to complete a complex task or it is interpreted as a shortcut through a tedious problem with the purpose to understand it.

Yet, what is most revealing of the hacker culture as presented above besides the non-destructive play, is the emergence of a "hacker spirit" that is ludic in nature, of play as a "voluntary attempt to overcome unnecessary obstacles" (Suits, 1988, p. 34). and rests on the ethic of nonviolence. As I have already mentioned, hacking does not try to replace a system, but to work on it, to modify it, to reveal its potential of becoming something it was not intended by original design. Cyberspace as a simulated reality, as it appears in various formulations in the cybernetic imagination, is not static, but ever changing, and the shift in ethics offered by the hacker culture is beyond the dichotomy of good (white hat) or bad (black hat), constantly reconfiguring it. In the next chapter on *Playing with Ethics*, we shall revisit the hacker culture in order to emphasize the ethics of hacking.

6. Playing with Ethics

In the literature explored in the previous chapters dealing with both fictional representation of artificial intelligence and cyberspace to scientific explanations of reality and machines, we find that cyberspace, where A.I. operates and coding turns digital objects into being, is a realm set between real and unreal influencing both artificial constructions and the physical, offline, world. As such, cyberspace is not completely disconnected from the physical world; instead, it seems, it acts as a bridge between two, or more, worlds.[75] One world is our physical reality which we still define according to Newtonian terms, while the other(s) are user-generated and can be as many worlds as there are users. This point located at the ontological intersection of real and unreal, artificial and natural, resembles a liminal space, or a "cyberliminal" space.

Cyberliminality is the point "in between and betwixt", a play between worlds. Cyberpunk protagonists such as Case in *Neuromancer* prefer the online world and see physical reality as a form of exile. In turn, cyberspace can be seen as either escapist, an artificial reality with the power to satisfy the illusions of desire unable to materialize in the real world, or a welcomed exile. In this sense, it would be inappropriate to define cyberspace as a new realm completely disconnected from the other by employing language such as "the fifth space" or, to deploy violent contest, "the fifth battlespace" alongside air, ground, sea, and space. Instead, the concept

[75] The experience of the individual in cyberspace makes it open to the development of multiple worlds within artificial reality. The example of gaming offered in this sense in the chapter Playing with Replacement is further supplemented by the services available online: YouTube channels are practically individualized worlds as well as content streaming services such as Twitch and others.

"cyberliminal" appears more appropriate when removed from the positive or negative outcomes of play and seen as an independent concept where play is free to develop its own ethics. A recurring theme is that of "exile" in cyberspace. However, the technological substratum of cyberspace, the physical networks of intercommunicating computers, removes the imperative of space for exile and may be interpreted rather as a form of escapism than exile. With the communication technology available today, one can be "exiled" in one place and have real time information about the place of origin, communicate directly with family and friends, and continue to be involved in local affairs from afar. Therefore, the idea of the *cosmopolis* is facilitated by contemporary communication technologies and also facilitates the global perspective.

Another important element to consider in this discussion is the "information overload" transposed to the effects of contemporary digital technology. While the concept alone reminds of the abundance of information available today, such as news and media, it is also pertinent to suggest that rather than having the effect of offering a plethora of data to solidify one's sense of reality, it has, in fact, greatly diminished it. We are at a point, today, when the often-called unreal objects of digital media, including A.I., affect our sense of the self and reality (as I have shown in previous chapters). As such, each individual contributes, in turn, to the effect of generalized confusion in the ongoing play of replacement and the perception of a total takeover by intelligent machines. It is assumed that the responsibility of developing moral technology rests on the engineer, which is undoubtedly true, but the usage of technology is, of course, the responsibility of each individual. Therefore, I would like to build on the Quantum Relations Principle, hacker ethics, and traditional virtue ethics, to redirect the discussion towards the individual. An

individual is not lost in the sea of information, she only requires a proper, functional ethical framework to coherently (and sanely) navigate its constant floods, storms, and calm waters.

6.1. Cyberliminality

Instances of liminality in literature go back to *Gilgamesh.* Within a literary work, liminality can be observed in the journey of the protagonist from point A to point B, as we can see in *Odyssey*. However, within the Homeric epic we can identify moments of liminality within the liminal process. For example, Odysseus descends into the Underworld and similar descents, *katabasis*, take different forms in most works of fiction. In cyberpunk literature, or in stories and novels like *Pygmalion's Spectacles*, liminality becomes a space, either a fully articulated environment set outside of physical reality of the protagonist, or an emerging reality that is acted upon and developed by the characters of the novel, such as cyberspace or metaverse. A brief overview of the concept of liminality may further clarify my approach.

Blanka Domagalska writes that "the space between the human and technology is occupied by new hybrid ontology." This new hybrid ontology is certainly related to the use of technology and the human body. "This territory has been traditionally theorized in the figure of the cyborg, an entity marrying in its body various orders of being such as human, plant, animal and technological. This conceptualization, however, still places the emphasis on the individual body. The liminality that interests me deemphasizes the individual ego and creates a connected world of objects as well as subjects." Liminality in this case is the merging of technology and the human body, and the result is that "the new entity is not human in a traditional sense and not an object, but an entity suspended between these ontologies" (Domagalska, 2014, p. 411).

Domagalska emphasizes the point of merger between the human body and technology. The cyborg takes a primary position in the cybernetic imagination of the second half of the 20th century. In turn, Donna Haraway defines the cyborg as "a hybrid of machine and organism, a creature of social reality as well as a creature of fiction" (Haraway, 1991, p. 149). Thus, this merger of human and machine is at the intersection of both nature and technology.[76] But cyberspace has always been seen as an independent space outside of physical reality, rather than something in-between. Nevertheless, a close second is the idea of a mixed reality which refers to the experience of the physical reality through technologies such as Google Glass or Oculus Rift (Greengard, 2019). Nevertheless, mixed reality is a balance between technologies of reality and physical reality allowing the user to be located in between the two realities.

Liminality was initially observed in the *rite of passage* trials and experiences that anthropologist Arnold van Gennep studied in primitive cultures. In a rite of passage, van Gennep noticed, the individuals experience a phase of separation from the community. Being in a point outside the community, the center of the individual's social and cultural experience, he experiences the second phase characterized by confusion and ambiguity defined by van Gennep as "limen" or "margin", a state of *aporia* or the katabasis journey, as we have mentioned. The third phase is the return, the "reaggregation" or "incorporation" which, as Victor Turner writes, "includes a symbolic phenomena and actions which represent the return of the subject to their new, relatively stable, well-defined position in society" (Turner, 1979, p. 16). These phases, according to van Gennep, constitute "the pattern of the rites of passage" (van Gennep, 2019, p. 191).

[76] The study of human-machine merger was studied briefly as "cyborg anthropology". See "Defining Cyborg Anthropology" here: http://cyborganthropology.com/Defining_Cyborg_Anthropology

Building on these patterns, Victor Turner discovers in his anthropological work that "liminal entities are neither here nor there; they are betwixt and between the positions assigned and arrayed by law, custom, convention, and ceremonial" (Turner, 1995, p. 95; Duffy, 2011). The liminal process, as such, can be seen as accessing a space of transition towards the new self.[77] Cyberspace or virtual reality serves this purpose. Cyberspace is not simply a whole, independent medium where nothing happens. On the contrary, the cybernetic imagination as well as current practices show that cyberspace is in perpetual development by users and innovators alike. Furthermore, the distinction between user-innovator is removed as one can become a creator-trickster and hacker.[78]

In this liminal space defined by technology and technical innovation, ethics take a new dimension. Just like the space itself is developing so its ethics are generative. Nevertheless, in this space we can also observe certain ethical patterns forming, while preserving the ability to experiment with new ethics as well, outside of traditional views. In this sense, we turn again to the Quantum Relations Principle and the hacker culture to evaluate the ethics of cyberspace in order to discover a feasible moral response to the problem of replacement

6.2. Ethics in Quantum Relations Principle

[77] August J. Cwik builds on the concept of active fantasy in Carl Jung to explore the development of imaginal play-spaces. Active imagination, as Cwik discovers, is a transitional phase and an important technique "which embodies the transcendent function and leads to the establishment of an imaginal play-space."See Cwik, August, J. Active Imagination as Imaginal Play-Space, in Nathan Schwarts-Salant and Murray Stein (eds.), *Liminality and Transitional Phenomena*, Chriron Publications, 1991, pp. 99 - 114

[78] This has been largely achieved due to efforts of technological education initiated by the early hackers through Phrack and 2600 Magazine. As such, the user is encouraged to become a developer of a service that will benefit others too. In addition, as hackers have been compared to trickster characters, their similarity is further accentuated by the concept of liminality. In the literature describing tricksters as liminal entities, they are either bound to exist in liminality, always being in between the world of the gods and that of men or play with boundaries and act in continual transit between realms. See Waddell, Terrie, *Wild/Lives. Trickster, Place and Liminality on Screen*, Routledge, 2010

The Quantum Relations Principle is located in this in-between space at the convergence point of computation and physical phenomena. Proceeding from quantum physics and the theory of relativity, it argues that their principles should apply to the physical world as well as to the human world. In essence, QRP conforms to "system thinking" that was highlighted in the previous chapter and which became, and still is, the dominant conceptual thinking, but adds that systems are not isolated but interact between them. In QRP, it is no longer a matter of distinction, as the age-old problem goes, between object and subject, but a mutual causal dependency for the development of reality. According to Schloer and Spariosu:

> "QR emphasizes the need to shift scientific thinking away from the quest for objective Truth towards the recognition that all scientific data are observer-dependent and that all approaches to reality, including scientific ones, are influenced by subjective experience. QR departs from Cartesian dualism of subject and object, or mind and matter, by placing mental and physical events or processes on the same experiential continuum. Consequently, it describes mind or "consciousness" by the same relational processes that contemporary theoretician physicists used to describe physical and other systems" (Schloer & Spariosu, 2016, p. 77).

Therefore, QRP moves away from a rigid understanding of reality, "matter and mind, body and consciousness are not ultimate reality. They have no rigid boundaries, being simply different conceptualizations to bring order into experience" (*Ibid.*, p. 78). In addition, concluding that the "traditional notion of causality" as linear, physical, and local is inadequate, Schloer and Spariosu propose a different direction uniting the "notion of linear causality with the nonlinear and nonlocal concept of mutual causality" for a computational theory that should provide "qualitative account of the reciprocal relations among the systemic networks that our minds and bodies constitute" (*Ibid.*, p. 79).

Conforming to system thinking and drawing on “feedback loops” as shown in the works of numerous scholars, the authors point to the Gaya hypothesis that suggests that the world interconnected, complex, self-organizing system, as well as the teaching of Gautama Buddha of “dependent co-arising" or “dependent origination”. Thus, “the Buddhist view of subjectivity or human agency applies to both general systems theory and Quantum Relations: it dissolves the subject/object dichotomy into reciprocal play of relations between fluid variables” (*Ibid.*, p. 84). In other words, the distinction between physical and digital object is no longer necessary to make. Instead, the informational properties of both harmonize ontologically. Furthermore, Schloer and Spariosu note that events and not substances form our ontological basis. “In this kind of process ontology, human events and agencies are not understood in terms of qualities, attributes, and substances, but in terms of specific contexts and relations” (*Ibid.*, p. 85). In this sense, the agent, observer, is both a consequence and the cause of action. Stating the influence of Daoism, the authors note that it “makes no substantialist distinction between subject and object, but operates instead with a dynamic process of fluid interrelationships, based on mutual causality” (*Ibid.*, p. 85). Schloer and Spariosu sum up their theory as follows:

> “[o]ur QR-models are based on the “web of life” in its most diverse and complex aspects, including human relations and interactions. Unlike reductionist scientific theories, which generate reductionist technological platforms, QR implicitly acknowledges diversity and alterity as the very conditions of existence. Whereas the reductionist theoretical models perpetuate the hegemonic pretensions of mainstream Western science, attempting to impose its dualistic, Cartesian perspective on all cultures in the guise of objective, universal knowledge, DFO/FOR models can take into account and process widely different cognitive perspectives, including linguistic, philosophical, cultural, sexual, and other observer-dependent variables. At the same time, they can continually and automatically update, reframe and reorganize their data as new global realities emerge” (*Ibid.*, p. 87).

The Quantum Relations Principle solves the ontological problem discussed in previous chapters by moving away from a static, Newtonian-style notion of reality and of an

understanding of being still immersed in Cartesian dualism, to a dynamic, "process ontology" in which digital and physical, mind and matter, contribute in mutual causality to produce being. However, the ethical implications of QR-based technologies, and the framework itself, are not fully articulated by Schloer and Spariosu and, therefore, is limited in the development of an ethical approach to the play of replacement. The authors do not provide deontological suggestions for the use of QRP and wisely resist the utilitarian temptation. Instead, Schloer and Spariosu seem to point towards virtue ethics: "Consequently, it is a particular human mentality that will determine the uses of QR technology, and not the technology itself that determines this mentality and the type of life-enhancing or self-destructive feedback loops it will generate. It bears stressing, again and again, that whether we move from a digital age to the quantum age and beyond, the future of human development will depend less on advanced technological platforms and much more on our collective ability, above all, willingness to shift to a mentality and practices that are oriented toward global intelligence and planetary wisdom" (*Ibid.*, p. 89).

The issue repeated by the authors, and Spariosu in particular in other works, is the change of mentality. Accordingly, if our power-based mentality seeking replacement continues through new technology, then, inevitably, the result will be "self-destructive feedback loops" rather than "life-enhancing" usages of technology. In addition, this problem seems to have less to do with the technology itself than the ethics mandating the usage of it. This change of mentality is a tremendously difficult task because the issue of replacement, of dualism, and the power-driven agonistic play of replacement is well articulated in Western culture, as I have shown in the chapter on "Playing with Replacement." The forces keeping up the development of potential civilizational changing technologies, as was emphasized by Capek, are great and active

confirming the power-driven play of replacement, the "either-them-or-us" mentality, shown by Forster and representative of much of the 20th and 21st cybernetic imagination.

Schloer and Spariosu's attempt to formulate a theory based on Daoism, Buddhism, and religious beliefs of the East reflects the loss of confidence in Western culture shown in the 1960s and 1970s with the import of Eastern philosophy, beliefs and gurus. However, with all the problems the Western culture has, to change its mentality I believe it is necessary to find within it the examples and resources to transform itself. In particular, and pertinent to our study, we need to move away from the philosophy of technology developed in the Western world, which is often mesmerized by ideological presuppositions [add Latour and others] and observe the practice of technology and the culture developed around it. In this sense, I do not mean to look at all cultures forming along all Western technologies, which is out of the limits for this study, but to observe the moral development and predispositions of the culture predominantly responsible for what we call today computer sciences: the hacker culture. The hacker's behavior, competence, use and development of tools for hacking, as well as the consequences of hacking rest on a form of hacker virtue. The moral analysis of the hacker culture reveals a set of ethical beliefs guiding the personal development of the hacker. Of course, since the 1970s, the computational sciences have been monopolized by the university and a stream of computer ethics papers began to flow. Furthermore, the works of Spariosu tend to focus on ancient Western philosophy and literature. As such, we can attempt to identify the concept of virtue in ancient thought and transpose it to our digital ethics.

6.3. Hacker Ethics

The temptation for investigating the ethics of the hacker culture is to look at contemporary ethical hacking or to the general topic of computer ethics. Accordingly, the former would be easily explained in terms of the "white hat" ethics that form the basis of adjacent coding professions such as system penetration tester, while the latter aligns or adjusts current ethical theories. We would then operate in a closed system and will inevitably produce conclusions similar to those in the field of cybercrime. So, we will instead look at the hacker culture in its own development in an exploratory fashion and observe what ethical directions can be noted from its original intention. Several works are relevant here, as well as my own investigations into the hacker culture. Drawing from the conclusions in this section and from the previous chapters, I shall then proceed to formulate an ethical framework.

Before proceeding, however, it is also important to mention that the very idea of the hacker culture being able to provide a set of ethical references pertinent to future positive development may sound problematic, to say the least. In the last decades, hacking was associated with crime, in one way or another, whether in political hacking or online identity fraud. This may be because, on the one hand, the computer security industry is actively promoting the idea of hacking as crime, thereby probably ensuring a growing client base, and, on the other hand, the media coverage of hacking is sensational or in line with the cyber apocalypse narrative. For example, Kevin Mitnick was portrayed in the New York Times as a top villain eluding the FBI for months, even being called a "computer terrorist." When on trial, Mitnick was held in solitary confinement for more than six months, because the judge was afraid that Mitnick was able to whisper a computer virus into the telephone system. Other hacking pioneers were convicted and served sentences in jail for acts that would be considered nuisances rather than actual, violent

crimes. Despite these well-publicized examples, behind the media curtain lies a story of hacking defined by personal responsibility, innovative curiosity and philosophical wonder.

The reaction to the technological abilities, knowledge, and understanding of how systems operate, and implicitly systems thinking, shown by the hacker culture is understandable, given the potential to cause severe harm to the national economy or even to activate military weapons. On the other hand, the fear exceeded the actual crimes committed by hackers, and legal punishment was largely based on it. Nevertheless, as shown in the films *Hackers* and *Take Down,* the hacker mind took the witch hunt as yet another challenge to understand the patterns shaping the fear of their craft and, even more, as motivation to continue their rather undefined moral mission. Emmanuel Goldstein, the founder of *2600 Magazine*, reflecting on the "hacker philosophy," notes that "the fear and suspicion that people were met with when they demonstrated more knowledge of technology than those who were allowed to use it was what convinced us that we were really on to something" (Goldstein, 2008, p. 208).[79]

One of the most revealing documents on the hacker ethical imperatives is "The Constitution of a Hacker" published in *2600 Magazine* in March, 1984. The document shows the interests of hackers in technology and attempts to answer the question of motivation while stressing their duty to attempt to penetrate systems to ensure better security before someone else

[79] The language the media used to portray hackers was spearheaded by John Markoff of the New York Time. John Markoff covered the story of Kevin Mitnick for the New York Times. The book *Takedown,* Voice, 1996, details FBI's pursuit of the hacker. Adapted to film, it explores the hacker as a criminal obsessed with the desire to know how systems work. Hackers did not respond well to the coverage, the book, or the film and hacked the New York Times calling out Markoff with messages such as "Hi John Markoff, this one is for you. We expect front page like you promised.... Do you have nightmares about helping imprison Kevin?" See Hessendahl, Arrik, All the News that's Fit to Hack, in *Wired*, 09.14.1998. Available here: https://www.wired.com/1998/09/all-the-news-thats-fit-to-hack/ Ever since the Markoff-Mitnick incident, the hacker community has shown distrust of reporters and are reluctant to share their knowledge with investigative reporters.

does according to a questionable agenda. Communication technologies, particularly the telephone (which one used to connect to the internet), are seen as objects of wonder and potential. The hacker mind intends to discover ways in which the technology can be used for other purposes than its original design, as the document says, "to play around" with it. But, "you're not supposed to have fun with them" so that "for the first time, these technological enthusiasts posed a 'threat' to technology by reaching out and touching it rather than simply using it without asking any questions" (*Ibid.*, p. 209). The curiosity about how it works, what it can do, and what are the vulnerabilities of a technology is a central motivation to the hacker. The document phrases the question as follows:

> "Why do they [hackers] do this? What do these people possibly have to gain by breaking into computer systems and seeing things that don't really concern them or that is of no possible use to them? In the great majority of cases, computer hackers don't gain anything material or financial from their explorations. Add to that the high risk of getting caught and it becomes very hard for the average citizen to understand what motivates these people" (*Ibid.*, p. 209).

The answer is that hacking is playing for its own sake. The intention is not financial gains and not the information that is locked behind strict security. Instead, the motivation comes from the challenge itself, an exploratory adventure into the systems forming cyberspace:

> "A genuine hacker breaks into computers for the challenge. He's not out to save the world or to destroy it. He is not out to make a profit out of what he's doing. Therefore, it's not fair to categorize him as a criminal and it's just as wrong to say he's some sort of a savior" (*Ibid.*, p. 210).

However, the Hollywood industry often appealed to the hacker culture to provide characters with criminal intentions coupled with apocalyptic messages such as *White House Down* (2013), or hacking was deployed to save the world from alien invasion in *Independence Day* (1996), albeit other films that have hacking as a central theme are mostly connected to shutting down the U.S. economic system or military equipment. As the document says, this is not

the concern of the hacker. Instead, the hacker is simply there to play with computers and challenge security systems. And this play reflects an attitude towards technology and life. It is this apparent careless or childish form of playing with technology that remains disconnected from serious power games. If hackers engage in such contests, then it is to mock them and reveal their absurdity in the process. Nevertheless, the major technological innovation of the past century, particularly after the Second World War, often attributed to North America ignoring similar quests for discovery in Europe and elsewhere, was done not in the universities nor by government, but by computer enthusiasts who would have an interesting life style and a commitment to playful exploration and ethical stability. Pekka Himanen describes the hacker way accurately:

> "Sandy Lerner is known not only for being one of the hackers behind the Internet routers but also for riding nacked on horseback. Richard Stallman, the bearded and longhaired hacker guru, attends computer gatherings in a robe, and he exorcises commercial programs from the machines brought to him by his followers. Eric Raymond, a well-known defender of hacker culture, is also known for his playful lifestyle: a fan of live role-playing games, he roams the streets of his Pennsylvania hometown and the surrounding woods attired as an ancient sage, a Roman senator, or a seventeenth-century cavalier" (Himanen, 2001, p. 5).

Probably this playful attitude towards life and technological exploration is enough to get anyone seriously developing unserious stories surrounding the unintelligibility, if not flat-out paradox, about someone capable of producing something like the "Internet router" but falling short of sober, suit-and-tie imagination about competence. In this regard, as Himanen notes, the hacker culture may have had a deeper impact on society than we expected on three important levels. First, hackers create enthusiastically, and so they have offered an alternative to traditional Weberian protestant work ethic; second, their way of developing free software, and even operating systems, has influenced the market economy; and third, through their culture of interconnectedness in constant peer-to-peer and peer-to-many online communication. As

Himanen argues, the hacker ethic can be observed in three categories posited outside conventional considerations: work ethic, money ethic, and nethic. Furthermore, to be a hacker is a state of mind with afferent ethics: "the more I thought about computer hackers, the more obvious it became that what was even more interesting about them, in human terms, was the fact that these hackers represented a much larger spiritual challenge in our time. Computer hackers themselves have always admitted their wider applicability of their ways" (*Ibid.*, p. VIII). It is, indeed, a wide applicability of hacker ethics and attitude: a person can have any profession, computer related or not, and be a hacker. The requirement is not the profession, but the philosophy motivating the individual to perform.

Linus Torvalds, the inventor of the operating system Linux, a free and open-source OS that forms the core of Google and Apple run devices, Android and MacOS, respectively, identified this philosophy of action in three stages. He playfully names it "Linus's Law," "in my normal humble and self-deprecating way" (*Ibid.*, p. XIV) and its process rests on three elements: survival, social life, and entertainment. The first one, as he notes, is a truism since "any living thing needs to survive as its first order of business" (*Ibid.*, p. XIV). And assuming this need, then, Torvalds inquires about the motivation to survive and for what one is prepared to die for. Social life as a motive has numerous examples in literature and history, but, as he points out, "entertainment" is a strange choice. Nobody in their right mind would want to die for entertainment. But Torvalds has something else in mind for his entertainment: "It's chess. It's painting. It's the mental gymnastics involved in trying to explain the universe" (*Ibid.*, p. XV). It is the type of entertainment that "gives your life meaning" (*Ibid.*, p. XVI). Indeed, by entertainment he means playing, and the computer serves as a tool for playful exploration of its creative potential. Here is how he puts it:

> "A "hacker" is a person who has gone past using his computer for survival ("I bring home the bread by programming") to the next two stages. He (or, in theory but all too seldom in practice, she) uses the computer for his social ties – e-mail and the Net are great ways to have a community. But to the hacker a computer is also entertainment. Not the games, not the pretty pictures on the Net. The computer itself is entertainment. This is how something like Linux comes about. You don't worry about making that much money. The reason that Linux hackers do something that they find it to be very interesting, and they like to share this interesting thing with others" (*Ibid.*, p. XVIII).

Far removed from the hacker as criminal, the narrative ever-present in the media, Linus Torvalds points to the essence of a culture of innovation outside of the Western cultural imperative of replacement.

"The Constitution of the Hacker" adds that attempting to break into computer systems is the duty of a hacker: "One could almost say that a person with hacking abilities has an *obligation* to try and get into as many different systems as he can" (Goldstein, 2008, p. 210). The justification is rather practical: all attempts to penetrate security systems, particularly if successful, would lead to better security. In doing so, the hacker invalidates the security imagination of "Sysops" (system operators) without attempting to replace the entire system. Hacking is playing in cyberspace with the tools designed for it, to expand it, to make it better, and "our world is turning into a technological playground" (Goldstein, 2008, p. 211). This play is rather disruptive and innovative at the same time. The play of hacking was associated with the trickster archetype and, as such, posits the hacker in the liminal space between online and the offline worlds. In this sense, like the trickster character, the hacker becomes liminal. However, the typical transfer of the individual into a liminal place, often *katabasis*, becomes in the case of the hacker a connection to the liminal space of cyberspace and, because the code behind cyberspace creates the world, *katabasis* becomes *kata (digital) physin,* according to (digital) nature. In this liminal space, hackers develop a new ethics, one that is based on individual

conduct, and formulate principles that (a) emphasize skill and innovation and (b) show the moral motivations of their activities in cyberspace.

As we have shown in the previous chapter, the first imperative of the hacker culture is impressive coding skills in order to achieve a hack. In this sense, scholars agree on the following:

1. Simplicity. The hack must provide a simple solution to a complex problem. Therefore, it is required to be simple but impressive.
2. Mastery. For a successful hack, the hacker needs to have formidable knowledge of the target system.
3. Illicitness. Once knowledge is gathered and solutions are formulated, the hack acts against the rules of the target system and, inevitably, the rules of society as well.

The 'thrill' of success comes, on the one hand, by following these skill-testing principles and, on the other, it acknowledges the mastery of the hacker, his elite (l33t) status, within the community. If, however, the hacker resorts to using software and tools designed by other people to perform his own hack, he becomes a script-kiddie – either a beginner or a hacker showing insignificant amount of knowledge about the system. In addition to these characteristics of a successful act, Steven Levy has come up with a list of motivations that take up the moral aspect of the culture and highlight the ethical duty of the hacker (Levy, 2010, pp. 28-29):

1. "Access to computers—and anything which might teach you something about the way the world works—should be unlimited and total. Always yield to the Hands-on Imperative!
2. All information should be free.
3. Mistrust authority—promote decentralization.
4. Hackers should be judged by their hacking, not bogus criteria such as degrees, age, race or position.
5. You can create art and beauty on a computer.
6. Computers can change your life for the better."

The framework provided by Pekka Himanen conforms to these already established ethical directions of hacker culture. The challenge to Protestant work ethic, as Himanen describes it by referencing Max Weber, is the hacker's enthusiastic programming as we can see

in (1): it is not a work concerned with the result but with the process itself. Points (1), (2), (5), and (6) show that the hacker culture does not innovate for profit, therefore the money ethic is challenged as well. Instead, hackers play with digital technology for a higher purpose. Point (3) shows that the distribution of information should not depend on a central hub or an institution, while (4) reflects the meritocratic nature of hacker culture. Furthermore, the activities of *2600 Magazine* and *Phrack Magazine*, as well as the political software development of groups such as Cult of the Dead Cow, show the hacker culture attempting to clarify or dispel the magic of communication technologies and instruct readers on how it properly works, while conforming to the other principles explained in (a) and (b) above and being concerned with the freedom of the individual in the digital age, where technology can benefit mass surveillance and curtail the spread of information, therefore hinder freedom of speech.

We find the hacker culture to be in appreciation of *togetherness* with the world. The principle that we recall here is *oikeiosis*, meaning an appropriation of the challenges of non-technical people in the digital world. Together, these principles provide the framework for personal ethical conduct in cyberspace, as well as ethical innovation for the age when reality becomes indistinguishable from its virtual reconstruction and artificial intelligence is imagined to replace humanity. As such, these ethics of innovation and conduct move away from agonistic play of replacement and promote, instead, a playful exploration of our digital age for the benefit of all. In doing so, and given the global aspect of contemporary communication technology, it may also provide further achievable steps towards “planetary wisdom” by making local problems global and global problems local.

6.4. Digital Virtue

As we have seen, hacker culture provides the elements for a virtue ethics for virtual reality and artificial intelligence. We thus propose a playful exploration of technology according to digital nature and social concern.[80] In order to do so, we have to reconsider the approach to

[80] Digital virtue is grounded in virtue ethics, an agent-centered branch of moral philosophy attributed to Aristotle (Nicomachean Ethics, Book II), although the four cardinal virtues (later supplemented with three theological virtues) are mentioned by Socrates in *Republic* (427e and 435b). Virtue ethics' primary supposition that good character fosters good action reduces the focus of action and behavior in the digital age to the individual. After Plato and Aristotle, virtue ethics found primary expression in Stoicism. Deterministic and materialist, the Stoic philosophy has found recent resurgence with particular impact on Silicon Valley, see Bowles, Nelly, Why Is Silicon Valley So Obsessed With the Virtue of Suffering? in *The New York Times*, March 26, 2019. And while the Silicon Valley prefers the psychological tools of Stoicism, the philosophy itself has a history of influencing the development of science and philosophy, and, as such, it could potentially be accepted as a user-innovator guiding moral philosophy framed on its major pillars of reason and social life. In addition, Stoicism influenced philosophers such as Spinoza and Leibniz who, in turn, provide the groundwork for contemporary computational sciences. Alexander Jones, writing *The Stoics and the Astronomical Sciences*, argues that "in Stoicism, as in Epicureanism, an understanding of the physical place of humanity in the universe was an integral component of a system of thought underpinning each sect's ethical commitments," in Brian Inwood (ed.), *The Cambridge Companion to the Stoics*, Cambridge University Press, 2003, p. 328, and proceeds to show the Stoic influence on science drawing from multiple ancient sources. Indeed, the Stoic tradition rests on three major pillars: physics, logic, and ethics. The quotation above is an example of Stoic physics. In this sense, it can serve as a framework to understand the place of humanity in the digital universe as well. Donald Rutherford explains the Stoic thought of Descartes in *On The Happy Life: Descartes vis-a-vis Seneca*, in Steven K. Strange and Jack Zupko (eds.), *Stoicism*, Cambridge University Press, 2004, pp. 177 – 197; The influence of Stoicism in modern psychotherapy is well documented, see DeBrander, Firmin, Psychotherapy and Moral Perfection: Spinoza and the Stoics on the Prospect of Happiness, in Steven K. Strange and Jack Zupko (eds.), *Stoicism*, Cambridge University Press, 2004, pp. 198 – 213, as well as Robertson, Donald, *The Philosophy of Cognitive-Behavioural Therapy*, Routledge, 2019. Particular to the Stoic influence on science is Peter Barker's chapter *Stoic Contributions to Early Modern Science* in Margaret J. Osler (ed.), *Atoms,* Pneuma, *and Tranquility,* Cambridge University Press, 1991, pp. 135 – 154. Here, Barker uses Cicero's *De natura deorum* to trace the Stoic inspiration of natural philosophy. As he notes, "scientific authors in this period often combine elements of Stoic ideas with other positions" (p. 140), but it is through Cicero's work that Rebelais, Jean Pena, Johannes Kepler, Sebastian Basso, have used elements of Stoicism for scientific inquiry. As such, Barker concludes that "Stoic ideals therefore re-entered scientific debates not as a complete system, distinguishable from an Aristotelian or an Epicrean system, but in eclectic combinations with ideas that in antiquity had been regarded as antithetic" (p. 154). For example, As Barker points out, Sebastian Basso combines Stoic cosmological ideas with atomism contrary to the strict opposition of the ancient Stoics to atomism. "While the orthodox Stoic view takes the earth to be the center of planetary motions, and Basso accepts this view, Brahe combines one Stoic-derived celestial substance with a geo-heliocentric world system, and Kepler combines another with heliocentrism" (p. 154). B. J. T. Dobbs attempts to explain the Stoic influences on Newton. His concern is to underline Newton's religious and scientific objections to Epicureanism while stating that "alchemy, in particular, carried much Stoicism into the seventeenth century, and, as is now fairly well known, Newton studied alchemical literature exhaustively." It appears that the Stoic concept of *pneuma* was particularly attractive to Newton. Dobbs quotes the following passage from Newton: "Those ancients who more rightly held unimpaired the mystical philosophy of Thales and the Stoics, taught that a certain infinite

questions such as "why this fear of robots?" as well as "what does it mean to be human?" or "what is reality?" by placing the focus on the individual. While these questions are worth pursuing, particularly an in-between position uniting literary studies, philosophy, and the computational sciences in an academic setting, they are too broad to be understood by individuals operating within the confusion and possibilities of our digital future. In addition, digital innovation rests on the individual acting and building upon the already established computational environments and frameworks for artificial intelligence research. Furthermore, as we can see from the hacker culture, history, and the cybernetic imagination, the impact of an individual in this liminal space can have major positive or negative consequences. Accordingly, to promote elaborate general ethical frameworks for A.I. and software development would fall short of acknowledging the individual in favor of a narrative that could resemble the fiction of the cybernetic imagination taken as fact. Digital virtue, therefore, is a practical philosophy for the digital life, providing guidance for playful innovation and use of software and technologies. Moreover, innovation-user is not, in this framework, a dichotomy of choice where one chooses to innovate while the other consumes. Instead, it is viewed as similar practice: the user can innovate on the innovation itself.[81]

---

spirit pervades all space *into infinity*, and contains and vivifies the entire world. And this spirit was their supreme divinity, according to the Poet cited by the Apostle. In him we live and move and have our being." Newton's reference here is to the Stoic conception of God and the universe. For the Stoic, as Dobbs notes, *pneuma*, "the breath of life," was a compound of air and creative fire thought to leave the body upon death to return to the universe. See Dobbs, B.J.T., Stoic and Epicurean Doctrines in Newton, in Margaret J. Osler (ed.), *Atoms,* Pneuma*, and Tranquility,* Cambridge University Press, 1991, pp. 221 – 238. For a comprehensive work on Stoic physics see Sambursky, Samuel, *Physics of the Stoics*, Princeton University Press, 1987. For a general overview of Stoicism see Brennan, Ted, *The Stoic Life*, Oxford University Press, 2005 and Hadot, Pierre, *The Inner Citadel: The Meditations of Marcus Aurelius*, Harvard University Press, 2001

[81] In the early days of computing, specific software had to be developed for a particular machine. Universal operating systems, such as Windows, Linux, or MacOS, have been developed later in an economic environment that favored networked computers. With limited computing time allocated to users, the software of the machine would receive small corrections and improvements from its users.

Digital virtue builds on two main premises: (1) physics – process ontology: to understand digital technology and incorporate its ontological effects outside of the tension of replacement between real and artificial, thus moving towards "process ontology" and (2) ethics – oikeiosis: playful innovation based on the principle of appropriation of world problems leading towards planetary wisdom.

Physics – Process Ontology

In *The Sciences of the Artificial,* Herbert A. Simon notes that "natural science is knowledge about natural objects and phenomena. We ask whether there cannot be "artificial" science – knowledge about artificial objects and phenomena" (Simon, 1997, p. 3). Since the first

---

These were considered "hacks." See Green, Leila, *The Internet: An Introduction to New Media,* Berg Publishers, 2010 pp. 19 – 42. Of course, intellectual property hinders the practical application for this process significantly. However, when copyrighted software and products allow at least a small portion or a feature of it to playful exploration by users, the results can be staggering. For example, Blizzard Entertainment's Warcraft III: Reign of Chaos (2002) provided a customizable multiplayer online battle arena (MOBA) titled Defense of the Ancients (DOTA). Users have built on it to such an extent that it became a highly successful independent game attracting millions of players and established a new genre of competitive gaming, MOBA. Other creators followed suit and developed successful games, including Blizzard Entertainment. The reticence of some developers to provide features such as this rests is financial, but the contribution the Warcraft III developers provided to the global gaming scene is extraordinary and facilitated, implicitly, the creation of a new genre that today competes with traditional sports. The money ethic aspect of the hacker culture and the enthusiastic programming and computer development shown by groups such as Homebrew Computer Club or by the contributors to 2600 Magazine redirect the focus from financial gain to the product itself and its potential. Thus, play for play itself. In the operating system world, the user-developer has built on Linux to develop Ubuntu, a cost free, open source OS with numerous enthusiasts dedicating their time and intellectual resources to facilitate access to computers and information. Ubuntu now offers free personal and professional versions of its OS for both computers and servers. The user-developer, in addition, provided free alternatives to the expensive Microsoft Office Suite, such as LibreOffice. Google's Android was developed on similar user-developer grounds to enable app development. Tim Rayner in *Hacker Culture and the new Rules for Innovation,* Routledge, 2018, notes that the success of the software industry rests on two principles: open standards and innovative culture. He writes that |the culture of openness has supercharged the startup industry. Open standards make it possible for entrepreneurs to build app-based services that combine data from different sources, to serve different needs" while the second factor while often overlooked, "it [innovative culture] concerns the culture of technological innovation, rather than technology itself. It concerns the way people coordinate, share, work and collaborate to produce innovation. This factor is hacking" (pp. 1 – 2).

publication of this book in 1969, our conception of natural/artificial objects has developed significantly. As we have seen in previous chapters, the cybernetic imagination carried this process to a point in which artificial, digital objects may not be artificial at all, and our reality rests more and more on perception. Of course, digital objects may not be natural particularly when considering a collection of digital objects enabling the mental processes of anthropoid machines. Simon adds that "computers have transported symbol systems from the Platonic heaven of ideas to the empirical world of actual processes carried out by machines or brains, or by the two of them working together" (*Ibid.*, p. 26).

However, process ontology considers the digital objects alongside other phenomena within the framework of data fusion objects and frame of reference producing superstructures that do not stand independently but return to further co-dependent causality with DFOs/FORs. In this sense, artificial intelligence and cyberspace have enhanced creative potential without creating standalone profiles of individual artificial/natural objects. Instead, I conceive of this ontology in QRP terms as well as traditional Stoic metaphysics where the existing world is an ongoing process of interconnected objects.

Jaques Brunschwig writes that "for them [Stoics], 'nature' (*phusis*) encompasses everything, including things, phenomena, and events which in other worldviews might seem to be 'super-natural' in some way" (Brunschwig. 2003, p. 206). In other words, Stoic metaphysics has no issue with incorporating artificial/digital objects within their philosophical system. Thus, reading Diogenes Laertius, Brunschwig concludes that the Stoics have not one but two metaphysics: "The one is merely a *part* of physics; the other is a study over and above their standard tripartition of philosophy" (*Ibid.*, p. 206). In this sense, Stoic metaphysics distinguish between a "specific" dimension and a "generic" dimension. According to the first, Stoics

propose five topics: bodies, principles, elements, gods, limits, place, and void. For the second, "generic" division they propose the world (*kosmos*), the elements, and the search for causes. The "void" is problematic here, because the Stoics traditionally do not accept any gaps within the wholeness of the world. Brunschwig notes in this matter that "given their belief in the unity, continuity, and cohesion of the *kosmos*, they denied the existence of any void within it; but they posited an extra-cosmic infinite void, in order to make room for its cyclical expansions and contradictions" (*Ibid.*, p. 207).

Bodies, in Stoic metaphysics, are existing physically as beings (*onta*), but most interesting is their acknowledgement of *incorporeals*, things that exist without being strictly *onta*, but "something." Among these things are place, void, and "sayables" or "things said" (*lekta*) (*Ibid.*, p. 213). Place refers to space that is something else "than the body, since the body can move to another place without ceasing to be what it is, and the place can be occupied by another body without ceasing to be what it is" (*Ibid.*, p. 214).[82] In essence, Stoics propose a "process ontology" that does not differ from the QRP co-dependent causality of objects of different ontological status.[83]

Ethics – Oikeiosis

The concept of oikeiosis is best shown in Philip K. Dick's *Do Androids Dream of Electric Sheep?* The Voight-Kampff test is described as a series of questions measuring response

[82] Of course, Stoic metaphysics are much more complex than presented here. But for the sake of clarity of argument, we resumed our discussion to the elements directly relevant for this study.
[83] Marcus Aurelius sums up Stoic ontology as *oikeiosis*: "Everything is interwoven and the connection is something holy, and one is not foreign to the other. Then it is weaved in and together builds the same Universe." Own translation from Mark Aurel, *Selbstbetrachtungen*, Patmos Verlag GmbH, 2010, p. 155

times and show of empathy. However, when observed in the context of the novel and the interaction of characters and expectations of android behavior, it becomes clear that empathy is not just what Philip K. Dick has in mind, but *oikeiosis*, that is, the ability to appropriate other people's problems as if our own and self-sacrifice as an act of virtue.[84] This becomes even more apparent in the sequel to the movie adaptation of the novel *Bladerunner 2049* where the protagonist, a replicant fully aware of his artificial nature, choses to self-sacrifice for humanity. Stoic *oikeiosis* means that the individual's actions are responsible towards the whole of humanity. A good action upon the self should be reflected towards humanity as well. In essence, there is little difference between the self and humanity, an idea that derives from the Stoic conception of the universe and of the "deity" that governs the causal interactions among things and incorporeals. It is, in addition, an early conception of the *cosmopolis.* Perhaps the clearest source on the subject is Hierocles who explains *oikeiosis* as overlapping circles:

> "The first and closest circle is the one which a person had drawn as though around a center, his own mind. This circle encloses the body and anything taken for the sake of the body. For it is virtually the smallest circle, and almost touches the center itself. Next, the second one further removed from the center but enclosing the first circle; this contains parents, siblings, wife, and children. The third one has in it uncles and aunts, grandparents, nephews, nieces, and cousins. The next circle includes the other relatives, and this is followed by the circle of local residents, then the circle of fellow-tribesmen, next that of fellow-citizens, and then in the same way the circle of people from neighbouring towns, and the circle of fellow-countrymen. The outermost and largest circle, which encompasses all the rest, is that of the whole human race. […] For although the greater distance in blood will remove some affection, we must still try to assimilate them. The right point will be reached if, through our own initiative, we reduce the distance of the relationship with each person" (*apud* Brennan, 2005, pp. 157 – 158).

[84] Troels Engberg-Pedersen has written probably the most detailed study of the theory of oikeiosis in Stoic philosophy. Engberg-Pedersen places oikeiosis at the center of Stoic ethics as "the key account of their freedom of the will" in "the world as unfolding in accordance with strict causal lawlikeness (fate)" (pp. 232 – 233). See, Engberg-Pedersen, Troels, *The Stoic Theory of Oikeiosis*, Aarhus University Press, 1990

This idea of a connected world resulting in the implicit concept of acting just on oneself-humanity by attempting to remove the distinction between self and the other is expressed by the hacker culture using an ontological metaphor: the world is a computer. Kevin F. Steinmetz observes that "the *world as a computer* metaphor allows for hackers (and others who adopt this metaphor) to perceive the world as being open and intricately connected. Everything has a logical function and place" (Steinmetz, 2016, p. 28). It is this sense of moral responsibility towards the interconnected world that allows for playful exploration of technology and hacktivism to uphold and implement the principles outlined by Steven Levy.

In this sense, digital virtue rests on (1) process ontology as a departure from traditional ontological frameworks that have difficulties incorporating artificial reality / cyberspace into their views of reality; (2) playful exploration of technology considering (a) user-innovator cooperation and (b) self-humanity action. In the next and last chapter, I shall present the conclusions of this study and propose a way forward for education in virtue-based digital innovation.

Conclusions

"Why this fear of robots?" was one of the research questions of this study. Technophobia, Asimov's answer, is not enough. As we have seen, the cybernetic imagination rests on the principle of playing with technology. If we consider the expectation of replacement by intelligent machines and cyberspace or virtual reality, then our path towards it is irrational. It is a path of self-defeat and auto annihilation. It hardly makes any sense. But at the essence of replacement is something irrational as well: play. Thus, the irrational fear of robots may also rest on, or may be carried by, irrational, agonistic play.

While there are numerous examples of intelligent machines designed to resemble humans in the ancient and medieval worlds, it is largely in the 19th century that the expectation of replacement is articulated in literature. Starting with Mary Shelley's *Frankenstein, the Modern Prometheus*, Western literary imagination moved towards science while simultaneously being anxious about its possibilities. In the novel, the Creature harasses his creator, is violent, kills, and eventually expires with his creator. Numerous other literary works imagine similar instances of violent artificial constructions posing an existential threat to humanity. In the chapter on *Playing with Replacement*, I have shown how this existential play was carried out within the parameters of violent contest. This form of play producing zero-sum games spread in the cybernetic imagination, which in turn influenced the development of science and technology. Therefore, the expectation of replacement is not confined to a single academic discipline or just to technology or science; instead, it touches almost all intellectual productions of Western culture. Spariosu's

conclusion that the history of the Western mentality is the history of violent contest is confirmed in the cybernetic imagination, which deploys the metaphor of the robot along these lines.

In this agonistic play, the intelligent machine is either at the forefront of narratives, when robots actively seek the replacement of people playing violent contests, such as in Capek's *Rossum's Universal Robots,* or looms in the background, guiding the plot towards agonistic expectations, as in the case of Meyrink's *Der Golem.* In the latter, even if the novel balances metaphysical questions and seeks ancient, mystical wisdom, the agonism represented by the Golem lurks behind the narrative. In doing so, the artificial construct becomes a mirror of humanity's expectations and fully reveals the underlying cultural values of the West. Thus, the cybernetic imagination exports the violent play of replacement, in a more subtle reformulation, to the sciences.

As we have seen in the chapter on *Playing with Replacement*, post-war disciplines formed around computing tend to monopolize the humanities or adjust them to their terminology in the pursuit of knowledge and truth. The result of this tendency is that literature is no longer representational and is expected to be explained, rather than understood. The methods of scientific inquiry transported into the humanities have developed the field of the digital humanities or similar fields where the term "computational" signals a reformation of and a protest against the traditional humanities. This approach favors the use of the quantitative methods of computational research to explain the qualitative features of literary works. Of course, such methods can produce valuable contributions to the study of the humanities, but the tendency is not to employ them as adjacent methods or opportunities of interdisciplinary studies; rather, they tend to replace the methods and object of study of the humanities, favoring measurable results over interpretative analysis.

With the arrival of cybernetics and the computational sciences, literary authors play with the concept of robots in the cybernetic imagination. The novels and stories portray a constant conflict between man and machine. In this sense, Asimov's laws of robots attempt to show that robots can be beneficial to humanity if designed according to a set of rules that ensure peaceful cohabitation. However, as robots progress in design, resembling with ever more fidelity the mind and body of humans, the question of fear of robots becomes ever more complicated as well. It gives rise to other questions, such as "what is reality?" and "what does it mean to be human?" (for example, in the work of Philip K. Dick). While these questions have long been of interest to philosophers and literary authors, they are rephrased in the context of the potential of technology in a century that has experienced two world wars defined by mechanical innovation. Furthermore, with the onset of the ideologies of collectivism, such as National Socialism and Communism, the metaphor of the machine is used to define the state. In this context, the cybernetic imagination portrays the violent contest between the individual and the state where the former is expected to become a functional piece of the grand machine stripped of personal identity, freedom of thought and will. The individual, then, becomes the robot.

In the chapter on *Playing with Reality,* I explored Philip K. Dick's questions in relation to the narratives of replacement. To answer the question of "what is real?" I proceeded from the father of metaphysics, Parmenides, through Descartes and Berkeley, to explore reality as "consensual hallucination" in William Gibson's definition of cyberspace / virtual reality. We have found that reality is a process and rests on perception. Berkeley's principle of *esse est percipi* is actualized in cyberspace, particularly in cyberpunk literature dealing with such themes as virtual reality, artificial memories, the influence of media, computation and so on. This understanding of reality does not dismiss but integrates the theories of digital ontologists who

seek out the physical real. Thus, despite the ontological problems that may result from seeing reality as arising in the co-dependent causality of physical and non-physical elements, the choice is set for the contemporary world, which provides a mixture of digital and physical at every step: digital objects and artificial things can be evaluated through their information properties interconnected with our sense of reality.

A recurring replacement narrative of utmost importance in the cybernetic imagination consists in the contest of domination between man and machine. In the chapter on *Playing with Affection*, I have asked what creates the possibility of acknowledging the consciousness of an intelligent machine, when we cannot be certain that its mental processes resemble human processes and emotions, being emulated rather than spontaneous and genuine. I have concluded that we do not think of other human beings as real; instead, we feel them into existence. This would explain the problem of the "uncanny" as noted by early 20th century psychologists. Our reticence in viewing the android as human originates in our inability to feel the android into human existence. As such, I have looked at literary examples of cybernetic imagination where the human creates emotional connections with artificial humans. The fiction in this field can be organized in three categories, according to the external appearances of the intelligent machine. In the first category, a mental affection is formed with intelligent machines who do not have a physical body but are able to express affection through voice. The second category concerns a partial mental-physical appearance of the intelligent machine, when the human observer is able to identify the android as artificial, but nevertheless develops romantic and sexual feelings towards it. Finally, the third category presents cases when it is almost impossible to distinguish between human and machine.

In all these cases, we have found that the contest consists of manipulating the human observer into feeling the android or replicant as human. Feelings, therefore, influence severely our conception of reality and of real things. In this context, additional problems and questions appear when considering robots utilized in healthcare. I have proposed that the solution is not necessarily to produce a framework to balance feelings and robots, but to rethink the distinction between the real and the unreal. This leads us to the problem of ethics in a computerized world.

If computational technology in all its variants is developed based on the agonistic play of replacement, then what is there to be done and how would one live in such a world? As we can see even in the *Iliad* and *Odyssey*, there are moments in which violent agonism is suspended. This is the case within the culture of computers as well, for example in the hacking "movement" in which the narrative of replacement is being transformed with a sense of exploratory wonder. Cyberpunk literature imagines technological dystopias, reflecting on what to do about them and how to conduct oneself within their parameters. Rather than operating in the same agonistic fashion of returning to an organic world free of technology and proclaiming the superiority of nature over technology, the cyberpunk protagonist appreciates the potential of technologies. Usually, this protagonist is a hacker deploying elements of ingenuity and moral action in cyberspace. In the chapter on *Playing with Computers* I have showcased the nature of the hacker culture and the moral principles it follows.

The play of the hacker generally steps away from violent contest, while his hacks intend to improve a computer system, instead of replacing it. Regarding technology, the hacker sees devices through their potential, rather than their strict design purpose. For him, technology is meant to be played with, to make it do things or serve a purpose not originally intended. Accordingly, hackers between the 1970s and the 1990s not only developed the personal

computer and the software to run it, but also produced software to further the freedoms of the individual and allow their voice to be heard on the global market of ideas.

As we have seen, to play the game for the game itself is another important feature of hacker culture, enabling passionate programming for exploratory innovation. In doing so, the hacker culture does not attempt to control the outcome of innovation or design it with the purpose of financial gain. Most software developed by hackers are open sources and free. Thus, the potential of playful innovation is present only when play is free.

Furthermore, in the chapter on *Playing with Ethics*, I have shown that the hacker culture rests on a series of principles formulated around the idea of the potential good of the computer. In addition, the principles that 'information wants to be free' and that people should have free access to computers, reflect the notion of duty as *oikeiosis*. *Oikeiosis* is perhaps what allows humans to be distinguished from androids in *Do Androids Dream of Electric Sheep?* It is not just empathy, but to care for the well-being of an unknown other. Cyberspace, for the hacker, takes the form of a liminal space able to facilitate the reconfiguration of violent contest towards irenic play. In this way, the hacker culture seems to operate on two major principles: (1) to understand technology for what it does, and not the perception it creates; and (2) *oikeiosis* as the principle of playful innovation.

In turn, I have considered the Quantum Relations Principle as the theoretical framework capable of providing a view of reality beyond violent agonism. QRP sees reality as arising co-dependently in mutually causal relationships among the informational properties of objects, entities and mental phenomena. Thus, it is no longer a problem of contesting one over the other, but rather observing reality as generative rather than static. In cybernetic experiments, QRP

allows the formation of instance-based ethics, that is, the free formation of ethical behavior resulting from mutual causation. We can thus explore with irenic ethics, rather than be confined to our agonistic cultural imperatives.

Finally, my assumption throughout this thesis has been that literature and artistic creation in general can provide the framework to think about and develop an irenic ethics. Here I would like to present several additional illustrations to support this assumption. In *Story of Your Life,* Ted Chiang explores the limits of language in communicating complex ideas. He imagines a situation in which alien visitors try to contact humanity. The heptapods, thus named because their bodies "looked like a barrel suspended at the intersection of seven limbs," speak a language completely unknown to their human hosts. Within their language, they are able to express the past, present, and the future simultaneously: "Humans have developed a sequential mode of awareness, while heptapods have developed a simultaneous mode of awareness. We experience events in an order and perceive their relationship as cause and effect. They experience all events at once, and perceive a purpose underlying them all. A minimizing, maximizing purpose" (Chiang, 2002, p. 97).

In turn, the film adaptation of this story, *Arrival* (2016), explores the mentality outside of replacement and violent contests when governments discover that aliens have promised a "weapon." Government representatives react as expected, worrying that the weapon could be used to destroy the world or, at least, be a serious threat to national security. However, it turns out that the "weapon" is the language itself, a gift for humanity to develop a simultaneous mode of awareness as well. As such, aliens are able to communicate in multiple dimensions, knowing the intent and outcome of the sentence, the past, present and the future. However, even if they know the future, they can be certain of it only if they act it out (Jecan, 2020, p. 102).

The tension between reality and virtual reality is explored in another motion picture, Spielberg's *Ready Player One*. In the film, a virtual reality called Oasis serves as an escape from the real world for many people. The Oasis is on the verge of replacing reality altogether. However, upon his death, the creator left a key for players to discover. The key would allow access to control the entire artificial reality. As it turns out, players around the world compete to find the clues leading to the key. It is only when the protagonist learns how to play the game for the game itself rather than to win (therefore, outside the traditional parameters of violent contest) that he is able to find the clues and control the Oasis. However, instead of shutting it down in favor of physical reality, the protagonist decides on a balance between real and virtual reality. The agonistic feature of play, therefore, is suspended and the contest between the real and the unreal is alleviated (*Ibid.*, p. 103).

As we have seen, narratives of replacement do indeed portray the robot as a metaphor for our own violent play, while the concerns of reality and the ontology of human persons derive from it. Cyberspace, however, provides the platform for ethical A.I. development and playful innovation outside of the cultural imperatives of replacement. It is not an easy journey because at its heart rests the change of mentality of a culture operating on the premises of replacement for at least three millennia. But as rapid developments in computational technologies call for changes of perception and seek out appropriate ethics for innovation, it is possible to reconsider our way forward.

A few positive steps that can be taken result from our study. Current ethical discussions on A.I. ethics and virtual reality try to establish rules for overall innovation. While these may turn out to have some beneficial results, which remain speculative due to the peculiarity of the problem at hand, they are unable to account for specific outcomes of individual usages and

effects. Therefore, it would be useful to move the discussion to the individual and propose a framework derived from the ethics of Stoic metaphysics, hacker culture, and QRP. The first step is to understand technology for what it is and what it does precisely. This principle follows the Stoic materialistic understanding of reality and the concept of "mastery" in hacker culture. In addition, drawing from QRP, we move away from the antagonism of real and unreal objects and see objects in their digital nature through their informational properties.

The second step is the principle of playful innovation on the premise of *oikeiosis*. In this context, playful innovation is regulated not by an external entity, but by extending the good for oneself to the interconnected people in the global technological world. Why should innovation be playful rather than meeting specific economic demands? While the answers can be many and some in accordance with specific scholarly fields, the solution I have offered is in line with Socrates' idea of philosophy as originating in wonder. To wonder how a system works and what it can be turned into is an ethical attitude towards technology itself that minimizes, if even takes into consideration at all, the tension of replacement by intelligent machines. Rather than seeing virtual reality and artificial intelligence as dangerous, seeking to replace the individual in the workplace or in the world, the playful attitude insists, instead, on the positive potential of the situation. In turn, the principle of *oikeiosis* serves as a deterrent to exploitation of technology for personal gain.

Education resting on these two principles must provide an understanding of the irenic alternative. In doing so, it may (1) provide information on the play of replacement to future engineers; (2) instruct on the ethics of the hacker culture; and (3) encourage playful innovation by hands-on experience with technology. Cyberspace can be utilized according to the hacker spirit. As cryptographer Bruce Schneier advised the readers of a young adult hacker novel, *Little*

Brother: “so, close the book and go. The world is full of security systems. Hack one of them” (Schneier, 2008, p. 302). In this sense, the narrative of replacement is changed to the playful pursuit of world-wide beneficial technologies.

Bibliography

Addison, Joseph (1868) *The Works of Joseph Addison*, New York: Harper & Brothers Publishers

Adolphs, Ralph (2005) "Could a Robot have Emotions? Theoretical Perspectives from Social Cognitive Neuroscience," in Michael A. Arbib and Jean-Mark Fellous (eds.), *Who Needs Emotions? The Brain Meets the Robot*, Oxford: Oxford University Press, 2005, pp. 9 – 27

Aristotle (2014) *Complete Works of Aristotle*, Vol. 2, Princeton, NJ: Princeton University Press

Asimov, Isaac (1983) "Runaround" in Isaac Asimov, Patricia S. Warrick, Martin H. Greenberg (eds.), *Machines that Think*, New York: Holt, Rinehart and Winston, pp. 209 - 232

Asimov, Isaac (1983a) "The Evitable Conflict," in Isaac Asimov, Patricia S. Warrick, Martin H. Greenberg (eds.), *Machines that Think*, New York: Holt, Rinehart and Winston, pp. pp. 251 - 277

Asimov, Isaac (1986) *The Dangers of Intelligence and Other Essays*, Boston, MA: Hough Mifflin Company

Isaac Asimov, Patricia S. Warrick, Martin H. Greenberg (1984) *Machines that Think*, New York: Holt, Rinehart and Winston

Bailey, Brian (1998), *The Luddite Rebellion*, Stroud, Gloucestershire: Sutton

Bates, Harry (1984) "Farewell to the Master," in Isaac Asimov, Patricia S. Warrick, Martin H. Greenberg (eds.), *Machines that Think*, New York: Holt, Rinehart and Winston, pp. 93 - 138

Beraldo-de-Araujo, Anderson and Baravalle, Lorenzo (2017) "The Ontology of Digital Physics," in *Erkenn*, No. 82, pp. 1211–1231

Berber, Gregory (2019) "Text-Savvy AI Is Here to Write Fiction, in *Wired*, Nov. 23. Available here: https://www.wired.com/story/nanogenmo-ai-novels-gpt2/

Black, Jeremy (2013), *War and Technology*, Bloomington, IN: Indiana University Press

Blackford, Russell (2004) “Try the *Blue* Pill: What’s Wrong with Life in a Simulation,” in Matthew Kapell and William G. Doty (eds.), *Jacking Into the Matrix Franchise: Cultural Reception and Interpretation*, New York: Continuum, p. 169 – 179

Bloch, Ivan (2017) *Future of War in Its Technical, Economic, and Political Aspects*, London: Forgotten Books

Blumenberg, Hans (1979), *Work on Myth*, Cambridge, MA: The MIT Press

Borup, Mads, Brown, Nick, Konrad, Kornelia, van Lente, Harro (2006) “The Sociology of Expectations in Science and Technology,” in *Technology Analysis & Strategic Management*, Vol. 18, No. 3-4, July-September 2006, pp. 285-298

Bostrom, Nick (2003) “Are You Living in a Computer Simulation?” in *Philosophical Quarterly*, Vol. 53, No. 211, 2003, pp. 243-255

Bradbury, Ray (1993) “The Veldt,” in Karie Jacobson (ed.), *Simulations*, Toronto: Citadel Press, pp. 3 - 20

Braezael, Cynthia and Brooks, Rodney (2005) “Robot Emotion. A Functional Perspective,” in Michael A. Arbib and Jean-Mark Fellous, *Who Needs Emotions? The Brain Meets the Robot*, Oxford: Oxford University Press, 2005, pp. 271 - 310

Brennan, Tad (2005) *The Stoic Life*, Oxford: Oxford University Press,

Brian Inwood (ed.) (2003) *The Cambridge Companion to the Stoics*, Cambridge, UK: Cambridge University Press

Brunschwig, Jacques (2003) Stoic Metaphysics, in Brad Inwood (ed.), *The Cambridge Companion to the Stoics*, Cambridge, UK: Cambridge University Press

Bullock, Seth (2008) “Charles Babbage and the Emergence of Automated Reason,” in Philip Husband, Owen Holland, and Michael Wheeler (eds.), *The Mechanical Mind in History*, Cambridge: The MIT Press, pp. 19 – 40

Cadigan, Pat (1991) *Synners*, New York: Harper Collins

Capek, Karel, *Rossum’s Universal Robots*, Mineola, NY: Dover Publications

Carrere, Emanuel (1993) *I Am Alive and You Are Dead. A Journey Into the Mind of Philip K. Dick*, New York: Metropolitan Books

Carter, Matt (2007) *Minds and Computers*, Edinburgh: Edinburgh University Press

Certeau, Michel de (2000) *The Possession at Loudun*, Chicago: University of Chicago Press

Chaitin, Gregory, (2005) "Epistemology as Information Theory: From Leibniz to Ω," Alan Turing Lecture on Computing and Philosophy, *E-CAP'05, European Computing and Philosophy Conference*, Mälardalen University, Sweden, June 2005. Available here: https://arxiv.org/pdf/math/0506552.pdf

Chiang, Ted (2002) "Story of Your Life," in *Stories of Your Life and Others*, New York:Vintage Books, pp. 91 – 148

Chorzempa, Martin (2018) "China's Social Credit System: A Mark of Progress or a Threat to Privacy?" *Policy Brief 18-14*, Peterson Institute for International Economics

Clarke, Arthur, C. (1961), Dial F for Frankenstein, in *The Wind from the Sun. Stories of Space of the Space Age*, Harcourt Brace Javanovich Inc.

Clarke, Arthur, C. (2000), *Profiles of the Future*, London: Orion Pub Co.

Coleman, Gabriella (2003) *Coding Freedom. The Ethics and Aesthetics of Hacking*, Princeton, NJ, Princeton University Press

Conboy, Martin (2004) *Journalism: A Critical History*, New York: Sage Publications

Cordero, Nestor-Luis (2004) *The Thesis of Parmenides*, Las Vegas: Parmenides Publishing

Coxon, A. H. (1986) *The Fragments of Parmenides*, Van Gorcum

Critical Arts Ensemble (1998) *Flesh Machines: Cyborgs, Designer Babies, and New Eugenic Consciousness*, Autonomedia

Critical Arts Ensemble (2001) *Digital Resistance*, Autonomedia

Critical Arts Ensemble (2018) *Necropolitics, Environmental Struggle*, Autonomedia

David Wall (ed.) (2001) *Crime and the Internet*, London: Routledge

de Solla Price, Derek (1964) “Automata and the Origins of Mechanism and Mechanistic Philosophy,” in *Technology and Culture*, Vol. 5, No. 1, Winter, 1964, pp. 9-23

del Rey, Lester (1977) “Helen O’Loy,” in Mike Ashley (ed.) *Souls in Metal. An Anthology of Robot Futures*, New York: St. Martin’s Press, pp. 19 – 36

del Rey. Lester (1968) “Instinct,” in Robert Silverberg (ed.), *Men and Machines*, New York: Meredith Press, pp. 95 – 112

Derrida, Jacques (2007) “Structure, Sign and Play in the Discourse of the Human Sciences,” in Richard Macksey and Eugenio Donato (eds.), *The Structuralist Controversy. Language of Criticism and the Sciences of Man*, John Hopkins University Press, 2007

Descartes, Rene (1988) *Selected Philosophical Writings,* Cambridge, UK: Cambridge

Deutsch, David (1997) *The Fabric of Reality*, London: Penguin Books

Dick, Philip, K. (1977) “Impostor,” in Mike Ashley (ed.), *Souls in Metal*, St. Martin’s Press, London, pp. 117 - 136

Dick, Philip, K. (2017) *Do Androids Dream of Electric Sheep?,* New York: Del Rey

Dick, Philip, K. (2017a) “The Android and the Human,” in Rob Latham (ed.), *Science Fiction Criticism. An Anthology of Essential Writings,* Bloomsbury, 2017, pp. 295 - 305

Dick, Philip, K. (2017b) “Human Is,” in Philip K. Dick, *Electric Dreams*, Boston: Houghton Mifflin Harcourt, pp. 163 – 180

Dobbs, B.J.T. (1991) “Stoic and Epicurean Doctrines in Newton,” in Margaret J. Osler (ed.), *Atoms*, Pneuma, *and Tranquility*, Cambridge, UK: Cambridge University Press, pp. 221 – 238

Domagalska, Blanka (2014) “Liminality and the Emergence of an Integrated Being,” in *Technoetic Arts: A Journal of Speculative Research*, Vol. 12, no. 2-3, pp. 409-414

Duffy, Jean, H. (2011) *Thresholds of Meaning. Passage, Ritual and Liminality in Contemporary French Narrative*, Liverpool: Liverpool University Press

Durzak, Manfred (2003) “Das Unheimliche in Gustav Meyrinks Roman 'Der Golem',” in Alexandr W. Belobratow (ed.), *Österreichische Literatur und Kultur: Tradition und Rezeption,* Peterburg XXI Vek, pp. 57 – 70

Eagleton, Terry (1976) *Marxism and Literary Criticism*, Berkeley, CA: University of California Press

Echevarria II, Antulio (2007) *Imagining Future War*, Praeger Security International

Ellison, Harlan (1984) "I Have No Mouth, and I Must Scream," in Isaac Asimov *et al.*, *Machines that Think*, New York: Rinehart and Winston, New York, pp. 233 – 250

Emerson, Ralph, Waldo (1849) *Nature; Adresses and Lectures*, James Munroe and Company

Engberg-Pedersen, Troels (1990) *The Stoic Theory of Oikeiosis*, Aarhus: Aarhus University Press

Erskine, Andrew (1992) "Review: Stoic Oikeiosis," in *The Classical Review*, New Series, Vol. 42, No. 1, pp. 77-79

Evans, Arthur, B., (ed.) (2014) *Vintage Visions: Essays on Early Science Fiction*, Middletown, CT: Wesleyan University Press

Even-Zohar, Itamar (1979) "Polysystem Theory," in *Poetics Today*, Vol. 1, No. 1/2, pp. 287-310

Everard, Jeremy (2000) *Virtual States. The Internet and the Boundaries of the Nation-State*, London: Routledge

Finn, Ed (2017) *What Algorithms Want*, Cambridge, MA: The MIT Press

Floridi, Luciano (2009) "Against Digital Ontology," in *Synthese*, Vol. 168, No. 1, May 2009, pp. 151–178

Freenberg, Andrew (1996) "Marcuse or Habermas: Two Critiques of Technology," in *Inquiry*, No. 39, pp. 45 – 70

Fusco, Coco (2003) "Electronic Disturbance. Ricardo Dominguez, Electronic Disturbance Theatre (US), Interviewed by Coco Fusco," in Joanne Richardson (ed.), *Anarchitexts: A Subsol Anthology*, Autonomedia, pp. 105 – 106. The interview is also available online here: http://subsol.c3.hu/subsol_2/contributors2/domingueztext2.html;.

Gibson, William (1981) "The Gernsback Continuum," in Terry Carr (ed.), *Universe 11*, New York: Doubleday & Company, pp. 81 – 90

Gibson, William (2000) *Neuromancer*, New York: Ace Books

Gilbert, Nigel, Troitzsch, Klaus (2005) *Simulation for the Social Scientist*, London: Open University Press

Goldstein, Emmanuel (2008) *2600: A Hacker Odyssey*, Hoboken, NJ: Wiley

Gragido, Will, Pirc, K. Jaishankar (2011) *Cyber Criminology. Exploring Internet Crimes and Criminal Behavior*, Boca Raton, FL: CRC Press

Gray, Colin S. (2004) *Strategy for Chaos: Revolutions in Military Affairs and the Evidence of History*, Londong: Routledge

Green, Lelia (2010) *The Internet: An Introduction to New Media*, Oxford: Berg Publishers

Greene, Preston (2019) "Are We living in a Computer Simulation? Let's Not Find Out," in *The New York Times*, August 16, Available here: https://www.nytimes.com/2019/08/10/opinion/sunday/are-we-living-in-a-computer-simulation-lets-not-find-out.html

Greengard, Samuel (2019) *Virtual Reality*, Cambridge, MA: The MIT Press

Grunkel, David, J. (2001) *Hacking Cyberspace*, Boulder, CO: Westview Press

Gunn, James (2003) "The Reality Paradox in *The Matrix*," in Glenn Yeffeth (ed.), *Taking the Red Pill. Science, Philosophy and Religion in the Matrix*, Dallas: Benbella Books, Dallas, pp. 59 – 70

Guyer, Paul, Horstmann, Rolf-Peter (2015) *Idealism*, Stanford Encyclopedia of Philosophy. Available at: https://plato.stanford.edu/entries/idealism/

Hafele, M., David (2004) *Three Different Shades of Ethical Hacking: Black, White, and Gray*, SANS Institute InfoSec Reading Room

Hafner, Katie (1998) *Where Wizards Stay Up Late: The Origins of the Internet*, New York: Simon & Schuster

Hammes, Thomas, X. (2006) *The Sling and the Stone: On War in the 21st Century*, Minneapolis, MN: Zenith Press

Haraway, Donna (2001) *Simians, Cyborgs and Women: The Reinvention of Nature,* London: Routledge

Harkova, Jana, Kelemen, Jozef (2008) "The Robot Story: Why Robots Were Born and How They Grew Up," in Philip Husbands, Owen Holland, and Michael Wheeler (eds.), *The Mechanical Mind in History,* Cambridge, MA: The MIT Press, p. 283 – 306

Heidegger, Martin (1977) *The Question of Technology and Other Essays,* New York: Harper Torchbooks

Heim, Michael (1994) *The Metaphysics of Virtual Reality,* Oxford: Oxford University Press

Himanen, Pekka (2001) *The Hacker Ethic,* London: Random House, 2001

Hoffmann, E.T.A. (1982) The Sandman, in Victor Lange (ed.), *Tales,* New York: Continuum

Huizinga, Johan (2016) *Homo Ludens. A Study of the Play-Element in Culture,* Brooklyn, NY: Angelico Press

Hyde, Lewis (1998) *Trickster Makes This World. Mischief, Myth, and Art,* New York: North Point Press

Idel, Moshe (2019) *Golem. Jewish Magical and Mystical Traditions of the Artificial Anthropoid,* Brooklyn, NY: Ktav Publishing House

Jecan, Vlad (2011) "Hacking Hollywood: Discussing the Hacker Community Reactions to Three Popular Films,) in *Journal of Media Research,* Vol. 2, No. 10, 2011, pp. 95 – 114

Jecan, Vlad, Meza Radu (2013) "Concept Mapping of Ideological Positioning in Cultural and Political Periodicals in the Interbellum Cluj," in *Romanian Journal of Information Science and Technology,* vol. 16, no. 2-3, 2013, pp. 237-250

Jecan, Vlad (2016) *The Play of Hacking and the Political Values of the Hacker Culture,* Unpublished Doctoral Dissertation, Cluj-Napoca: University of Babes-Bolyai

Jecan, Vlad, Meza, Radu (2017) "Co-Citation Mapping and the Intercultural Dialogue of the Intellectual Communities in Arad and Timisoara (19th to early 20th centuries)," in Spariosu,

Mihai, I., (ed.), *Intercultural Conflict and Harmony in Central European Borderlands. The Cases of Banat and Transylvania 1849 – 1939*, Göttingen: V&R unipress, pp. 355 – 367

Jecan, Vlad (2020) "Towards Irenic Mixed Reality: An Experiment in Play and the Cybernetic Imagination," in Dorothy M. Figueira (ed.), *Rebuilding A Profession: Comparative Literature, Intercultural Studies and the Humanities in the Age of Globalization*, Göttingen: V&R unipress, pp. 91 – 106

Jentsch, Ernst (1906) "Zur Psychologie des Unheimlichen," in *Psychiatrisch-Neurologische Wochenschrift*, Vol. 8, No. 22, pp. 195 – 205

Jordan, Tim (2013) *Hacking. Digital Media and Technological Determinism*, Cambridge, UK: Polity

Jordan, Tim (2015) *Information Politics: Liberation and Exploitation in the Digital Society*, London: Pluto Press

Jordan, Tim, Taylor, Paul, A. (2004) *Hacktivism and Cyberwars. Rebels with a cause*, London: Routledge

Kadar, Zoltan, Toth, Janos (2013), "The Critique of Technology in 20th Century Philosophy and Dystopias," in *Procedia: Social and Behavioral Sciences*, Vol. 71, 2013, pp. 53 – 60

Kahn, David (1996) *The Code Breakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*, New York: Scribner

Kaplan, David, M. (ed.) (2009) *Readings in the Philosophy of Technology*, Lanham, MD: Rowman & Littlefield Publishers

Karie Jacobson (ed.) (1993) *Simulations*, Toronto: Citadel Press

Kazemi, Darius, (2013) *Teens Wander Around a House*, Available here: http://tinysubversions.com/nanogenmo/novel-2.pdf

Keohane, Robert, O., Nye, Joseph S., Jr (2006). *Power and interdependence and the information age*, in Richard Little and Michael Smith (eds.), *Perspectives on World Politics*, London: Routledge, pp. 207 – 216

Kidder, Tracy (2000) *The Soul of a Machine*, New York: Back Bay Books, 2000

Kingsley, Peter (1999) *In the Dark Places of Wisdom*, Point Reyers, CA: The Golden Sufi Center

Kingsley, Peter (2003) *Reality*, Point Reyers, CA: The Golden Sufi Center

Klaus, Eric (2010) "Allegorical Slumber: Somnambulism and Salvation in Gustav Meyrink's Der Golem," in *Seminar: A Journal of Germanic Studies*; Vol. 46, No. 2, pp. 131 – 145

Konrad Zuse (1967) "Rechnender Raum," in *Elektronische Datenverarbeitung*, vol. 8, pp. 336-344

Kostka, Genia (2019) "China's Social Credit Systems and Public Opinion: Explaining High Levels of Approval," in *New Media & Society*, Vol. 21, No. 7, pp. 1565–1593

Krieger, Arndt (1996) "Wege der Erkenntnis : in Gustav Meyrinks Roman 'Der Golem' und Franz Kafkas Erzählung 'Die Verwandlung'," in *Brücken. Germanistisches Jahrbuch Tschechien – Slowakei*, Vol.6, No. 1-2, pp. 153 – 174

Kuberski, Philip (1994) *Chaosmos. Literature, Science, and Theory*, New York: State University of New York Press

Kurzweil, Ray (2009) "Superintelligence and Singularity," in Susan Schneider (ed.), Science Fiction and Philosophy: From Time Travel to Superintelligence, Hoboken, NJ: Blackwell Publishing, Chichester, pp. 201 - 224

Levy, Steven (2010) *Hackers. Heroes of the Computer Revolution*, Sebastopol, CA: O'Reilly Media

Link, Eric, Carl (2010) *Understanding Philip K. Dick*, Columbia, SC: The University of South Carolina Press

Lozano-Lopez, Miguel (2007) *Utopian Dreams, Apocalyptic Nightmares: Globalization in Recent Mexican and Chicano Narrative*, West Lafayette, IN: Purdue University Press

Mackey, Douglas (1988) A., *Philip K. Dick*, Woodbridge, CT: Twayne Publishers, 1988

Manovich, Lev (2016) "The Science of Culture? Social Computing, Digital Humanities and Cultural Analytics," in *Journal of Cultural Analytics,* May 23. Available here: https://culturalanalytics.org/article/11060

Marcuse, Herbert (1991) *One-Dimensional Man,* Boston, MA: Beacon Press

María Belén (2013) "Der Raum in Der Golem von Gustav Meyrink: Transgression der Gothic Novel," in *Estudios Filológicos Alemanes,* No. 26, pp. 191 – 198

Mark Zunac (ed.) (2016) *Literature and the Conservative Ideal,* Washington, DC: Lexington Books

Mayor, Adrienne (2018) *Gods and Robots. Myths, Machines, and Ancient Dreams of Technology*, Princeton, NJ: Princeton University Press

McLuhan, Marshall (1994) *Understanding Media: The Extensions of Man,* Cambridge, MA: The MIT Press

Meacham, Darian, Studley, Matthew (2017) "Could a Robot Care? It's All in the Movement," in Patrick Lin, Ryan Jenkins, and Keith Abney (eds.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, Oxford, UK: Oxford: Oxford University Press, pp. 97 - 112

Melville, Herman (1855) "The Bell-Tower," in *Putnam's Magazine: Original Papers on Literature, Science, Art and National Interests,* Volume 6, December 30

Meyrink, Gustav (1985) *The Golem,* Mineola, NY: Dover Publications

Mitnick, Kevin (2012) *Ghost in the Wires*, New York: Back Bay Books

Mitnick, Kevin, D., Simon, William, L. (2005) *The Art of Intrusion,* Hoboken, NJ: Wiley., 2005

Montfort, Nick (2015) "Computational Literature," in Patrik Svensson and David Theo Goldberg (eds.), *Between Humanities and the Digital*, Cambridge, MA: The MIT Press, pp. 205 – 216

Moretti, Franco (2007) *Graphs, Maps, Trees: Abstract Models for Literary History,* London: Verso

Moretti, Franco (2013) *Distant Reading*, London: Verso

Mourelatos, Alexander P. D. (1970) *The Route of Parmenides*, New Haven, CT: Yale University Press

Nan Z. Da (2019) "The Digital Humanities Debacle," in *The Chronicle for Higher Education,* March 27. Available here: https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986

Nan Z. Da (2019a) "The Computational Case against Computational Literary Studies," *Critical Inquiry,* Vol. 45, No. 3, pp. 601-639

Nathan Schwarts-Salant and Murray Stein (eds.) (1991), *Liminality and Transitional Phenomena*, Asheville, NC: Chiron Publications

Nicolae Jurau, Two Specialists in Cybernetics: Stefan Odobleja and Norbert Weiner. Common and Different Features, available here: http://www.bu.edu/wcp/Papers/Comp/CompJurc.htm

Nietzsche, Friedrich (1974) *The Gay Science*, New York: Knopf Publications

Nozick, Robert (2013) *Anarchy, State, and Utopia*, Philadelphia: Basic Books

O'Malley, George (2013) "Hacktivism: Cyber Activism or Cyber Crime?," in *Trinity College Law Review*, Vol. 16

Odobleja, Ştefan (1938-1939) *Psychologie consonantiste*, Paris: Librairie Maloine

OpenAI, (2019) "Better Language Models and Their Implications," Available here: https://openai.com/blog/better-language-models

Orwell, George (1946) "Review of 'WE' by E.I. Zamyatin," in *Tribune*, January 4, 1946. Available here: http://orwell.ru/library/reviews/zamyatin/english/e_zamy

Osler, Margaret J. (ed.) (1991) *Atoms,* Pneuma*, and Tranquility,* Cambridge, UK: Cambridge University Press,

Pastuszka, Anna (1996) "Einige Bemerkungen zu Raumdarstellung und -deutung in 'Golem' von Gustav Meyrink," in *Lubelskie Materiały Neofilologiczne,* No. 20, pp. 57 – 73

Paul di Filippo (ed.) (2017) *Critical Survey of Science Fiction and Fantasy Literature,* Pasadena, CA: Salem Press, 2017

Picard, Rosalind, W. (1997) *Affective Computing*, Cambridge, MA: The MIT Press

Plato, (1992) *Republic*, Indianapolis, IN: Hackett Publishing Company, Grube, G.M.A. (trans.)

Plato, (2000) *Timaeus*, Indianapolis, IN: Hackett Publishing Company, Zeyl, Donald J. (trans.)

Ramsay, Stephen (2013) "Who's In and Who's Out," in Melissa Terras, Julianne Nyhan, Edward Vanhoutte (eds.), *Defining Digital Humanities*, Farnham, UK: Ashgate, pp. 243 – 247

Randford, Alec, et. Al (2019) "Language Models are Unsupervised Multitask Learners," OpenAI, 2019. Available here: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

Ravetto-Biagioli, Kriss (2019) *Digital Uncanny*, Oxford: Oxford University Press

Rayner, Tim (2018) *Hacker Culture and the new Rules for Innovation*, London: Routledge

Richard Brown and Kevin S. Decker (eds.) (2009) *Terminator and Philosophy*, Hoboken, NJ: Wiley

Richard J. Ketchum (1990) "Parmenides on what There Is," in *Canadian Journal of Philosophy*, Vol. 20, No. 2, pp. 167 – 190

Ricoeur, Paul (1976, *Interpretation Theory. Discourse and the Surplus of Meaning*, Fort Worth, TX: TCU Press

Rid, Thomas (2016) *Rise of the Machines: A Cybernetic History*, New York: W. W. Norton & Company

Roberts, Adam (2016) *The History of Science Fiction*, London: Palgrave Macmillan

Robertson, Donald (2019) *The Philosophy of Cognitive-Behavioural Therapy*, London: Routledge

Rolls, Edmund, T., (2005) "What are Emotions. Why do We have Emotions, and What is Their Computational Basis in the Brain?" In Michael A. Arbib and Jean-Mark Fellous, *Who Needs Emotions? The Brain Meets the Robot*, Oxford: Oxford University Press, pp. 117 - 146

Rubin, Charles, T. (2019) "The Ancients' Tech Anxiety. On the Shallowness of Reading Mythology as Sci-Fi," in *The New Atlantis*, Fall 2019, pp. 80 – 86

Ryan, Johnny (2013) *A History of the Internet and the Digital Future*, London: Reaktion Books

Sambursky, Samuel (1987) *Physics of the Stoics*, Princeton, NJ: Princeton University Press

Scharre, Paul (2019) *Army of None: Autonomous Weapons and the Future of War,* New York: W. W. Norton & Company

Schell, Bernadette, H., Dodge, John, L. (2002) *The Hacking of America,* Quorum Books, 2002

Scheub, Harold (2012) *Trickster and Hero. Two Characters in the Oral and Written Traditions of the World*, Madison, WI: University of Wisconsin Press

Schloer, Hardy, Spariosu, Mihai (2016) *The Quantum Relations Principle,* Göttingen: V&R unipress

Schneier, Bruce (2008) "Afterword to Little Brother," in Cory Doctorow, *Little Brother,* Feedbooks

Sebag-Montefiore, Hugh (2004) *Enigma: The Battle for the Code*, Hoboken, NJ: Wiley

Seneca (1925) *Seneca's Letters from a Stoic,* New York: G. P. Putnam's Sons, 1925, Gummere, Richard Mott (trans.)

Sharkey, Noel, Sharkey, Amanda (2010) "Living with Robots: Ethical Tradeoffs in Eldercare," in Yorick Wilks (ed.), *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues,* John Benjamins, 2010

Shelley, Mary (2012) *Frankenstein*, New York: W.W. Norton & Company, 2012

Shimomura, Tsutomu, Markoff, John (1996), *Takedown: The Pursuit and Capture of Kevin Mitnick, America's Most Wanted Computer Outlaw - By the Man Who Did It*, New York:Voice

Siemens. Ray (ed.) (2013) *A Companion to Digital Literary Studies,* Hoboken, NJ: Wiley-Blackwell, 2013

Simon, Herbert, A. (1997) *The Sciences of the Artificial*, Cambridge, MA: MIT Press

Singer, Peter, W. (2009) *Wired for War: The Robotic Revolution and Conflict in the 21st Century*, London: Penguin Books

Singer, Peter, W. (2016) *Ghost Fleet: A Novel of the Next World War*, Boston, MA: Mariner Books

Singer, Ryan (2010) "Richard Clarke's Cyberwar: File under Fiction," in *Wired Magazine*, April 22, 2010. Available here: https://www.wired.com/2010/04/cyberwar-richard-clarke/

Sithigh, Daithi, Mac, Siems, Mathias (2019) *The Chinese Social Credit System: A Model for Other Countries?*, Fiesole FI, Italy: European University Institute

Skolnick, Evan (2014) *Video Game Storytelling*, New York: Watson-Guptill

Smart Bell, Madison (2000) *Narrative Design. Working with Imagination, Craft, and Form*, New York: W.W. Norton

Spariosu, Mihai, I. (1989) *Dionysus Reborn. Play and the Aesthetic Dimension in Modern Philosophical and Scientific Discourse*, Ithaca, NY: Cornell University Press

Spariosu, Mihai, I. (1991) *God of Many Names. Play, Poetry and Power in Hellenic Thought from Homer to Aristotle*, Durham, NC: Duke University Press

Spariosu, Mihai, I. (1997) *The Wreath of Wild Olive*, New York: State University of New York Press

Spariosu, Mihai, I. (2015) *Modernism and Exile. Play, Liminality, and the Exilic-Utopian Imagination*, London: Palgrave Macmillan

Spariosu, Mihai, I. (ed.) (2017) I*ntercultural Conflict and Harmony in the Central European Borderlands: The Cases of Banat and Transylvania 1849-1939*, Göttingen: V&R unipress

Standage, Tom (1998) *The Victorian Internet*, London: Bloomsbury

Stefan, Gheorghe, M. (1981) "The Cybernetic View of Stefan Odobleja," in Mihai Draganescu (ed.), *Odobleja: between Ampère and Wiener*, Bucharest: Editura Academiei Romane

Steinmetz, Kevin, F. (2016) *Hacked. A Radical Approach to Hacker Culture and Crime*, New York: New York University Press

Sterling, Bruce (1992) *The Hacker Crackdown*, New York: Bantam Books

Steven K. Strange and Jack Zupko (eds.) (2004) *Stoicism*, Cambridge University Press

Stimson, Frederic, Jesup (1995) D. Materialismus, in H. Bruce Franklin (ed.), *Future Perfect*, Rutgers University Press

Strickland, Lloyd (2014) *Leibniz's Monadology*, Edinburgh: Edinburgh University Press

Suits, Bernard (1988) "Words on Play," in W. J. Morgan and K.V. Meier (ed.) *Philosophic Inquiry in Sport*, Champaign, pp. 117 – 131

Svensson, Patrik (2013) "Humanities Computing as Digital Humanities," in Melissa Terras, Julianne Nyhan, Edward Vanhoutte, (eds.), *Defining Digital Humanities*, Farnham, UK: Ashgate, pp. 159 – 186

Taran, L. (1965) *Parmenides*, Princeton University Press, 1965

Taylor, Paul, A. (1999) *Hackers. Crime in the digital sublime*, London: Routledge, 1999

Taylor, Paul, A. (2005) From Hackers to Hacktivists: Speed Bumps on the Global Superhighway? In *New Media & Society*, Vol. 7, No. 5

Tegmark, Max (2017) *Life 3.0. Being Human in the Age of Artificial Intelligence*, New York: Knopf

Teschner, George, Grace, Patrick (2011) "Human or Machine, Does it Mind or Matter?" In D. E. Wittkower (ed.), *Philip K. Dick and Philosophy*, Chicago, IL: Open Court, pp. 89 – 98

Thomas, Douglas (1998) *Hacking L.A.: Exploring Los Angeles' Digital Underground,* Annenberg School for Communication, University of Southern California, 1998, p. 3 Available here: http://www.usc.edu/dept/LAS/SC2/pdf/thomas.pdf

Thomas, Douglas (2002) *Hacker Culture*, Minneapolis, MN: University of Minnesota Press

Togelius, Julian (2018) *Playing Smart. On Games, Intelligence, and Artificial Intelligence*, Cambridge, MA: The MIT Press

Truitt, E. R. (2015) *Medieval Robots. Mechanism, Magic, Nature, and Art*, Philadelphia, PA: University of Pennsylvania Press

Turgeman-Goldschmidt, Orly (2011) “Identity Construction Among Hackers,” in K. Jaishankar, *Cyber Criminology. Exploring Internet Crimes and Criminal Behavior*, Boca Raton, FL: CRC Press

Turing, Alan (1950) “Computing Machinery and Intelligence,” in *Mind*, No. 49, pp. 433 – 460

Turkle, Sherry (2005) *The Second Self: Computers and the Human Spirit*, Cambridge, MA: The MIT Press

Turner, Victor (1979) *Process, Performance, and Pilgrimate: A Study in Comparative Symbology*, Concept Publishing

Turner, Victor (1995) *The Ritual Process: Structure and Anti-Structure*, Aldine de Gruyter

Vaingurt, Julia (2012) “Human Machines and the Pains of Penmanship in Yevgeny Zamyatin’s *We*,” in *Cultural Critique*, No. 80, pp. 108 – 130

van Gennep, Arnold (2019) *The Rites of Passage*, University of Chicago Press, Second Edition

Vinge, Vernor (2017) The Coming Technological Singularity: How to Survive in a Post-Human Era, in Rob Lantham (ed.), *Science Fiction Criticism*, London: Bloomsbury 2017

Waddell, Terrie (2010) *Wild/Lives. Trickster, Place and Liminality on Screen*, London: Routledge

Walter, Gerard (1953) *Caesar: A Biography*, New York: Cassell

Warrick, Patricia, S. (1980) *The Cybernetic Imagination in Science Fiction*, Cambridge, MA: The MIT Press

Wheeler, Michael (2008) “God’s Machines: Descartes on the Mechanization of Mind,” in Philip Husband, Owen Holland, and Michael Wheeler (eds.), *The Mechanical Mind in History*, Cambridge, MA: The MIT Press

Wiener, Norbert (1956) *The Human Use of Human Beings: Cybernetics and Society*, Doubleday Anchor Books

Wright, Robert (2001) “The Man Who Invented the Web,” in *Time*, June 24, 2001. Available here: http://content.time.com/time/magazine/article/0,9171,137689,00.html

Wrobel, Dieter (2010) “Gustav Meyrink, 'Der Golem': zwischen Traum und Wissen, zwischen Vision und Realität,” in Wrobel, Dieter, *Vergessene Texte der Moderne: Wiederentdeckungen für den Literaturunterricht,* Wissenschaftlicher Verlag Trier, pp. 49 – 63

Xiuli, Jin (2012) “Phantasie und Wirklichkeit : Gustav Meyrinks Traummotiv im Roman 'Golem',” in *Literaturstraße : chinesisch-deutsche Zeitschrift für Sprach- und Literaturwissenschaft,* Band 13, pp. 151-172

Yannaras, Christos (2015) *The Schism in Philosophy*, Brookline, MA: Holy Cross Orthodox Press

Yvonne Jewkes (ed.) (2007) *Crime Online,* London: Willan Publishing

Zamyatin, Yuvgeny (1952) *We*, New York: E. P. Dutton