2

The Stupidity of Intelligence

As soon as intelligence takes itself as its object, it is destined to transform into stupidity, either as the g factor or as intellect. If the psychologists' intelligence is stupid, then, in the end, that of the philosophers may be equally so. The philosophical self-assertion of the mind, claiming the sovereignty of the mind or intellect, always seems to result in a ridiculous form of celebration of the self that is no better than the reductionism of psychologists.
—Catherine Malabou *(2019, pp. 51–52)*

2.1 AI: Failure, Trauma, Dupery

The dictionary definition of intelligence, from the Latin 'intus' (between) and 'legere' (to choose), entails variously the ability to discern, make decisions, comprehend, to have skill, art, taste and ultimately knowledge.[1] But the term 'artificial' preceding 'intelligence' gives one the immediate

[1] Indeed, the question of the genealogy of the concept of intelligence has a rich history in the fields of philosophy, psychology and cognitive science, the literature on which I shall not delve into here. Suffice it to say, once we interrogate the concept of intelligence, very quickly, we are in the realms of philosophy and the humanities in general and not science (much less computer science) *stricto sensu*.

I. Millar, *The Psychoanalysis of Artificial Intelligence*, The Palgrave Lacan Series,
https://doi.org/10.1007/978-3-030-67981-1_2

impression that we have a sound definition of what real intelligence is; a conceit that in philosophical terms elides *technē* and *epistēmē* and in psychoanalytic terms takes for granted the distinct categories of knowledge, truth, and most significantly, enjoyment.

Human intelligence entails various interactions between different skills, for example, a combination and interaction of visual perception, motor skills, memory, speech, spatial reasoning, auditory processing may be utilised at any given moment. These skills are of course not all transparently understandable to the 'intelligent human' utilizing them. This is the paradox at the heart of debates between neuroscience and philosophical accounts of consciousness[2] which start from fundamentally different premises on how we may talk about subjective phenomena. For example, at its crudest level, just because you can see, it does not mean that you 'know' how vision works. And conversely, knowing how vision works does not guarantee that you will be able to see. This same sort of combination of functions will be present in any complex AI program, which will have integrated elements the core processor knows how to access. This may include evidence-based reasoning, language skills, text analysis, sensors, decision making, data analysis and so on.

For example, can a program that processes visual information be said to understand how vision works? Or does facial recognition software operate in the same intuitive way for computers as it does for humans? According to some computer scientists the answer is yes, computers like humans have "intuition" (Hammond 2018). So, in very basic terms we can see that there is a discrepancy between technical or machinic capacity and theoretical knowledge, which is integral to the concept of Artificial Intelligence as it is commonly used. However, as Lydia Liu (2010) notes, it is one thing to argue that the human brain can be augmented by the computer chip and another to imagine that the logic of the computer and

[2] Classically exemplified by John Searle's (1980) *Chinese Room Argument*; a thought experiment in which a hypothetical language processing machine would be constructed by a non-Chinese speaking person inside a room manipulating Chinese characters according to a set of rules and outputting them to another Chinese-speaking person outside the room. Searle used the example to critique Strong AI on the basis that the person inside could not be said to understand Chinese despite being able to convince its interlocuter otherwise. The argument has been refuted since in various ways, objecting that metaphorically speaking the person inside the room may logically correspond to a part of the brain but not consciousness *per se*.

the human psyche are analogous, this prosthetic argument, as she calls it, is an 'alibi for something more fundamental that has been going on since the mid-twentieth century, and this is the cybernetic conception of the human psyche as a computing machine' (p. 8). What remains to date the pervasive paradigm for thinking about Artificial Intelligence is the relationship between the computer and the brain.[3] But as we shall see this specific focus on the analogy of the brain and the computer leaves important psychoanalytic questions unbroached.

In recent times, the field of Artificial Intelligence has grown exponentially with a multitude of technical approaches provoking interest from various philosophical perspectives surrounding its conceptual ramifications. Today, AI encapsulates a huge range of phenomena including those that fall under goal-oriented "narrow" AI or Artificial Narrow Intelligence (ANI), which perform limited tasks such as sorting, tracking, predicting or recognising data patterns based on a range of different algorithms. These may be used in applications like Google Search and Amazon's Alexa, or more sophisticated AI in the development of autonomous vehicles or AVs (Apple's Project Titan and Tesla's Autopilot), preventative medicine (Microsoft AI) and, more controversially, autonomous weapon systems. At the other end of the spectrum, however, one finds robotics, deep learning and neural networks, which find a more complex outlet in simulations and perhaps most intriguingly, the Blue Brain Project as we shall discuss.

A recent volume of essays *Alleys of Your Mind: Augmented Intelligence and its Traumas* tackles the question of Artificial Intelligence from the perspective of contemporary critical theory.[4] The collection gathers contributions from leading thinkers who all in diverse ways criticise popular conceptions of AI and ask the questions; what thinking means in the age of Artificial Intelligence, and how does big-scale computation transform the ways in which our brains function. The book's main conceit is to

[3] The equation, or rather the evolutionary development of the human brain and the computer in neuroscientific research and computer science first emerged as a result of John Von Neumann's (2012) contributions to the fields of information technology and computation, best captured in his 1958 *The Computer & The Brain*. See also Kurzweil (2012).

[4] The authors gathered in this collection are predominately philosophers or new media theorists but include also historians of science and art theorists.

uncover the 'positive role played by error and trauma in the construction of our contemporary technological minds' (Pasquinelli 2015).

In his contribution to the volume, Matteo Pasquinelli (2015) argues that current philosophical debates on the question of technology and reason fall between, on the one hand, neomaterialism and, on the other, neorationalism (or in other words, positions that draw on either Whitehead or Sellars). The former being a proponent of the 'agency of technical objects, matter and affects' and the latter addressing rather the 'primacy of reason and its potential forms of autonomization' (p. 8), or the ability for synthetic reason to become autonomous. Pasquinelli sees this as the wrong distinction to be making however, and believes instead that no paradigm of cognition and computation can be assessed without the recognition of the 'epistemic abnormal' or what he calls 'noetic failure' which we could translate as mental or intellectual error (p. 8). This would mean a distinction should be made between philosophies that acknowledge a positive role for 'error, abnormality, pathology, trauma and catastrophe' and those who support a flat ontology without these 'dynamic, self-organizing and constitutive ruptures on the other' (ibid.). Following the Frankfurt School's lessons on the trauma of reason, Pasquinelli asserts that the *reason of trauma* must be 'rediscovered as the actual inner logic of the age of intelligent machines' (ibid.). In his introduction to the volume, he asserts that:

> One day, it will not be arbitrary to reframe twentieth century thought and its intelligent machines as a quest for the positive definition of error, abnormality, trauma and catastrophe—a set of concepts that need to be understood in their cognitive, technological and political composition. (p. 7)

Drawing parallels with Foucault's history of biopower and technologies of the self as sharing common roots with cybernetics and its 'error-friendly machines' (p. 7) and arguing that Deleuze and Guattari's desiring machines were in fact echoing research on war trauma and brain plasticity from the First World War, he states that:

> Across the history of computation (from early cybernetics to Artificial Intelligence and current algorithmic capitalism) both mainstream technol-

> ogy and critical responses to it have shared a common belief in the determinism and positivism of the *instrumental* or *technological rationality* to use the formulation of The Frankfurt School. (Ibid.)

The aim of the collection therefore is to highlight conversely the role of 'error, trauma and catastrophe in the design of intelligent machines and the theory of augmented cognition' (p. 7). Pasquinelli argues that the definition of intelligence remains an open problem since from a philosophical point of view, human intelligence is always artificial in the first place, engendering as it does novel dimensions of cognition (ibid.). Intelligence is a combination of various complex and multifaceted capacities, which defy easy definition. Pasquinelli discerns three major "fallacies" in the current debates on the state of Artificial Intelligence. Firstly, the anthropocentric fallacy, which naively assumes Artificial Intelligence would be analogous to human intelligence in the superficial sense of attributing menacing and threatening motives to AI. In this vision of AI, it is a malevolent predator whose aim is to expunge us or at the very least make us suffer. Secondly, the bootstrapping fallacy, which imagines a seamless exponential growth of machine intelligence similar to the progress of human psychological development, in the sense that the relative and progressive complexity of cognitive tasks for humans is directly mapped onto the progress of machinic intelligence. This fallacy errs when it fails to reckon with the different forms of cognitive capability that human intelligence entails, and the ways in which some processes that are very difficult for humans are easily replicable by algorithms while other seemingly simple human tasks are hugely complex engineering feats for AI. And finally, the third problem is the Singularity fallacy, which, combining elements of the previous two problems, entails the belief that there will be a decisive point of unification and synchronisation of different technologies simulating, augmenting and ultimately surpassing human intelligence which taken together form a homogenous and all-powerful mode of thought, capable of rendering human intelligence (and, by extension, the species in general) redundant.

Pasquinelli's criticisms of the current fallacies surrounding conceptions of AI and of the need to approach AI from the basis of error, trauma, and catastrophe, points to the implicitly psychoanalytic dimensions and

implications of AI. Firstly, the issue Pasquinelli points to with imaginary identification, which seems to be pervasive in our misapprehensions of Artificial Intelligence. This entails the assumption that the ability to communicate symbolically entails an element of intersubjectivity, which is fundamentally misguided. In Lacanian terms this relates to the question of misrecognition or imaginary *méconnaisance* and, ultimately, transference. Secondly, we should consider more closely the characterization of the concept of trauma and pathology as the unacknowledged ontologically significant factor in the positivist project of AI research. This is the concern that the whole of the Freudian enterprise was built on; the concept of trauma and the re-formulation of what was previously conceived as pathology as inherent to psychic structure. This was not least because it was via trauma and the symptom that Freud found the way into deciphering his very first patient's hysterical symptoms, and neither merely in the sense of the discovery of an unfortunate and disturbing event, but rather that trauma, as Freud described it (and Lacan formalised it), was structurally constitutive of the psychoanalytic subject. The subject, in psychoanalytic terms, is characterised by a constitutive failure.

Benjamin Bratton (2015) likewise warns that contemporary polemics around the meaning of thinking and living with radically other forms of synthetic intelligence crucially misunderstand the real issues at stake, stuck as they are in anthropocentric appreciations of AI. As far as Bratton is concerned, we should resist the temptation to understand AI through the lens of human intelligence, for to do so would be 'self-defeating, unethical and perhaps even dangerous' (p. 70). For this reason, he advocates a broadening up of the concept of intelligence to situate human intelligence as just one form of intelligence along a larger continuum. In this way he promotes a form of "anti-bigotry" in our understanding of what may count as intelligent thought; a step towards a better appreciation of the challenges involved in living and thinking with the Other of Artificial Intelligence.

Bratton suggests that our fantasies of AI as either desperately wishing to be human e.g. Steven Spielberg's *A.I.* (2001) or Chris Columbus' *Bicentennial* (1999) *Man* or malevolently fixated on our destruction e.g. James Cameron's (1984) *The Terminator* (1984) or more recently, various episodes of Charlie Brooker's series *Black Mirror* merely reflect our own

desires, paranoia and narcissistic self-image. That AI would be so invested in getting our attention is perhaps pure wishful thinking; the worst-case scenario may well be that they wouldn't even notice us at all:

> the real nightmare, even worse that the one in which the Big machine wants to kill you, is the one in which it sees you as irrelevant, or not even as a discrete thing to know. Worse than being seen as an enemy, is not being seen at all. (p. 70)

The idea of Artificial Intelligence being perceived as like humans, Bratton says, is a valid point of departure, but not a valid conclusion to end up with. Furthermore an important distinction Bratton makes in AI research is between the terms artificial stupidity and artificial idiocy, where the first term refers to faults deliberately programmed into AI in order to make it more realistically human (i.e. not winning every time in a game in which it could very easily beat a human every time). The second term refers to the problem that occurs when an AI performs its tasks too well, to the detriment of other factors. The apocalyptic example used originally by Nick Bostrom (2014), is the paperclip generator that follows its orders to continue making paperclips until the world is overrun by an avalanche of paper clips. This is the idiocy of AI so called. In his first example, AI stupidity has the function of creating a social bond between itself and its human companion. AI idiocy on the other hand has the aim of following its master's instructions so "ideologically" that it allows all other factors to pail into insignificance. So here the blind spot functions in different places. Or one could say the enjoyment produced is different in each case. The stupidity or idiocy of AI as Bratton calls it may therefore be seen as a discursive structure, in which enjoyment is produced according the relative positions of truth and knowledge. The concept of stupidity therefore is one which has fundamental structural importance in relation to the question of enjoyment.

Bratton also draws attention to what he sees as the inherently 'bigoted' approach humans have to AI, since as he (2015) puts it, the Turing Test's aim was to fool an interlocutor that an AI was human. So the AI had to be 'in drag' (p. 76). But as Bratton remarks, the thing about actual drag is, you are not supposed to convince someone you are a member of the

opposite sex, but merely get them to suspend their belief about your sexual or gender identity (in the disavowed form of 'I know very well, but…'). This he discusses with relation to the scandal surrounding Turing's outing as a homosexual man (illegal at the time) and his subsequent chemical castration, causing Turing an unspeakable suffering which in the end led to him commit suicide. Given that Turing's Imitation Game was based on the logic of tricking an interlocutor into believing that you were a particular gender (at which the AI had to perform as well as the human) the parallels are significant.

> One notes the sour ironic correspondence between asking an AI to pass the test in order to qualify as intelligent—to pass as human intelligence—with Turing's own need to hide his sexuality and to pass as a straight man. The demands of both bluffs are unnecessary and profoundly unfair. (p. 72)

However, the question of the relationship between sexual identity and AI is in fact far more psychoanalytically complex than a political irony about the potential problems with anthropocentrism in AI research as Bratton suggests. As we shall discus further in Chap. 5, the Turing Test, when closely examined, demonstrates the fundamental basis for the relationship between sexuation in psychoanalytic terms and Artificial Intelligence.

In essence, Bratton's arguably utopian message is that instead of forcing AI into prefigured stereotypes about the type of thinking that we may recognize as "real" we should allow AI in all its diverse forms, about which we are continually learning, to teach humans a more 'fuller and truer range of what thinking can be' (ibid., p. 72). Diversifying what counts as "thought" is in keeping with the realist turn in contemporary theory and offers us an important perspective in relation to pushing psychoanalytic paradigms beyond post-structuralist tendencies. But furthermore we should pay attention to the fact that (as Bratton himself points out in his example of AI stupidity) there is a constitutive part of "thought" that involves a dialectical relationship between the positions of truth and knowledge in any given discursive framework.

2.2 Omega Numbers and Suture

Let us press further the question of thinking in philosophical approaches to AI. In her essay '*Instrumental Reason, Algorithmic Capitalism and the Incomputable*' Luciana Parisi (2015) asks:

> Can the critique of instrumental rationality—as addressed by critical theory—still be based on the distinction between critical thinking and automation? Can one truly argue that algorithmic automation is always already a static reduction of critical thinking? (p. 126)

She argues that with the all-machine phase transition of digital capitalism, we are bearing witness to a new mode of thought and control. This phase identified by a group of physicists from the University of Miami, coincides with the introduction of high frequency stock trading after 2006 and entails sub-millisecond speed algorithm to algorithm interactions that exceed human response time. Having analysed the millisecond scale data at the core of financial markets, they discovered a series of sub second extreme events caused by these algorithms. Given this state of affairs, Parisi argues that due to the inhuman scale on which these events unfold, this changes significantly our capacity to analyse them using the tools of traditional critical theory. Crucially she asserts that the traditional critical theory reproach that accuses computation of reducing human thought to mechanical operations is no longer sufficient as an analytical paradigm for our current state of affairs. Since these events are outside the limits of human control and comprehension, she cites computer scientist and mathematician Gregory Chaitlin's conviction that incomputability and randomness are in fact the very condition of computation. This means that the incomputable forms part of instrumental rationality itself. Parisi identifies the dilemma thus:

> Both philosophical thought and digitality, rely on principals of indetermination and uncertainty while featuring these principles in their core complexity theories. As such, both challenge and define the neoliberal order at the same time—a paradox. (p. 126)

Addressing this paradox, she argues for a turn to Chatlin's concept of incomputability, and specifically his discovery of the incomputable 'Omega number'; a number which is definable, but not computable. As Parisi explains: 'Omega defines at once a discrete and an infinite state of computation occupying the space between zero and one' (p. 126). She sees this as bringing not only the philosophical critique of technical rationalization into question but also the instrumentalization of reason. This signals, she says, an 'irreversible transformation in the history of critical thought in which the incomputable function of reason has entered into the automated infrastructure of cognition' (p. 127).

However, this recognition from a Lacanian point of view is actually not so new at all. Parisi's identification of the mysterious Omega number which unsettles all attempts to ground rationality on computation, being both a discrete yet infinite space between zero and one, and therefore incomputable, uncannily brings to mind the fundamental psychoanalytic notion of suture. In Jacques-Alain Miller's (2012) '*Suture*' paper, 'the first great Lacanian text not to be written by Lacan himself' (Badiou 2008, p. 25) Miller presents a logic of the signifier, now widely accepted as the strict formalization of the *Lacanian* logic of the signifier, although never systematized by Lacan himself. The paper grounds the logic with recourse to Frege's attempts at a logical conception of the sequence of whole natural numbers in which the categories of concept, object and number are distinguished in order to establish the zero as the only object subsumable under the concept of 'non-identity to itself'. As the editors of *Concept and Form* note, two relations are furthermore assumed by Miller: 'the subsumption of the object under the concept' and the 'assignation of a number to the concept' (Cahiers Kingston 2012, online).

By virtue of the thing non identical with itself being evoked only to then be excluded from the dimension of truth, Miller (2012) shows how Frege grounds the concept of number on a performative contradiction:

> This concept, by virtue of being a concept, has an extension, subsumes an object. *Which object? None.* Since truth is, no object falls into the place of the subsumed of this concept, and the number which qualifies its extension is zero. In this engendering of the zero, I have stressed that it is supported by the proposition that truth is. If no object falls under the concept of

> non-identical-with-itself, it is because truth must be saved. If there are no things which are not identical with themselves, it is because non-identity with itself is contradictory to the very dimension of truth. To its concept, we assign the zero. (p. 97)

Via this reading of Frege, Miller elaborates the paradoxical logic of the signifier and the subject of lack present in all discourses that aim at truth. Without entering into a comparative set-theoretical or conceptual discussion of Parisi's position versus Frege, it suffices to highlight that significant critical approaches to the question of AI seem to coincide with a logical paradox about the position of truth within discourse. In Parisi's case the supposed discrepancy that lies between the opposed discourses of rationality and computation are undermined by the apparition of a mysterious *stain on reality* that does not seem to fit into either side of the binary opposition of subjectivity and objectivity as she characterizes it. In reference to the Turing Machine qua 'absolute mechanism of iteration based on step by step procedures' (p. 130), Parisi argues that:

> Nothing is more opposed to pure thought—'or the being of the sensible' as Deleuze (1994: 68) called it—than this discrete—based machine of universal calculation. The Turing architecture of pre-arranged units that could be interchangeably exchanged along a sequence is effectively the opposite of an ontogenetic thought moving through a differential continuum, through intensive encounters and affect. (pp. 130–131)

What does Parisi here mean by the Deleuzian reference to ontogenetic thought, intensive encounters and affect? How does this translate into psychoanalytic terms?

This concern with the mysterious Omega number we may also read with reference to Lacan's early engagements with cybernetics in *Seminar II* and his (2006a) *écrit* '*Logical Time and the Assertion of Anticipated Certainty*' in which he draws attention to the significance of the algorithmic calculation of possibilities central to cybernetic theory and its possible use for constructing a psychoanalytic theory of causation. In his essay on logical time, Lacan illustrates via the thought experiment of the prisoner's dilemma, how rational calculation also contained within it a

temporal dimension that was entirely subjective and conversely how free action is only ever constituted via a universal rule:

> What makes this act so remarkable in the subjective assertion demonstrated by the sophism is that it anticipates its own certainty owing to the temporal tension with which it is subjectively charged; and that, based on this very anticipation, its certainty is verified in a logical precipitation that is determined by the discharge of this tension—so that in the end the conclusion is no longer grounded on anything but completely objectified temporal instances, and the assertion is desubjectified to the utmost. (Lacan 2006a, p. 171)

When deciding how to act in order to secure their freedom, the prisoners were entirely dependent on their intersubjective positioning towards one another, this involved not the flat ontology of hyper-rationality but, as Wright (2018) observes, the temporality of 'living scansion' (p. 75). The remarkable thing here is that the any act of freedom retroactively constitutes a completely determined mode of causality but is nonetheless experienced as subjectively authored. *In Seminar II* Lacan uses Edgar Allen Poe's *Purloined Letter* to illustrate the fundamentally illusory effect of chance that resides in the game of odds and evens, highlighting how, we tend to see coincidence in certain numerical patterns where mathematically speaking there is only probability; chance being purely an effect of structure. This becomes significant clinically when in *Seminar XI* Lacan makes an important distinction between two different kinds of psychoanalytic causality broadly corresponding to necessity and contingency. The first Lacan calls *automaton* and refers to a type of programmed repetition of behavior, and the second, *tuché*, corresponds to the chance and unpredictable encounter with the real. In terms of the speech of the analysand this is experienced as an unprecedented event, a radical subjective freedom that erupts from of the chain of pure machinic determination by the symbolic order. It is this unpredictable encounter with the real that we may call 'stupid' in that it by necessity lacks any rational explanation. It seems that Parisi envisions rationality as precisely this continuum of immanent affect, totally transparent to itself and self-positing, in other words *being* and *thinking* here are the same thing. Parisi diagnoses the

problem as one of discrete units versus continuous movement, perhaps a problem one could put differently in Lacanian terms as tuché versus automaton (Lacan 1977), or mathematization versus topology (Lacan 2016)?

Whilst Parisi attempts to obviate the question of the subject in relation to AI, it nevertheless slips in via the back door. For what is the function of the Omega number if not to guarantee absolute contingency in necessity and absolute freedom within determinism? It seems the Omega number, like the Lacanian real, *always returns to its place* and *never stops not being written*.

2.3 Malabou and the Blue Brain

In 2005 the Blue Brain Project was set up by Professor Henry Makram at the EPFL (2020), the Swiss Federal institute of technology in Lausanne, Switzerland. The aim of the project is to 'build biologically detailed digital reconstructions and simulations of the rodent brain and ultimately the human brain'. The project's super-computer based reconstructions promise to offer a radically new approach to understanding and simulating the multilevel structure and functioning of the brain. Whilst the progress of the project continues a pace, due to the immense complexity of the human brain, the BBP still only manages to simulate tiny portions of the brain in order to extrapolate larger findings from their limited but ever-growing harvest of data. The aim, according to their mission statement, however, is to be able to simulate in ever more detail the complex multilevel activity across different parts of the brain that would be impossible to investigate in living tissue. Researchers at the BBP may manipulate tissue in various ways, for example by 'knocking out' or 'lesioning' parts of the circuit.

As it stands, the researchers are capable of digitally reconstructing brain tissue 'in silico' to simulate a snapshot of the anatomy and physiology of the brain at any one moment in time. They can use these digital reconstructions for a virtually unlimited range of simulations in order to replicate the spontaneous electrical activity of a real brain. The project is currently building neurorobotic tools, which will enable the researchers

to replicate cognitive behavioral experiments in animals, with the aim of extrapolating this data to generate further understanding and simulations of the human brain. Ultimately, BBP is attempting what has hitherto been the preserve of science fiction and fantasy; building a fully functional non-biological brain. Even though the technology needed is still in its early stages, in theory with enough computer power, the Blue Brain Project envisages the possibility of complete digital simulation of the human brain. It is illuminating, therefore, to consider the *raison d'être* of their project:

> Understanding the brain is vital, not just to understand the biological mechanisms which give us our thoughts and emotions and which make us human, but for practical reasons. Understanding how the brain processes information can make a fundamental contribution to the development of new computing technology. (EPFL 2020, online)

Whilst the BBP recognizes the 'value' of understanding the brain, since as they put it, 'it gives us' thoughts, emotions and ultimately 'humanity', it appears as an almost secondary benefit. The most important driver in the development of the Blue Brain Project is less about understanding the human mind and more about the development of new more powerful computing technology. So, we have a curious reflexivity at play. By modelling the technology on the structure of the brain, we may come to better understand the functioning of the brain, which in turn helps us to develop new and better computing technology, in order to then model the brain in more detail. It is a perfectly circular argument which depicts the persistent computational metaphor at play in popular understandings of the human brain ever since the beginning of cybernetics.[5] Furthermore, the Blue Brain Project, with its absolute faith in the power of digital

[5] Thanks to the work of mathematician and creator of one of the early computers John Von Neumann and his model of computation, his definition of the essential equivalence of the human brain and the computer continues to influence computing to this day. As Kurzweil (2012) puts it: 'He acknowledges the apparently deep structural differences, but by applying Turing's principle of the equivalence of all computation, Von Neumann envisions a strategy to understands the brain's methods as computation, to re-create those methods, and ultimately to expand its powers' (p. xii). See also the Macy Conferences, set up by Joshua Macy Junior and Frank Fremont Smith in New York between 1946 and 1953.

simulation promises the erasure of the boundary between the functioning of the virtual and the actual and opens up the conceptual question of the possibility of a completely simulated intelligence. But if the brain is nothing more than a highly complex system which we are gaining the ability to reproduce digitally, at what point would the simulation cross over from virtual to real? Or in Hegelian terms, when would it pass from substance to subject?

The Blue Brain Project and other significant developments in the field of brain science (e.g. Neuralink) pose profound questions across many fields of research not just in Artificial Intelligence and engineering but in philosophy, ethics, and even cosmology (Lovelock 2019). In the wake of these groundbreaking developments challenging our notions of biological intelligence, the response from some significant thinkers has been suitably radical. One of them comes from Catherine Malabou. In her (2019) book *Morphing Intelligence: From IQ measurements to Artificial Brains*, Malabou mounts not just a critique of, but a rejection of her previous position on the status of the concept of intelligence as laid out in her (2008) seminal work *What Should We Do with Our Brain?* where she explores the neuro-scientific notion of brain plasticity. In her latest book, Malabou continues what she describes as her examination of the space between biological and symbolic life.

Malabou argues that in light of the development of synaptic chips and the Blue Brain Project, her former thesis on brain plasticity was chauvinistically humanist. The arrival on the scene of the artificial brain signals, in Malabou's view, a watershed moment that changes the status of human intelligence *tout court*. She goes as far as to state that her previous stance on plasticity was wrong, privileging as it did a biologically based model of intelligence. Her new perspective reconceives the brain as something not purely biological with the capacity for plasticity, but rather an entity whose symbolic character does not necessarily require a biological substrate at all.

Malabou lays out a genealogy of the scientific formulation of the concept of intelligence starting in the nineteenth century and tracks its appropriation by historians, psychologists, biologists and philosophers and the disputes that arose from their discussion over its meaning. Malabou goes on to outline what she calls the three main metamorphoses

of intelligence, which she calls firstly; genetic fate; secondly epigenesis and synaptic simulation, and lastly the power of automatism. The first metamorphosis is that pertaining to the invention of the IQ test, and the conceptualization of intelligence as a measurable entity for the first time in the nineteenth century. This led to the beginning of eugenics from the work of proto-geneticist Francis Galton followed by Alfred Binet and Theodore Simons research in the field of molecular biology and the sequencing of the human genome.

The second metamorphoses is supposedly our current configuration, that which is encompassed by the field of epigenetics, in other words, our generally and widely accepted understanding of intelligence that takes into account the effects of environmental factors, education, culture on the plastic functioning of the brain as the locus of intelligence. This second metamorphoses is largely locatable to the moment of the twentieth century when the relation between history and biology, thanks in part to French epistemologists such as Bachelard, Canguilhem and the latter's students Foucault and Simondon, started to be appreciated in a new light. It is the start of this reappraisal of the 'innatist and preformationist' (Malabou 2019, p. 15) character of intelligence that, Malabou argues, opens the door towards the possibility of the cybernetic brain that came to dominate the late twentieth century giving weight to the computer-brain metaphor.

The third metamorphoses is still to come. This would entail the removal of the boundary between nature and artifice; a conceptualization of intelligence that allowed for the simulation of brain activity to constitute something more complex and significant than a biopolitical 'threat to humanity'. Rather, it would signal a next stage in the evolution of the concept of intelligent life itself. On this score, we would have to approach Artificial Intelligence in a completely different way, and ask: what do we really mean when we call intelligence 'artificial'? This is in part what Malabou's book attempts to address; the unsaid and undisclosed dimensions of the concept of intelligence that lead us into prejudice as to what *counts* as intelligent life. She asks:

> how, then, should we situate artificial life in relation to biological and symbolic life? Is it an intruder, ever foreign and heterogeneous to them both,

> existing only as a threatening replica? Or is it, rather, the necessary intermediary that enables their dialectical interrelation? (p. xvii)

Indeed, how should we situate artificial life? Malabou does not attempt to resolve this question in the book, and the question of life in relation to intelligence remains indeterminate. In the postscript to *Morphing Intelligence*, Malabou admits that while writing the book, she had not yet realised 'the extent to which the issue of intelligence, particularly Artificial Intelligence, had become a pressing issue, one bound up with significant social, political, legal and economic implications' (p. 145). It was the enthusiastic reception of her book that made her aware that society once again expresses a 'deep and urgent *need for philosophy*' (p. 145; *italics in original*); a need for new tools to help it address the urgent questions posed by AI as a 'transformational technology' that, by virtue of its challenge to the traditional structures of information systems brings about 'a total upheaval of being-in-the-world' (p. 146). The nature of these challenges, Malabou argues, require a philosophical approach that would allow us to apprehend:

> rationally and without delusion […] a radical revolution not only in the conditions of thought, knowledge and expertise—notions commonly associated with intelligence—but in every field of activity, affectivity and the human psyche. (pp. 145–6)

This point is underscored by Malabou's parting shot, a quote from Freud's (1917) *Introductory Lectures on Psychoanalysis* on the third blow to human narcissism that was psychoanalysis.[6] Reflecting on the transformational nature of AI, Malabou (2019) asserts that the "fourth blow", as she sees it, will be the 'capturing of intelligence by its own simulation, exceeding and transcending it' (p. 164).[7] After reflecting on the upheaval that is AI, Malabou posits that:

[6] The first and second being the Copernican cosmological blow and second Darwinian evolutionary blow.

[7] It must be noted however that a fourth blow to human narcissism had previously been characterized by Donna Haraway (2008) as the 'informatic or cyborgian wound' in *When Species Meet*.

> the challenge is to invent a community with machines together, even when we share nothing in common with them. Never will there be a community of machines. The automatic creation they are capable of will have a political platform and ethical texture only if we endow them with it. (p. 161)

It is this final summation of her position that seems paradoxically at odds with the spirit of Malabou's book. For why, given AI's status here as the fourth blow to humanity, would machines not be capable of politics, an ethical act or the founding a community? It seems that whilst Malabou rightly refrains from reifying intelligence she falls into precisely the trap she is warning others of: viewing Artificial Intelligence as some perfectly spherical and discrete Other, unknowable and absolute.

It is interesting then that it is the concept of *stupidity* which Malabou notes, becomes absolutely pivotal in the tradition of French thinking on intelligence. Significantly, as she points out, the genealogy of modern French thought on the question of intelligence from Proust to Flaubert to Valery has always been accompanied by a reflection on stupidity, to which Deleuze's *Difference and Repetition* and Derrida's *The Beast and the Sovereign* would subsequently devote much thought, making stupidity the 'object of a properly transcendental question' (Deleuze cited by Malabou 2019, p. 7). As she puts it: 'a single word, "intelligence," characterizes both genius—natural intelligence—and machines—Artificial Intelligence. A gift is like a motor: it works by itself and does not come of itself. In this sense, then, it is stupid' (p. 8). Malabou concludes that intelligence and stupidity are but one, qualifying their respective fruitless striving for an essence as nothing more than 'ontological stubbornness' (p. 54).

> If such is the case, then why not give up intelligence as an independent philosophical question? The defenses are finally gathering to form this position: ultimately, the ontological void of intelligence is never as evident as in the stupidity of ontology. A stupidity that is, perhaps, not so very distinct from the stupidity of psychology. (p. 54)

Despite her emphasis here on the question of stupidity, Malabou maintains her commitment to rationality as key to a better understanding of

the possibilities of artificial forms of intelligence. It is curious then that Malabou, given her long standing dialogue with Freudian and Lacanian psychoanalysis,[8] raises the question of stupidity, given its relationship to enjoyment. For stupidity appears at various times in Lacan's seminar. In fact, he (1998) says plainly that '[t]he signifier is stupid' (p. 20) and more poetically: 'if an angel has such a stupid smile, that is because it is up to its ears in the supreme signifier. To find itself on dry land would do it some good—perhaps it wouldn't smile anymore' (p. 20). Stupidity, whilst not systematized by Lacan in theoretical terms is nevertheless emblematic of much of his conceptual endeavors. His (1988) three passions, love, hate and ignorance could all be called *stupid*. The latter passion, ignorance is perhaps the most self-evidently stupid, of which Lacan (2018) speaks in *Seminar XIX* in relation to the savoir of the analyst; ignorance being the basis of the analytic relationship; a level of knowledge attributed to the analyst, which in fact resides in a blind spot in the analysand themselves.

Then in *Seminar XX* ignorance becomes a divine attribute. Lacan (1998) reminds us that 'Freud arms himself with Empedocles statement that God must be the most ignorant of all beings, since he does not know hatred' (p. 91). To this ignorant and impotent Other, which forms the basis for both love and hate, Lacan gives pride of place in his graphs of sexuation as the barred Other. Formally speaking the three passions love, hate and ignorance as the three primordial forms of enjoyment are all stupid. I would argue that stupidity is precisely the blind spot necessary for 'real' intelligence. Not because of some mysterious attribute of "soul" or "consciousness" but because of the strictly formal operation of stupidity constitutive of enjoyment. The relationship between intelligence and stupidity is therefore not a simple question of opposites. Stupidity is precisely the part of intelligence that intelligence itself cannot see, the place from which the subject of intelligence *enjoys* itself.

While Malabou admirably revises what she perceives as her previous conceptual errors regarding the possibilities of machine intelligence, our

[8] See for example Adrian Johnston and Catherine Malabou (2013) *Self and Emotional Life: Philosophy, Psychoanalysis and Neuroscience,* in which the two authors engage in a debate on the potential challenges that the materialist turn in continental philosophy poses for psychoanalytic accounts of the subject.

conceptions of thinking and its relationship to the brain, she fails to consider the irreducible relationship between intelligence and its *constitutive failure*, which enters her discourse under the rubric of *stupidity*. In this sense, it is not the failure to conceive of a 'community of machines' that makes Malabou fall back on the kind of human exceptionalism that she attempted to escape in the first place, but the idea that intelligence can, by means of a careful process of reverse engineering, be eclipsed by its own simulation. This points us not to another poorly disguised argument for human exceptionalism (i.e. 'not all of the subject can be captured by simulation' or 'there is always a remainder that escapes simulation'), but a concern with the failure inherent to intelligence as such. A failure, furthermore, that recuperates the dimension of enjoyment which Malabou omits from her epistemology and which I see as vital to the challenge that Malabou sets us.

Malabou's opening gambit is to critique her previously held position on the question of human intelligence in light of new discoveries about the brain and new technologies, which put the unique status of the human brain under erasure. However, her acceptance of the genuinely philosophical problem of the synthetic brain still evades the central problem of subjectivity, which resides in the structure of enjoyment, or in psychoanalytic terms the sexual non-relation.

Malabou begins the conclusion to her investigation into the metamorphoses of intelligence with the following statement:

> IN THE END, intelligence is not ours, and it's not theirs either. This resistance to appropriation derives from the ontological paradox that constitutes it: intelligence has no being and cannot, therefore, belong to anyone. (p. 139)

By 'theirs' she is referring to those who wish to instrumentalize and appropriate the wholesale concept of intelligence in the synthetic recreation and simulation of the human brain (i.e. the neuroscientists of the Blue Brain Project). So Malabou admits intelligence has no *being*. In a sense an archetypal Lacanian position. Being is lost by virtue of existence, that is to say as soon as intelligence is reified and instrumentalised it ceases to be intelligent and becomes stupid. Yet paradoxically it is

precisely this stupidity which gives rise to the possibility of the conceptualization of intelligence itself. Intelligence is that which is always trying to coincide with itself and fails; the ontological void, constituted by the dialectic as we shall see between sex and knowledge.

As I will attempt to show towards the final part of this chapter, this third "metamorphosis" in the concept of intelligence that Malabou refers to then is the position which Lacanian psychoanalytic theory arguably already occupies. Intelligence on this score is, as we shall see, from the outset a paradoxical category at once artificial and stupid.

2.4 Anti-Philosophy: Thinking or Being?

Following Malabou's characterization of the stupidity of ontology; the enjoying blind spot of intelligence; Pasquinelli's constitutive failure; the originary trauma of rationality, Bratton's gendered and duplicitous AI and Parisi's incomputable (but real) Omega number it seems there is a constellation of crucial psychoanalytic issues emerging. Whilst these thinkers have articulated important and distinct dimensions of the critical treatment of Artificial Intelligence, they all seem to share a common theme; the unspoken omission of the psychoanalytic subject in philosophical and scientific discourses on AI. Lacan describes this problematic thus;

> The philosopher is inscribed […] in the discourse of the master. This doesn't mean that what he says is foolish; it's even more than usable … Nor does it mean, mind you, that he knows what he's saying. The court fool has a role to play: being truth's stand-in. He can play it by speaking like a language, just like the unconscious. That he is himself unconscious is of secondary importance; the important thing is that the role should be played. (Cited by Badiou 2018, p. 28)[9]

[9] Cormac Gallagher translates the last part of the final sentence of this quote from Lacan's (2009) *L'Étourdit* as 'the role should be held' (p. 42). Because 'played' in the above translation is more consistent with the relationship between 'role' and 'play', I have chosen not to reference Gallagher's version here, which is the edition I otherwise refer to throughout the remainder of the thesis.

In then to begin articulating how psychoanalysis (as distinct from philosophy) would approach the question of 'thinking' and 'knowledge' and their relationship to AI it is necessary to ground the discussion in the logic of the Lacanian subject. This will ultimately lead us to the underlying concept around which this book revolves; the sexual non-rapport.

Lacan's (2006b) 1966 paper '*Science and Truth*', written especially for the *Cahiers pour l'Analyse* and delivered at the beginning of *Seminar XIII The Object of Psychoanalysis* contains the first textual exposition of Lacan's subject of science, which he equates with the subject of psychoanalysis as that which is represented by 'a signifier for another signifier' (p. 875). Lacan's overarching aim in the paper is to identify the break in modern science heralded by Newtonian physics that leads to the eventual emergence of the subject of science. This involves delineating the passage that starts from Galileo and Descartes in the seventeenth century, 'the century of genius' (p. 857) and culminates in Freud's encounter with the unconscious at the turn of the twentieth century. Fundamentally, the subject for Lacan must be 'rigorously distinguished from the biological individual as from any psychological evolution subsumable under the subject of understanding' (p. 875).

As Ed Pluth (2019) points out, at this stage in his teaching Lacan is trying to work out the status of psychoanalysis as an autonomous discourse, aware of its debt to both structuralism and the natural sciences. He considers the subject to be constituted by a *Spaltung* or splitting; a gap between the subject of the enunciated content and the subject of enunciation, which is encountered in psychoanalytic practice in the form of the symptom. According to Lacan (2006b) psychoanalysis is situated at a specific junction in the history of science, making Freud not an anomaly of the scientific tradition but a direct product of it:

> I am saying, contrary to what has been trumped up about a supposed break on Freud's part with the scientism of his time, that it was this very scientism—which one might designate by its allegiance to the ideals of Brücke, themselves passed down from Helmholtz and Du Bois-Reymond's pact to reduce physiology, and the mental functions considered to be included therein, to the mathematically determined terms of thermodynamics (the latter having attained virtual completion during their

> lifetimes)—that led Freud, as his writings show, to pave the way that shall forever bear his name. (p. 857)

That said, the contribution that Freud makes was in a sense not to scientific knowledge, but to its absence. Psychoanalysis highlights the persistence in the subject of a relation to a 'not-knowing'. It is also in this essay where Lacan makes reference to Descartes' *cogito* as the correlate of modern science, a correlation which Descartes himself had misapprehended given his commitment to the belief in a non-deceitful God. The true significance of the Cartesian cogito is not perceived therefore until Freud, who realised that the real subject revealed in the cogito is not the ego of the 'I think', but the subject of the unconscious who postulates this punctual and vanishing moment of enunciation.

Any attempt to incarnate the subject (i.e. in any biologically determined sense, for example, as 'man') is therefore incompatible with the discoveries of both modern science and psychoanalysis. So fervent is Lacan in his conviction in the futility of pinning down the study of the subject using any of the currently available 'scientific' methods, that he makes reference to psychology as the only discipline deserving of that repugnant appellation "science of man", which he points out has found ways to outlive itself by 'providing services to the technocracy' (p. 730). With reference to Canguilhem's (2016) 1958 essay '*What is Psychology?*' Lacan notes that psychology 'slides like a toboggan from the Pantheon to the Prefecture of Police' (p. 730). Given the biopolitical ring to Lacan's distain for psychology, with its philosophical pretentions to know something universal about man only to then march him down to the police station, one can see how the work of '*Science and Truth*' prefigures Lacan's (2007) elaboration of the four discourses two years later in *Seminar XVII*. The essay contains Lacan's precursor to the four discourses in his elaboration of the three distinct modes of truth—formal, efficient and final—which correspond in their structure to science, magic and religion. These categories of thought, Lacan hints, will in turn find their expression in the differential subjective structures of neurosis, perversion and

psychosis and their respective epistemological drives of *Verdrängung*, *Verleugnung* and *Verwerfung*.[10]

Current thinking on the status of 'intelligent life' in the universe from purportedly scientific mindsets, often consists of a baffling mix of magical, religious and scientific thinking, shifting the position of truth as cause within their discourse. Take for example Tegmark's (2017) reflections on the conscious awakening of the universe:

> Before the universe awoke, there was no beauty. This makes our cosmic awakening all the more wonderful and worthy of celebrating: it transformed our universe from a mindless zombie with no self-awareness into a living ecosystem harbouring self-reflection, beauty and hope—and the pursuit of goals meaning and purpose. Had our universe never awoken, then, as far as I'm concerned, it would have been completely pointless, merely a gigantic waste of space. (p. 22)

Tegmark's quasi-religious sentiments of wonder seem to merge seamlessly with his valoration of instrumental rationality and genuine surprise at the sheer miraculousness of it all. Magical thinking, Lacan (2006b) notes, entails a dissimulation of knowledge for the subject of science. This, he says, is one of magic's conditions; it 'involves the truth as cause in its guise as efficient cause' (p. 742). Religious thought, on the other hand, involves an eschatological operation as truth as final cause in the form of revelation. As for science, Lacan says it 'does not want to know anything about the truth as cause'; a form of foreclosure is thus the way in which science must proceed in order to produce knowledge, what he terms a 'successful paranoia' (p. 742). In correspondence with Aristotle's four categories of causation, the missing fourth category would then be material causality, which subsequently will come to be occupied by the discourse of psychoanalysis in *Seminar XVII*. All four relations to truth will later become

[10] As Pluth (2019) remarks in his commentary on the essay:

> [Lacan's] remarks about magic and religion earlier can be thought of in terms of repression and negation respectively. While he was explicit about religion involving a negation of the truth as cause, the connection between magic and repression was never made clear. Perhaps his references to the status of knowledge in magic as obscure, could be read of one of the consequences of this repression. Since the truth as cause is repressed in magic—knowledge about magic's efficacy—why it works would remain a mystery to its practitioners and participants. (p. 299)

formalised in a slightly different configuration as the Master's, University, Analysts, and Hysteric's discourse, with science becoming associated with the Hysteric's discourse as that which ceaselessly produces new knowledge.

This formal distinction between truth and knowledge finds its most precise theoretical iteration in the characterization of Lacan's teaching as anti-philosophy, a term which he first used in reference to himself in 1975 in *Peut-être à Vincennes* (Lacan 2001) and subsequently mentioned on only two occasions throughout his seminar to refer to his method. The term however was not invented by Lacan himself. Despite its origins in religious critiques of rationalism, it is crucial to understand that for Lacan, anti-philosophy was not a mysticism nor a religious concern with the ineffable, the human soul or other such metaphysical concepts, rather the point for Lacan was to reassert the primacy of formalization, and the structure of the production of knowledge as distinct from philosophical truth.

In terms of the psychoanalytic approach to AI, anti-philosophy offers us an alternative perspective on the stakes of the debate on AI than we have thus far encountered. It is not so much a question of finding 'new paradigms' to think AI but rather of elucidating exactly how the Lacanian subject is already present in our conception of Artificial Intelligence, and what this mean for the relationship between the two. These inherent 'impossibilities' that we have identified reside at the very kernel of both psychoanalysis and AI.

Whilst anti-philosophy originated in the eighteenth century as a reaction to French enlightenment thought, by critics who sought to defend church authority and religious dogma against the rationalist tide that was descending over Europe, for Lacan, as Samo Tomšič (2016) notes, the term took on a completely different, not to say opposite meaning. Anti-philosophy in this new post-Freudian context was on the contrary linked by Lacan to the modern scientific revolution and its ramifications for the pre-modern Aristotelian orientation in philosophy and science. For Lacan, the term anti-philosophy was not a mere superficial rejection of philosophy, but rather an interrogation of its formal structure; the discursive position of truth in relation to knowledge, to be precise it is an interrogation of the 'imbecilities of philosophy' (Lacan 2001) in Lacan's view (the stupidity of enjoyment). As Tomšič (2016) notes, antiphilosophy

was to form one of the pillars of Lacan's *quadrivium* in his proposition for the transition of psychoanalytic knowledge, the other three disciplines being linguistics, mathematical logic and topology. According to Tomšič these other three disciplines are dependent upon:

> three crucial decentralizations conditioned by scientific modernity: decentralisation of language which suspends the organonic (pragmatic) theory of language; decentralization of knowledge, which detached it from the human observer; and finally decentralization of space which progressively gave rise to non-Euclidian geometries and restructured the space of thinking. (p. 102)

These three domains, Tomšič concludes, linguistics, mathematics and topology each could be said to be sciences uniquely concerned with specific dimensions of the Lacanian organon, symbolic, real and imaginary respectively. The fourth term in this quadrivium anti-philosophy, therefore, would function as the *sinthome* tying together the borromean trio as the discipline which resumes the fundamental lesson of psychoanalysis the 'decentralization of thinking' (ibid.).

Lacan's consistent engagement with the work of important philosophical figures was integral to the different stages of development of his psychoanalytic theory. But it was Lacan's endeavor to mathematize the transmission of psychoanalysis, using a plethora of symbols, letters and algebraic formulas and diagrams, which was a direct attempt to escape the bastardisation and misappropriation, which befell the Freudian discovery given its deceptive appearance as "ordinary language" open to hermeneutic abuse. The matheme therefore was, in the Greek sense of '*ta mathemata*', that which can be transmitted without loss (Johnston 2014, p. 254). One of the key agendas of Lacan's anti-philosophy according to Adrian Johnston is:

> [T]o draw attention to the fact that the philosophical (and quotidian) conception of psychoanalysis as a depth hermeneutics in search of the profound meaning of psychical suffering is a hopelessly wrongheaded misreading of Freud and his place in the history of ideas. (p. 255)

And as Johnston notes, it is thanks to the scientific rupture signified by the name Galileo, that according to Lacan the conditions for Freud's discovery became possible in the first place. Superficial rebuttals of the claim to call Lacan an anti-philosopher then often rely on an objection that what Lacan meant by philosophy was either a proxy for university discourse (as he elaborated in his four discourses in *Seminar XVII*), or that on the other hand the philosophy he had in mind was that of the Deleuzo-Guattarian kind, which was ostensibly mounting its anti-Oedipal attack on psychoanalysis in the post '68 revolutionary era. These strands of straw man argumentation, however, do not get to the crux of what is really at stake in Lacan's significance in relation to the discipline of philosophy. Alain Badiou, who devoted an entire seminar series to Lacan's anti-philosophical endeavors (2018) designates him anti-philosopher exemplar along with figures such as Gorgias, Pascal, Rousseau, Kierkegaard, Nietzsche, and Wittgenstein before him, for Badiou modern philosophy is only possible after first traversing Lacan. Indeed, he places Lacan as third in a succession of such pivotal figures in the history of philosophy as Plato and Kant.

According to Matthew Sharpe's (2015) reading, Lacan's claim to be an anti-philosopher in his later teachings has a number of related origins. Firstly, the concern with philosophy as a historical cultural practice of extracting knowledge. This Lacan derives from the influence of Alexandre Kojève and his reading of Hegel's Master-Slave dialectic. On this account, the master was never the man of knowledge, but rather the warrior prepared to fight to the death in the struggle for recognition. The master converts the slave's practical knowledge or *technai* (his lived *savoir faire*) into theoretical knowledge (*epistēmē*). Sharpe points us here to the primal scene of this paradigm of philosophy, which Lacan (2007) locates in *Seminar XVII* as Plato's *Meno*, in which the slave boy is made to recall mathematical knowledge that he didn't know he had through Socrates questioning: 'refer to the *Meno*, where it is a question of the square root of 2 and its incommensurable. There's someone who says. "Hey look, get the slave to come over, that little fellow can't you see he knows"' (Lacan 2007, p. 22). For Lacan it is this transubstantiation of knowledge from technical know-how, (the slave's drawings in the sand) into theoretical episteme (i.e. Euclidean geometry) that constitutes pre-modern

philosophy. 'Philosophy in its historical function is this extraction, I would almost say this betrayal, of the slave's knowledge, in order to obtain its transmutation into the master's knowledge' (ibid.). Whilst this characterization of philosophy as the Master's discourse, is susceptible to the critique that it homogenizes the whole of philosophy into one type of thinking, or one disciplinary form, Lacan will later shift his anti-philosophical critique to a slightly different target. In *Seminar XX*, Sharpe (2015) shows how Lacan mounts his critique again, this time for the types of metaphysical claims philosophers *per se* are prone to making. Again, Lacan (1998) reaches back into the history of philosophy, this time to the pre-Socratics. First of all, it can be said that we have changed the thinking subject considerably:

> Since, the "I am thinking" that presupposes itself, grounds existence, we have had to take a step, that of the unconscious … The subject is not the one who thinks. The subject is precisely the one we encourage, not to say it all (*tout dire*), as we tell him in order to charm him—one cannot say it all—but rather utter stupidities. That is the key. (p. 22)

Here Lacan is referring to what he sees as the fundamental presupposition of pre-modern philosophy deriving from Parmenides of Elea's poem, specifically this passage:

> For without the being in relation to which it is uttered you cannot find thinking. For there neither is nor shall be anything outside of being, since fate (Moira) bound it to be whole and immovable. For that reason, all these will be mere names which mortals have laid down, convinced that they were true: coming-to be as well as passing away, Being as well as non-Being, and also change of place and change of shining color. (Cited by Sharpe 2015, p. 11)

In response to this, Lacan (1998) notes that:

> It is precisely because he was a poet that Parmenides says what he has to say to us in the least stupid of manners. Otherwise, the idea that being is and non-being is not, I don't know what that means to you, but personally, I find that stupid. (p. 22)

The elision of thinking and being, Lacan takes as paradigmatic of the whole of Western philosophy right up through to Hegel's 'the real is the rational' (Sharpe 2015, pp. 11–12). According to Lacan it is an error to assume that all that can be thought necessarily can also be or that the world is a mirror of the subject who thinks it. As Sharpe (2015) points out, this imaginary gestalt of the oneness of being and thinking Lacan ascribes to some fundamental pre-modern axioms that derive from the Aristotelian notion of the spherical nature of the universe. He sees the decisive break with the advent of modern science taking his cue however, not from the Copernican revolution that displaces the earth from the center of the universe, but rather from the Keplerian discovery of the elliptical orbit of stars. Displacing all the celestial bodies instead to one side of a bipolar ellipsis, the other point always being an empty space.

Ultimately, Lacan sees the beginning of the subject of modern science as Descartes' cogito. Lacan (2006b) regards the Cartesian cogito not as hypostatized being that thinks but a punctual evanescent moment. This performative character of the cogito guarantees only that there is thinking, not that there is a thing that thinks. As he remarks in 'Science and Truth':

> Which is why it is worth restating that in the test of writing *I am thinking: "therefore I am"*, with quotes around the second clause, it is legible that thought only ground being by knotting itself in speech where every operation goes right to the essence of language. (p. 734)

In *Seminar XI* Lacan gives his account of the cogito 'as a forced choice between cogito and sum' (Dolar 1998, p. 18). Here Lacan makes the distinction between thinking and being; one must choose between the two. If one chooses thinking one must give up being and vice versa. Lacan's point, in this forced choice is that, sum doesn't follow once one has made the first step. Thought depends on the signifier, which turns the subject into the empty point of enunciation, instead of founding his/her being. In the place of the supposed certainty of the subject's being, there is just a void. As Dolar explains: 'It is not the same subject that thinks and that is; the one that is, is not the one that thinks, even more the one that is, is ultimately not a subject at all' (p. 19). Lacan's (1977) point, simply

put, is that being and thinking are different concepts. Returning to our earlier philosophical approaches to Artificial Intelligence, their concerns with failure, dupery, incomputability and stupidity we may recognize a certain impossibility emerging. This impossibility arises by virtue of what Lacan would call the split or barred subject. It is precisely this split between being and thinking that radically changes the way we may conceptualize AI as a 'thinking thing'. This split between being and thinking at the heart of Lacan's theory of the subject, casts a different light on philosophical conceptualizations of AI as a discrete form of intelligence in contradistinction to so called 'real intelligence'. The subject of knowledge is already one riven between the truth of the enunciation and the enjoyment produced discursively. Lacan's (1998) main contention against philosophy then was its claim to sense, which for him was what differed about the *truth* of psychoanalysis, since 'the unconscious is the fact that being, by speaking, enjoys' (pp. 118–119). Perhaps AI is best conceived not as a thing which thinks, but a thing which is thought, or as Lacan (1998) puts it:

> The problem with the kind of science I qualify as traditional, because it comes to us from Aristotle's thought, is that it implies that what is thought of (*le pensé*) is in the image of thought, in other words that being thinks. (p. 105)

The human subject is *spoken by* language. It is because of this enjoyment, that the truth of psychoanalysis, as he (2009) articulates in *L'Étourdit* is an *ab-sens*; a lack of sense. This ab-sens is the hole in symbolization that is otherwise referred to as the *non-existent sexual relation*. The Lacanian sexual non-relation is ultimately located within the dimension of language as the organ of castration, and thus the creation of jouissance. The philosophical question of truth is thereby supplanted by the *anti-philosophical* psychoanalytic question of enjoyment. This means that metaphysical knowledge becomes sexual knowledge. It is the question of enjoyment, in lieu of 'sense' that is the pursuit of the psychanalysis (as opposed to the philosophy) of artificial Intelligence. In the following chapter I will examine the *object* of this enjoyment.

Bibliography

Badiou, A. (2008) *Number and Numbers*. Cambridge: Polity Press.
Badiou, A. (2018) *Lacan: Anti-Philosophy 3*. New York: Columbia University Press.
Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
Bratton, B. (2015) 'Outing Artificial Intelligence: Reckoning with the Turing Test' in M. Pasquinelli (Ed.), *Augmented Intelligence and its Traumas*, pp. 69–80. Lüneburg: Meson Press.
Cahiers Kingston (2012) Synopsis of Jacques-Alain Miller, 'La Suture: Elements de la logique du signifiant'. Available (01.03.20) at: http://cahiers.kingston.ac.uk/synopses/syn1.3.html.
Canguilhem, G. (2016) What is Psychology? *Foucault Studies* 21: pp. 200–213.
Dolar, M. (1998) 'Cogito as the Subject of the Unconscious' in S. Žižek (Ed.), *Cogito and the Unconscious*, pp. 11–40. Durham: Duke University Press.
EPFL (2020) Blue Brain Project. Available (01.03.20) at: https://www.epfl.ch/research/domains/bluebrain/.
Freud, S. (1917) 'Introductory Lectures on Psycho-Analysis ' in (2001) *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XVI (1916–1917): Introductory Lectures on Psycho-Analysis (Part III)*, pp. 243–496. London: Vintage.
Hammond, K. (2018) TedX: A New Philosophy on Artificial Intelligence [Video File]. Available (01.03.20) at: https://www.youtube.com/watch?v=tr9oe2TZiJw&t=179s.
Haraway, D. (2008) *When Species Meet*. Minneapolis: University of Minnesota Press.
Johnston, A. (2014) *Adventures in Transcendental Materialism: Dialogues with Contemporary Thinkers*. Edinburgh: Edinburgh University Press.
Johnston, A., & Malabou, C. (2013) *Self and Emotional Life: Philosophy, Psychoanalysis and Neuroscience*. New York: Columbia University Press.
Kurzweil, R. (2012) Foreword. In J. Neumann (Ed.), *The Computer and the Brain*. New Haven: Yale University Press.
Lacan, J. (1977) *The Seminar of Jacques Lacan Book XI: The Four Fundamental Concepts of Psychoanalysis*. London: Karnac Books.
Lacan, J. (1988) *The Seminar of Jacques Lacan Book I: Freud's Papers on Technique 1953–1954*. London: W.W. Norton & Company.

Lacan, J. (1998) *The Seminar of Jacques Lacan Book XX: Encore—On Feminine Sexuality, the Limits of Love and Knowledge 1972–1973*. London: W.W. Norton & Company.

Lacan, J. (2001) 'Peut-être à Vincennes…' in J. Lacan (Ed.), *Autres écrits*, pp. 313–315. Paris: Éditions du Seuil.

Lacan, J. (2006a) 'Logical Time and the Assertion of Anticipated Certainty' in J. Lacan (Ed.), *Écrits*, pp. 161–175. London: W.W. Norton & Company.

Lacan, J. (2006b) 'Science and Truth' in *Écrits*, pp. 726–745. London: W.W. Norton & Company.

Lacan, J. (2007) *The Seminar of Jacques Lacan Book XVII: The Other Side of Psychoanalysis*. London: W.W. Norton & Company.

Lacan, J. (2009) L'Étourdit. *The Letter (Irish Journal for Lacanian Psychoanalysis)* 41: pp. 31–80.

Lacan, J. (2016) *The Seminar of Jacques Lacan Book XXIII: The sinthome*. Cambridge: Polity Press.

Lacan, J. (2018) *The Seminar of Jacques Lacan Book XIX: … Or Worse*. Cambridge: Polity Press.

Liu, L.H. (2010) *The Freudian Robot: Digital Media and the Future of the Unconscious*. Chicago: The University of Chicago Press.

Lovelock, J. (2019) *Novacene: The Coming Age of Hyperintelligene*. London: Penguin.

Malabou, C. (2008) *What Should We Do with Our Brain?* New York: Fordham University Press.

Malabou, C. (2019) *Morphing Intelligence: From IQ Measurement to Artificial Brains*. New York: Columbia University Press.

Miller, J-A. (2012) 'Suture (Elements of the Logic of the Signifier)' in P. Hallward & K. Peden (Eds.), *Concept and Form Volume One: Key Texts from the Cahiers pour l'Analyse*, pp. 91–101. London: Verso.

Parisi, L. (2015) 'Instrumental Reason, Algorithmic Capitalism and the Incomputable' in M. Pasquinelli (Ed.), *Alleys of Your Mind: Augmented Intelligence and its Traumas*, pp. 125–137. Lüneburg: Meson Press.

Pasquinelli, M. (2015) 'Introduction' in M. Pasquinelli (Ed.), *Alleys of Your Mind: Augmented Intelligence and its Traumas*, pp. 7–18. Lüneburg: Meson Press.

Pluth, E. (2019) 'Science and Truth' in S. Vanheule, D. Hook & C. Neill (Eds.), *Reading Lacan's Écrits: From 'Signification of the Phallus' to 'Metaphor of the Subject'*, pp. 268–307. London: Routledge.

Searle, J.R. (1980) Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3(3): pp. 417–457.
Sharpe, M. (2015) Killing the Father, Parminedes: On Lacan's Antiphilosophy. *Continental Philosophy Review* 52: pp. 51–74.
Tegmark, M. (2017) *Life 3.0: Being Human in the Age of Artificial Intelligence*. London: Penguin.
Tomšič, S. (2016) 'Psychoanalysis and Antiphilosophy: The Case of Jacques Lacan' in A. Cerda-Rueda (Ed.), *Sex and Nothing: Bridges from Psychoanalysis to Philosophy*, pp. 81–103. London: Karnac Books.
Von Neumann, J. (2012) *The Computer and the Brain*. New Haven: Yale University Press.
Wright, C. (2018) 'Lacan's Cybernetic Theory of Causality: Repetition and the Unconscious in Duncan Jones' *Source Code*' in S. Matviyenko & J. Roof (Eds.), *Lacan and the Posthuman*, pp. 67–88. Basingstoke: Palgrave Macmillan.

Filmography

Cameron, J. (Director) (1984) The Terminator [Film]. US: Orion Pictures.
Columbus, C. (Director) (1999) Bicentennial Man [Film]. US: Buena Vista Pictures.
Spielberg, S. (Director) (2001) A.I. Artificial Intelligence [Film]. US: Warner Bros. Pictures.