MODELING THE HUMAN MIND AS AN INFORMATION PROCESSOR:

A HEIDEGGERIAN CRITIQUE

EMRE ALPAGUT

BOĞAZİÇİ UNIVERSITY

2020

MODELING THE HUMAN MIND AS AN INFORMATION PROCESSOR:

A HEIDEGGERIAN CRITIQUE

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Cognitive Science

by

Emre Alpagut

Boğaziçi University

2020

DECLARATION OF ORIGINALITY

I, Emre Alpagut, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature..

Date ......25/09/2025..............................

ABSTRACT

Modeling the Human Mind as an Information Processor: A Heideggerian Critique

Research efforts seeking to construct a scientific study of mind became institutionalized as cognitive science after World War II. These research efforts were inspired by the technological developments of that era in three domains: cybernetics, Turing computation and information theory. Although originally designed for engineering applications in controlled environments, these technological frameworks and concepts were adopted by researchers to give an account of the human mind. Consequently, "mind is an information processor" emerged as the central premise of cognitive science. Despite this information-related premise, cognitive science lacks a rigorous definition of the term information. Furthermore, the information processor approach necessitates a conception of mind consisting of discrete, enumerable constituents whose relations are well-defined. Such a conception recognizes mental capabilities insofar as they can be portrayed as a computational process and explains away metaphysical conceptualizations of subjectivity by portraying mental processes as emergent effects produced from complex interactions among simple entities. This relationship between technology and science during cognitive science's proliferation can be examined from a Heideggerian perspective. Heidegger conceptualizes technology as a mode of being and argues that technology uses science to fulfill its essence. The technological mode of being conceives all entities as standing reserve (*Bestand*), resources to be extracted and ordered. Cognitive science models mental processes as computational operations and thereby represents the human mind as an orderable resource.

ÖZET

İnsan Zihninin Bilgi İşlemci Olarak Modellenmesi: Heideggerci bir Eleştiri

Zihin hakkında yapılan bilimsel araştırmalar, İkinci Dünya Savaşı sonrasında bilişsel bilim disiplini etrafında kurumsallaştı. Bu araştırmalar, o dönemde üç alanda gerçekleşen teknolojik ilerlemeden ilham aldı: sibernetik, Turing'in hesaplama konusundaki araştırmaları ve enformasyon teorisi. Esasen kontrollü ortamlarda yapılan mühendislik uygulamaları için geliştirilen bu kuramlar, insan zihnini incelemek için uyarlandı. Böylece, "zihin bir bilgi işlemcidir" yargısı bilişsel bilimin temel önkabulu haline geldi. Böyle bir önkabule ragmen, bilişsel bilim hala iyice tartışılıp titizce oluşturulmuş bir *bilgi* tanımına sahip değildir. Dahası, bilgi işlemci yaklaşımı, zihni kesin tanımlı bağıntılarla birbirne bağlı, ayrık ve sayılabilir bileşenlerden oluşan bir bütün olarak tasavvur etmeyi zorunlu kılar. Bu kavrayış zihinsel süreçleri hesaplanabilir bir süreç olarak tarif edildiği ölçüde kabul eder ve öznelliği basit bileşenlerin karmaşık etkileşimleri sonucunda ortaya çıkan bir zihinsel süreç olarak açıklar. Bilişsel bilimin doğuşu sırasında bilim ve teknoloji arasında süregelen ilişkiyi Heideggerci bir bakış açısından incelemek mümkündür. Heidegger teknolojiyi bir varoluş modu olarak kavramsallaştırır ve bu varoluş modunun diğer varlıkları bir duran-stok (*Bestand*), yani çıkarılıp düzenlenecek kaynaklar olarak algılar. Bilişsel bilim, zihinsel sürüçleri hesaplanabilir süreçler olarak modelleyerek insan zihnini düzenlenebilir bir kaynak olarak temsil eder.

ACKNOWLEDGEMENTS

I am grateful for the presence of Professor Yağmur Denizhan in my life for the past three years. In my attempt to describe her impact on me and my academic work, titles such as teacher, advisor and mentor come off as understatements. She has been a source of inspiration with her erudition, compassion, and infectious enthusiasm for intellectual endeavors. To me, she embodies the Heideggerian notions “thinking is thanking” and “questioning is the piety of thought”.

I am also thankful for the guidance of Professor Johannes Fritsche, who was initially my co-advisor for the thesis, however, is unfortunately no longer with us. His wisdom and patience is the reason I was able to develop a thesis project seeking to connect cognitive science and Heidegger’s work on technology.

I would like to thank Doç. Dr. Chryssi Sidiropoulou for accepting the role of co-advisor upon Professor Fritsche’s untimely death and her help with the last sections of the thesis.

The Chaos etc. seminar at Boğaziçi University has been a source of emotional and intellectual support throughout my master’s education. They have truly fostered a sense of community in me and I am thankful for their presence. I would especially like to thank Özge Dağ for her friendship and all her help. I would also like to thank Galip Ümit Yolcu for always being intellectually engaged with my thesis and his feedback on my work.

I would like to thank the Core Curriculum at Boston University. They have fostered a sense of community very similar to that of Chaos etc. while showing me the value of an education in the humanities and the value of foundational texts.

I would also like to thank those closest to my heart: Yasemin Alpagut, Aliye Taner, Mustafa Alpagut and Ece Özsü. I am very lucky to have been blessed with their unending love and support.

I am thankful to my previous teachers that have impacted me greatly. I would especially like to thank Stephanie Nelson, Barak Caine, David Roochnik, İsmail Coşkun and Candaş Sert.

Finally, I would like to thank Müjgan, the ill-tempered stray cat that kept me company while I wrote my thesis on the deck behind my house.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Cognitive science is described as the scientific interdisciplinary study of mind in contemporary textbooks (Friedenberg & Silverman, 2006). The field's main premise is that the mind is an information processor that performs computations on representations (Friedenberg & Silverman, 2006).

The current historic period is often called the *Information Age*. Developed countries have transitioned into post-industrial economies that utilize information technologies to create wealth. Companies commercializing digital technologies such as Microsoft, Apple, Amazon and Google are now the most valuable companies in the world.

Unsurprisingly, cognitive science's conceptualization of mind followed the development of information technologies. Thus, cognitive science characterizes the efforts studying the mind as information processor in the Information Age. In contrast, at its inception, cognitive science was intended to harbor also alternative approaches to the mind as well. The first Center for Cognitive Studies, founded at Harvard University in 1960, invited figures such as Eric Hobsbawm, a prominent scholar of social-political history, and Ernst Gombrich, a prominent art historian, to work on the study of cognition with the psychologists who had founded the center (Boden, 2008, p. 345). Such outreach to the different disciplines of humanities are no longer the case.

The founders of the Center envisioned cognitive science as an umbrella discipline that would provoke fruitful interactions among psychology, linguistics, artificial intelligence, anthropology, and philosophy. Today, it is criticized for not developing into the interdisciplinary melting pot that it was intended to become. A recent paper takes up this matter through analyzing bibliometric and socio-institutional indicators and asserts that “cognitive science failed to move from a collection of enthusiastic multidisciplinary efforts to an integrated coherent interdisciplinary field” (Núñez et al., 2019, p. 9). The paper demonstrates that 63% percent of the authors published in the journal *Cognitive Science* are affiliated with a psychology department, 65% percent of the citations in the journal are from psychology journals, nearly half of the contemporary cognitive science faculty have their Ph.D. training in psychology and contemporary cognitive science programs have on average 60% of their coursework in psychology. The researchers’ findings suggest that compared to the other related disciplines, psychology is disproportionately represented in cognitive science.

Like cognitive science, the discipline of psychology itself has transformed over cognitive science’s lifetime. Until the 1950s, psychology was dominated by behaviorism, an approach focusing on analyzing externally observable behavior. Behaviorist psychology did not study mental activity because data about mental activity could only be obtained through introspection, making it unfit for scientific study. Starting in the 1950s, encouraged by the mind-as-computer paradigm and the developments in the field of brain imaging that claim to render mental activity externally observable, the discipline of psychology went through a “cognitive revolution” shifting its focus from behavior to mental activity (Friedenberg & Silverman, 2006, p. 96).[1]

[1] The decline of the behaviorist approach can also be explained it terms of behaviorism’s own theoretical difficulties and methodological shortcomings. Due to its focus on the historical relationship between cogniton and computation, this thesis will focus on the role of the computational paradigm in cognitive science’s historical development.

Psychology's cognitive turn was simultaneous with the inception of artificial intelligence as a full-fledged research program. While cognitive psychology sought to reverse-engineer cognitive capacities, artificial intelligence sought to engineer them. The complementary missions of these two fields compelled them to engage with each other. In the 1960s cognitive psychologists "saw cognitive science as essentially the marriage between psychology and artificial intelligence" (Núñez et al., 2019, p. 7).

Cognitive science's information-processing paradigm is rather unwelcoming towards alternative approaches such as those that focus on the integrity of the human subject. In contrast, cognitive science – as it is one of the claims of this thesis – defines itself as the endeavor of formulating quantifiable components and the relations among them, in order to account for cognitive processes. This contrast, combined with cognitive science's unsuccessful interdisciplinarity, creates an environment where studies and approaches from disciplines other than psychology are considered a part of cognitive science insofar as they reaffirm the notion of mind as an information processor. Alternative conceptualizations of mind are discouraged, if not excluded from the scope of the field.

Another paper analyzes how cognitive scientists use the term *information processing*. Authors point out that this term has become highly polysemous: different cognitive scientists use the term to mean different things – which itself may be another testament to cognitive science's failure to become a mature interdisciplinary science (Piccinini & Scarantino, 2010). The authors also state that most cognitive scientists regard the terms information processing and *computation* to be equivalent and use them interchangeably. The assumption of such equivalence eliminates the possibility of analyzing the similarities and differences between cognition and computation.

Sticking to the information processor premise implies in all instances that cognitive science must insist on a computational framework while giving an account of mind. Its dedication to a computational approach can be seen in the announcement of the 2017 meeting of the *Cognitive Science Society*: "Computation can serve as the foundational theory of how people actively process information in service of control and decision making … greater effort must be made to connect cognitive science theories to computational foundations" (as cited in Núñez et al., 2019, p. 7). The computational interpretation of the field's information processing premise provokes cognitive science to jettison non-computational views of mind.

Another noteworthy issue regarding cognitive science's theoretical grounding is that, while persistently employing the notion information processor as a model for mind, cognitive science lacks a rigorous definition of the term *information*. Contemporary textbooks, despite explicitly subscribing to the information processor premise, do not present a definition for the term information (Friedenberg & Silverman, 2006). During the field's inception, George Miller, one of the founding fathers of cognitive science, imported the terminology of Claude Shannon's *Information Theory* and used Shannon's conceptual toolkit to give an account of the human mind. Shannon, however, had developed his theory to tackle a communications engineering problem and did not consider his theory to be necessarily relevant to the field of psychology. Furthermore, Shannon excluded any notion of semantics from his theory. In contrast, cognitive science employs the term information to give a comprehensive account of all human mental activity which must include semantics. Shannon's theory and Miller's adoption of it is examined in greater detail in Chapter Two and Chapter Three.

Cognitive science's adoption of the term information is representative of the field's proliferation: a newly developed exciting technological framework is employed to reverse engineer and give an account of the human mind. There were very significant developments in the theory of computation, cybernetics and information theory around the time of the World War II. After the war, these technologies were utilized to create a science of the mind and these efforts were institutionalized as cognitive science in the 1960s. This process is also recounted in Chapter Two and Chapter Three.

Chapter Four presents Heidegger's analysis of the relationship between technology and science as given in his essay "The Question Concerning Technology", as well as some other components of his thinking that are relevant for a critique of cognitive science. In his essay, Heidegger analyzes technology as a mode of relating to the world. The technological mode of being challenges us to encounter the world as *standing reserve (Bestand*), a collection of orderable resources that is to be extracted, stored, and later called upon on demand. In this context, cognitive science's computational framework can be regarded as the intellectual enterprise seeking to conceptualize the human being as standing reserve (Bestand), a resource to be steered and ordered in accordance with the organization of other resources.

In his technology essay, Heidegger claims that technology uses science to fulfill its essence. To this end, technology employs modern science to represent nature as a coherence of forces calculable in advance. Such a representation allows science to achieve maximum manipulability of its object. This, in turn, creates fertile ground for the technological mode of being to encounter the world as a collection of orderable resources.

Heidegger also claims that the conception of human being as standing reserve (Bestand) is enabled by the conception of thought in terms of logic. Claiming such logic-based thinking closes us up to the essence of thinking, Heidegger argues for a conception of thinking based on *logos*. For Heidegger, logos describes the human being's capacity to gather, disclose and shelter things in their being. Understood as speech, logos describes making matters manifest by talking about them. Understood as careful consideration, logos describes one's capacity to thoughtfully engage with one's world to start things on a way to be manifested. Such a conception of thinking, Heidegger argues, can shelter us from the danger posed by the technological mode of being.

Historically, Heidegger's association with cognitive science is mostly due to the discussion of *Being and Time* in the context of artificial intelligence. The primacy attributed to practicality over theorizing behavior in *Being and Time* is also touched upon in Chapter Four to expand the discussion on the adoption of technological developments to give an account of the human mind.

In conclusion, the present thesis argues that cognitive science should develop a rigorous definition of the term information. If it is preferred to keep this definition insensitive to meaning as in the case of Shannon information, a region of validity for the 'mind as information processor' premise needs to be confined to leave room for the development of complementary approaches that harbor meaning, relevance and multidimensionality of human thought. Bounding the premise with a region of validity also provides an opportunity for cognitive science to delimit itself and harbor metaphysical conceptualizations by admitting what it cannot quantify or objectify into existence. Only then can cognitive science accommodate an adequately rich conception of the human being.

CHAPTER 2

HISTORICAL BACKGROUND OF THE TECHNOLOGICAL DEVELOPMENT

2.1. Turing computation

2.1.1 A brief history of the term *algorithm*

In 1900, mathematician David Hilbert gave a now-famous speech at the International Congress of Mathematicians. There, he identified 23 mathematical problems that should be taken on in the 20th century. The tenth problem on his list inquired about the formal definition of the term *algorithm*. Indeed, the rigorous study of this concept brought forth the digital technological revolution that has been taking place throughout the last century.

Although the notion algorithm has been used in mathematics for a very long time, there was only an intuitive appreciation of algorithm and it was not formally studied until the 20th century. Before then, the notion simply meant recipe or procedure and was used to designate sets of well-defined discrete steps to carry out a certain mathematical task. The rigorous study of the notion opened up multiple different frontiers in the history of mathematics, including but not limited to logic, theoretical computer science and analytic philosophy.

The groundwork for a formal definition of algorithm came out of David Hilbert's and Kurt Gödel's research in early 20th century. David Hilbert started a research program in the 1920s that proposed a new foundation for classical mathematics. The program hoped to reduce mathematics to a finite number of axioms, from which all possible theorems could be deduced. Hence, the program saw mathematics as a complete formal system.

First published in 1931, Gödel's incompleteness theorems proved Hilbert's program to be impossible. Gödel demonstrated that, beyond a moderate level of complexity, all logic-based formal systems that are made up of finite, sound and consistent set of axioms will have true but unprovable statements. [2] Hence, Hilbert's program could not have ever deduced every possible mathematical theorem from a set of axioms. Furthermore, Gödel's theorems accomplished two major breakthroughs. First, it demonstrated that "logic can be arithmetized" (Dupuy, 2009, p. 37); logic can be expressed solely in terms of arithmetic, as operations on numbers. This revelation isolated and coined the term "effective computability" (Dupuy, 2009, p. 36), denoting a subset of logic-based systems in which every true statement is provable. [3] This term later created the ground for the contemporary notion of an algorithm. In doing so, it laid the foundation of computer science.

Building on the notion of effective computability, Alan Turing proposed the mathematical formulation of an abstract machine that can calculate every effectively computable procedure. Proposed in 1936, the *Turing machine* formalized the notion of algorithm; if there is an algorithm for a task, the task can be simulated by a Turing Machine and if a task can be carried out by a Turing machine, there exists an algorithm for the task.

Turing machines receive inputs, apply a set of logical rules on the inputs and compute outputs.[4] Thus, every unique Turing machine corresponds to a unique algorithm that has a different set of input-output pairings.

Later in 1936, Turing further proposed the idea of a universal machine, namely "a machine capable of imitating, mimicking, reproducing, *simulating* the

---

[2] Beyond a level of complexity that is sufficiently rich to accommodate arithmetic.
[3] "Effective computability" can also be referred to as "recursively enumerable".
[4] Meanwhile, all of the machine's intermediary states are reflected in the successive changes made on the input.

behavior of any other Turing machine" (Dupuy, 2009, p. 40, emphasis Dupuy). This is possible, since the logical rules that govern a given Turing machine's behavior can be arithmetically represented and passed on to the universal Turing machine.

2.1.2 Formal definition of a Turing Machine

A Turing Machine (hereafter TM) is effectively the model of an abstract computer that has an infinite amount of memory. TM's formal definition is summarized in Figure 1 at the end of the section. TMs consist of three main parts: tape, tape head and control unit.

The tape makes up the TM's working memory. Prior to any computation, the tape contains the input and continues to contain the modifications as the input gets processed. The tape is made up of infinitely many individual cells, each of which can contain only one symbol. The tape head designates the specific cell the TM is working on at the moment. The tape head can read, write and overwrite the content of a cell when ordered to do so by the control unit. Likewise, the control unit can command the tape head to move along or back to a different cell.

The control unit is a finite state device that embodies the TM's *transition function*. It enumerates the TM's set of possible states as well as its current state, processes the input, updates the current state, changes the tape cell's content when necessary and orders the tape head its next move. Among the set of possible states, two states have special designations, namely the accept state and the reject state. If the TM reaches this state while computing, the TM either accepts or rejects the input according to its corresponding state.

TMs are always defined with respect to an alphabet that bounds the range of symbols the TM can process. This is often referred to as the *input alphabet*. Hence, the set of possible inputs consists of strings made up of different permutations of members of the input alphabet with unlimited repetitions. The set of inputs accepted by a specific TM is described as the *language* of that TM. In essence, in the context of a fixed alphabet, every unique language has a corresponding TM and denotes a unique algorithm.

Lastly, the *tape alphabet* consists of the input alphabet and a unique blank symbol that is not in the input alphabet. Since there are infinitely many cells and every cell after the input has a blank symbol, this distinction enables the TM to detect the end of its input.

DEFINITION

A ***Turing machine*** is a 7-tuple, $(Q, \Sigma, \Gamma, \delta, q_0, q_{\text{accept}}, q_{\text{reject}})$, where Q, Σ, Γ are all finite sets and

1. Q is the set of states,
2. Σ is the input alphabet not containing the ***blank symbol*** $\sqcup$,
3. Γ is the tape alphabet, where $\sqcup \in \Gamma$ and $\Sigma \subseteq \Gamma$,
4. $\delta : Q \times \Gamma \longrightarrow Q \times \Gamma \times \{L, R\}$ is the transition function,
5. $q_0 \in Q$ is the start state,
6. $q_{\text{accept}} \in Q$ is the accept state, and
7. $q_{\text{reject}} \in Q$ is the reject state, where $q_{\text{reject}} \neq q_{\text{accept}}$.

Fig. 1 Formal definition of a Turing Machine
Source: Sipser, M. (2012). *Introduction to the theory of computation*. Boston: Cengage Learning.

2.1.3 The limits of computation

As it has been popularized by the famous Imitation Game, Turing was not humble about the capabilities of the Turing Machine. [5] Likewise, daring perspectives such as

[5] First proposed by Alan Turing in 1950, the Imitation Game has the following form: “There are three players: a machine, a human being, and an interrogator who, being unable to see or hear either of the two, must try to determine which is which through conversation with them by means, for example, of a teleprinter. The machine’s strategy is to try to mislead the interrogator into thinking that it is the human being, while the latter attempts to affirm his or her human identity. The machine will have sufficiently well *simulated* the behavior of the human player if in the end the interrogator cannot tell them apart.” (Dupuy, 2009, p. 41, emphasis Dupuy).

pancomputationalism became popular following the emergence of the Turing Machine. [6] However, from a philosophical perspective, the limits of Turing's construct have to be scrutinized. The endeavor to understand those limits can start with an investigation of whether there are languages that cannot be recognized by a Turing Machine or an investigation into how many Turing Machines there are.

Such an investigation relies on the mathematical findings of the 19th century, specifically the work of Georg Cantor. Cantor rigorously studied infinite sets and proved that there exist infinite sets of different sizes. The smallest infinite set is said to be *countably infinite* and is denoted $\aleph_0$ within the hierarchy of infinities constructed by Cantor. This type of infinite set can be enumerated, written as a list and makes up a bijective function when paired with the set of positive integers.

Cantor initially proved that the set of real numbers is an infinite set that is more populated than countably infinite sets. He referred to such sets as *uncountably infinite*. Cantor proved this distinction through his diagonalization argument, in which he demonstrated that upon any attempt to enumerate an uncountably infinite set A, an element that is a member of A and is not on the list can be constructed.

Cantor's findings have important implications for the theory of computation: it has been established that there are countably infinitely many Turing Machines and uncountably infinite languages. There are more languages than there are Turing Machines. Since every Turing Machine describes one language, there exist languages that cannot be recognized by a Turing Machine. In other words, there exist problems in the world that cannot be accommodated by an algorithm.[7]

---

[6] Pancomputationalism describes the claim that every phenomenon in the universe is essentially computational and, hence, every phenomenon can be described in computational terms.

[7] Uncountable infinity is indeed a construction of the mind in the attempt to capture all aspects of being including the unknown. One can dismiss uncountable infinity as an irrelevant abstraction by arguing that it is not encountered in nature. Likewise, one could claim that uncountable infinity and countable infinity are both relevant abstractions in the attempt to understand the unknown. This

2.2 Cybernetics

2.2.1 Background

Cybernetics is the study of "self-regulating systems, in which information about the results of the system's actions is fed back so as to cease, adjust, or prolong the original activity" (Boden, 2008, p. 198). Since this regulation process is principally guided by flow of information, an abstract notion, as opposed to matter or energy, the cybernetic conceptual structure was easily applied to a diverse set of phenomena, including mental.

Although the design of self-regulating, autonomous systems dates back to antiquity, its conceptualization as a science-technological treatise happened in mid-20th century. Norbert Wiener's 1948 work *Cybernetics: Or Control and Communication in the Animal and the Machine* pioneered the field, formalized the notion of feedback and sought to describe how regular behavior is generated without distinguishing between animate and inanimate. [8]

Wiener's chosen name for the field is borrowed from the Greek word *kybernetes* meaning literally 'steersman' and metaphorically 'governor'. Wiener chose the name also as a tribute to James Clerk Maxwell who had presented a paper called *On Governors* to the Royal Society in 1868. This paper discussed negative feedback in mathematical terms and was inspired by James Watt's device called

---

discussion aside, this framework definitively establishes that the number of problems of exceeds the number of algorithms.

[8] As it will be mentioned in the next section, cybernetics' proliferation owes to the Second World War. Margaret Boden (2006) notes that until after the war much of the work related to this field was classified. Hence, it is expected that most work on cybernetics, including Wiener's, had matured prior to their publication during times of war but were kept classified.

Governor which automatically regulated the speed of a steam locomotive via feedback. [9]

2.2.2 The Second World War and the post-war transition

The Second World War provided previously unimaginable opportunities for cyberneticians, both in terms of funding and practical challenges. The relentless arms race between the conflicting countries initiated the launch of the first radar-guided missiles. These weapons involved self-corrective servomechanisms; they computed their current trajectory, the position of their target and autonomously took action to change their trajectory in line with the target. Naturally, the arms race also initiated the development of anti-ballistic missiles, which also anticipated the movement of the missile's tracked target and adapted automatically.

The predictive adaptations accomplished by military equipment encouraged cyberneticians to extrapolate their established mechanism as the ground for 'purposive' and 'teleological' behavior. Indeed, in a famous paper co-authored by Wiener (1943), teleological behavior is defined as "behavior controlled by negative feedback". Thus, the notion of purpose was analyzed as an attempt to reduce the difference between the agent's current state and goal state. Furthermore, cyberneticians did not differentiate between biological organisms and machines in terms of purpose and self-equilibration; cat chasing a mouse, missile updating its trajectory and human regulating its body temperature were represented in the same manner within this framework.

[9] The term *negative feedback* has different operational definitions within cybernetics and contemporary psychology. In cybernetics *negative feedback* means error reducing feedback, whereas in the context of B.F. Skinner's operant conditioning framework, *negative feedback* (or *negative reinforcement*) means the removal of an unpleasant stimulus. In other psychological experiments *negative feedback* is also used to describe the emotional valence of the feedback.

In the post-war period, there were further attempts to develop cybernetic devices.[10] No longer confined to military applications alone, researchers now sought to build self-regulating artefacts to understand psychological and biological phenomena. In the 1950s, there were multiple attempts to make a mechanical maze-runner rat that attempted to navigate through a maze and registered its errors while doing so. Likewise, the first attempts at automated chess players emerged out of these efforts.

2.2.3 Macy conferences and later revisions

As stated, cyberneticians felt encouraged to apply their framework to a wide range of phenomena and further develop the self-regulating automata they were designing. To that end, the Macy Conferences provided a fruitful environment where the brightest minds of the post-war period came together.

Organized as ten separate gatherings from 1946 to 1953, the Macy Conferences brought together prominent mathematicians, logicians, engineers, physiologists, neurophysiologists, psychologists, anthropologists and economists of its time. These meetings were held under the name *the Cybernetics Group*. The group declared its mission as constructing a general science of how the human mind works. Jean Pierre Dupuy, who has studied the conferences extensively, lists the two convictions that guided the group in its mission:

> 1. Thinking is a form of computation. The computation involved is not the mental operation of a human being who manipulates symbols in applying rules, such as those of addition and multiplication; instead it is what a particular class of machines do –machines technically referred to as 'algorithms.'
> 2. Physical laws can explain why and how nature – in certain of its manifestations, not restricted exclusively to the human world – appears to us

[10] It is worth noting that, by then, most cybernetic devices were analog, not digital (Boden, 2008).

> to contain meaning, finality, directionality and intentionality. (Dupuy, 2009, p. 3-4)

Thus, the Macy group sought to investigate the human mind as a self-regulating system whose subjective experiences can be 'naturalized' in accordance with the laws of physics by regarding the subjective experiences of thinking as algorithmic in nature. [11]

Some of the participants of the conferences later sought to revise the framework outlined in the Macy Conferences. One of these projects, led by Heinz von Foerster, Margaret Mead and Gregory Bateson, is referred to as *second-order cybernetics*. They took up an issue with the fact that the observer is never a part of the model constructed by cyberneticians. All of the models assume a 'God's eye view' in which the observer and the observation are not a part of the model. Foerster and his colleagues sought to integrate the observer into the model. Hence, their academic project is often referred to as the 'cybernetics of cybernetics' or *second-order cybernetics*. This approach is particularly important for cybernetics' study of human minds, since it emphasizes the fact that human minds attempting to theorize the governing principles of human minds, must account for their capacity to theorize the governing principles of human minds.[12]

---

[11] Cybernetics' attempt to study humans emerged in the context of a substance dualist conception of humans. This conception owes to the legacy of Descartes, whose work presents mental and physical beings as distinct and separable beings. To naturalize, in this context, means to describe a mental phenomenon in terms of physical beings.

[12] If this capacity is not accounted for the scholar falls into the trap of a 'Cartesian theater' where the framework can only be explained in terms of a higher order framework. A similar example, namely the frame problem, will be discussed at length in Chapter Four.

2.3 Information theory

2.3.1 Key terms

Unlike most detailed mathematical theories, information theory has been almost completely laid out in a single publication, namely C.E. Shannon's 1948 paper titled "A Mathematical Theory of Communication". Shannon's theoretical framework has been crucial for the development and proliferation of communication technologies such as mobile phones and the internet.

Shannon's theory is solely concerned with the logistics of communicating a message between a *sender* and a *receiver*. The theory assumes that the set of all possible messages contains a finite number of elements that are known to both sender and receiver prior to the communication. Hence, the communication event is simply the referral to an element of the message set. Furthermore, to carry out this process optimally, the parties can *code* the message as a shorter, unique string of characters chosen from a finite alphabet, referring to an indexing of the elements of the message set. [13] Consequently, the sender does not even have to reiterate the message itself but can refer to it with the shorter codeword. The receiver will simply *decode* the message.

Messages containing codewords will be transmitted bit by bit. In an efficient scenario, every transmitted bit will narrow the set of possible codewords transmitted by the message. In this context Shannon has introduced *information* and *entropy* as complementary terms. Every transmitted bit provides additional information, thereby 'reducing the uncertainty' at the receiver end, 'reducing the ignorance', thus 'increasing knowledge'. and incrementally reducing entropy which can be

[13] The strings do not necessarily have to be binary; any finite base larger than 1 would serve the same purpose. The larger the alphabet is, the shorter the codeword can be. Though, for technical applications, binary notation is almost always preferred. In binary notation, the smallest unit of data is called a bit (binary digit).

interpreted as ‘lack of knowledge’.[14] We can thereby compute the information content of any given message.

2.3.2 Quantifying Information

As stated, Shannon’s framework aspires to transmit the message with the shortest possible coding. The most obvious heuristic method is giving the message’s ordinal in the ensemble rather than reconstructing the message; i.e., for a message set with *t* messages, the ordinal can be stated in $\log_2 t$ bits.

Assume a random variable X in a message of length k that can take on s alternative values $\{x_1, x_2, \ldots, x_s,\}$. Let k_i denote the number of times the value x_i occurs in a message $x_1 x_2 \ldots x_k$ and that $k_1 + k_2 + \ldots + k_s = k$. We can calculate the number of possible messages satisfying the given conditions as follows:

$$\binom{k}{k_1, k_2, \ldots, k_s} = \frac{k!}{k_1!\, k_2! \ldots !\, k_s!}$$

As stated, the $\log_2$ of this number gives us the number of bits we need to refer to the ordinal of any one of these messages.

$$\log_2 \frac{k!}{k_1!\, k_2! \ldots !\, k_s!}$$

Under the stated conditions, the *relative frequency* of a symbol x_i is $p_i = k_i / k$.[15] Using this equation and Stirling’s formula[16], we can simplify the equation further:

$$h(x) \approx k \sum p_i \log \frac{1}{p_i}$$

---

[14] It should be noted that such expressions are typically used by cyberneticians, philosophers computer scientists, and cognitive scientists as if they have an agreement about the expressions’ meaning and relevance. This is one of the pitfalls of interdisciplinary studies. A computer scientist or a cybernetician would accept that receiving information increases knowledge, whereas such a statement would be blasphemous for a philosopher.

[15] It should be noted that “relative frequency” and “probability”, although different notions, are used equivalently and represented by p_i. In *An Introduction to Kolmogorov Complexity and Its Applications*, the authors warn: “With abuse of terminology and notions, henceforth we use ‘probability’ for ‘frequency.’ Under certain conditions on the stochastics nature of the source this transition can be rigorously justified” (Li & Vitányi , 2008, p.67).

[16] $\log(k!) \approx k \log(k) + O(k)$.

This formula enables us to compute the Shannon entropy of a message prior to its transmission. We can simply refer to this value as a measure of our ignorance about the content of the message. As the message is transmitted bit by bit, the entropy- the receiver's ignorance- decreases and information content increases by the same amount. By the completion of reception, information content of an amount of h(x) bits is received and the receiver's entropy is 0.

2.3.3 Meaning

Shannon's framework is not concerned with the *meaning* of the message; Shannon says "[f]requently the *messages* have meaning: that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem" (Shannon, 1949, p. 31, emphasis Shannon). Shannon limits the engineering problem to the efficient reproduction of isomorphic signal structures by transmitting only the ordinal of the intended message within the message set such that the receiver can recreate the message.

What renders a communication event meaningful is the semantic appreciation of the message. By disintegrating the semantic aspects from the communication process, Shannon recognizes that the capacity for semantic appreciation is irreducible to operations on the transmitted codewords according to a set of prescribed rules. Indeed, from its outset, Shannon's conceptual toolkit is developed to improve the technical communication of sender and receiver whose semantic appreciation is taken for granted.[17]

Since the quantity of information is defined independently of the meaning of the message in Shannon's framework, a hundred–letter excerpt from a Dostoyevsky novel and a sequence of random hundred letters have the same information content.[18] Information theory's radical exclusion of meaning encourages the treatment of syntax and semantics as separate realms. As a result, information theory states that syntax can be investigated separately, tacitly implying that semantics can be added later on.

2.4. A new composite science

As discussed in the previous sections, three new technological frameworks emerged around the time of the Second World War. Firstly, Alan Turing formalized the notion of *algorithm*. Since Turing successfully represented logical propositions in terms of arithmetic and performed operations on them, his framework was able to autonomously represent and navigate through questions of high computational complexity addressed at complete formal systems.[19] Therefore, Turing's breakthrough promised the capacity to construct autonomous decision-making algorithms for every condition which had formally defined the necessary and sufficient conditions for its fulfillment.

---

[17] In his paper, Shannon makes the following remark: "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point" (p. 31, Shannon) 1949). This clearly marks the transmission of isomorphic signal structures as the goal of his project and places semantic appreciation outside of the intended scope.
[18] Given that they exhibit the same statistical dependencies.
[19] Computational complexity is a theory used to classify computable problems based on their inherent difficulty, based on the time and storage resources a Turing Machine needs to solve the problem. Without any abuse of terms, it would be fair to say that a problem of high computational complexity is a complicated logic problem.

Secondly, the cybernetic framework enabled devices to exhibit self-governing behavior by tracking and responding to the discrepancy between the device's current state and goal state. This framework relied on the formalization of the term feedback by Norbert Wiener. Initially developed to satisfy military concerns, the framework was extended to take on non-military topics after its initial success. This process triggered the proliferation of the field of control and systems engineering in the second half of the 20^{th} century.

Lastly, Shannon's information theory provided a robust framework to quantify the amount of information in a message and provided the most efficient means to transmit that message. Since Turing's and cybernetics' frameworks relied on arithmetic representation and computational procedures, Shannon's methodology provided the means to communicate the necessary messages in the procedure. As a result, computer science, control and systems engineering were complemented by a robust mathematical framework for communication. That said, in information theory, communication was viewed as a purely syntactical manner and, hence, was reduced to the efficient reproduction of isomorphic signal structures at the receiver end.

Turing and the Cybernetics Group were not humble about the limits of their frameworks. Both considered their framework to be sufficiently rich to account for the human experience. Turing's 1950 paper *Computing Machinery and Intelligence* puts forth the Turing test and further suggests providing a digital machine "with the best sense organs that money can buy, and then teach it to understand and speak English" (Turing, 1950, p. 460). Likewise, the Cybernetics Group argued their control mechanisms can account for teleological and purposive behaviors exhibited by humans. Moreover, like Turing, they considered thinking to be a purely computational procedure. Further empowered by Shannon's information theory, these frameworks aspired to take on metaphysical terms such as subjectivity, agency, intention and goal.

The intellectual framework constructed by these three disciplines blurred any possible distinction between the animate and inanimate. For them, an artifact could be engineered to be intelligent and the intelligence of a living being could be reverse engineered. After all, intelligence, in this framework, is regarded as a matter of information-processing and information as a completely quantifiable signal structure.

Shannon's framework utilized the *bit* as its smallest, indivisible, discrete unit. Based on Boolean logic, Shannon's digitized framework's most basic component is essentially a Yes/No question.[20] Likewise, Turing's framework considered the symbols in a Turing Machine's alphabet as its smallest, discrete, indivisible unit. In this sense, both frameworks are constructed based on the assumption that their object of representation can always be successfully reconstructed and represented by manipulations on digitized components with discrete values. In contrast, the first prominent cybernetic devices such as the Bush differential analyzer, the Kerrison predictor and Kelvin's tidal predictor were mechanical analog devices (Boden, 2006). Unlike digital devices, analog devices can register signals of any value and do so continuously over time.

Jean Pierre Dupuy (2005) notes that the discussion over whether the mind should be represented as a digital system or an analog system made up one of the most heated discussions of the Macy Conferences. Ultimately, the digital conception prevailed. Indeed, the technological revolution that took place in the second half of the 20th century also involved electronic digital computers rather than analog. This is perhaps because digital computers are more robust, not susceptible to noise and

[20] Traditionally 1 corresponds to Yes and 0 corresponds to No.

economically feasible for scaling up in production. As a result, an information and communications technology (ICT) revolution came to define the rest of the century. Furthermore, these conditions created a fertile ground for this framework to attempt to give an account of the human experience. The Macy Conferences were the first attempt in this regard. Later efforts became institutionalized under the name *cognitive science*.

CHAPTER 3

COGNITIVE SCIENCE: ITS BIRTH AND DEVELOPMENT

3.1 Information theoretic psychology

3.1.1 George Miller's research

In *Mind as Machine: A History of Cognitive Science*, Margaret Boden describes George Miller as the "first information-theoretic psychologist" (Boden, 2006, p. 286). Boden examines Miller's research in the late 1950's at length; fascinated by Shannon's framework, Miller sought to import Shannon's terminology to the realm of psychology. Throughout the decade Miller collaborated with almost every foundational figure in cognitive science. Finally, in 1960, Miller founded, along with Jerome Bruner, The Center for Cognitive Studies at Harvard University, the first research institution dedicated to cognitive science.

George Miller's 1956 paper *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information* is his first attempt to bring Shannon's terminology to the realm of psychology. Boden describes this paper as the most-cited paper in the whole of cognitive psychology (Boden, 2006, p. 288). Throughout the paper, Miller presents some of Shannon's key terms such as *information*, *channel capacity*, *bit* and *recoding*, while explaining how they can be utilized in the design and analysis of psychological experiments. According to Miller, these terms complement the experimental setup so well that he considers it to be a "historical accident" that psychologists have not been using Shannon's framework up to that point (Miller, 1956, p. 343).

Miller's adoption of information theory deviates from Shannon's initial setup in one fundamental way; while Shannon envisioned the communications channel as a device that transmitted signals from sender to receiver, Miller states that, in his experimental setup, the research participant "is considered to be a communication channel" (Miller, 1956, p. 344). Miller would expose the participant to unidimensional stimuli such as tones that differ in frequency. The participants are then expected to order the tones according to similarity. Miller states that subjects make mistakes when the experimental setup includes more than four tones. Since two binary units are required to enumerate four different items, Miller concludes that "2.5 bits…is…the channel capacity of the listener for absolute judgments of pitch" (Miller, 1956, p. 345). [21]

Upon discussing other examples, Miller discusses immediate memory experiments, where the researcher presents stimuli such as binary digits, decimal digits, letters of the alphabet, monosyllabic words etc. After all stimuli are presented, the participant is required to restate all he has retained. Miller states that in such experiments the bits of retained information varies drastically. However, he still manages to explain the phenomenon with the jargon of communication theory; he declares that the participants are recoding the stimuli. Miller states that the participants organize the input into familiar units or chunks. The capacity of immediate memory is limited to seven chunks. Efficient recoding patterns allows one to retain more bits per chunk.

Miller cherished the fact that information is a "dimensionless quantity" (Miller, 1956, p. 343) since this allowed him to compare channel capacities across modalities. He noted that, previously, such processing capacities had been stated in terms of units of measurement. However, information in a discrete statistical

[21] Miller then discusses multidimensional stimuli. He states that adding further dimensions to the stimuli adds some capacity to the subject's ability to accurately characterize the stimuli. However, the increase is lower than the sum of the capacities of the dimensions combined.

distribution does not depend on the unit of measurement. Consequently, using the amount information as a benchmark allowed him to compare measurements based on different metrics. Miller applied three different tasks and ended up with the 'magical number' while trying to measure the mind's channel capacity in three different tasks.

George Miller truly believed that the discipline of psychology had to model the human as a communication channel and accept information as its core concept. He maintained this conviction throughout the rest of his academic career. In one of his final papers, he makes the following remark:

> Erwin Schrödinger pointed out that organisms survive by ingesting, not food, not calories, but negative entropy. It is no accident, of course, that the mathematics of entropy are also the mathematics of information. The analogy is obvious: Just as the body survives by ingesting negative entropy, so the mind survives by ingesting information. In a very general sense, *all higher organisms are informavores*. (Miller, 1983, p. 113, emphasis Miller)

For Miller, humans were merely a special case of the more general theory of information processing (Miller, 1983). In fact, Miller indeed posited the logical extension of his position and speculates whether anything of important be lost if the study of minds were restricted to computer simulations (Miller, 1983).

3.1.2 Chomsky comes on the scene

George Miller gave a seminar on mathematical psychology at Stanford University in 1957. He chose the young Noam Chomsky as his assistant for the seminar. Miller was already a tenured professor by then and Chomsky had earned his doctorate just two years before the seminar due to his work on transformational grammar. Chomsky had successfully described the containment hierarchy of formal languages and the corresponding grammars by then. Chomsky had thereby demonstrated that a set of finite, precisely described set of rules can generate the infinite number of expressions attributed to a formal language.[22] Eager to investigate the "psychological reality of syntax" (Boden, 2008, p. 286) and extend Chomsky's framework to natural language, Miller worked with Chomsky and introduced him to contemporary psychologists.

Chomsky's work inspired Miller and changed his approach to psychology: "it was largely because of Chomsky that Miller replaced 'information' by 'computation' as the core concept of psychology" (Boden, 2008, p. 296). In 1965, Chomsky published *Aspects of the Theory of Syntax* in which he outlined his transformational generative grammar as applied to natural language. Boden states that the greatest legacy of Chomsky may be the fact that he transformed the psychological discussion about minds and language and encouraged talking about them in computational terms. Boden says, because of Chomsky:

> [psychologists] began to think of language as a precisely describable generative system. Chomsky didn't try to measure information, nor to locate it on a flow chart, nor to incorporate it in a program. But he asked how it was structured, how it could be generated and transformed, and what (very abstractly defined) types of computational system would be able to represent this. That's why he was an important voice in the rise of computational psychology. (Boden, 2008, p. 298)

3.2 The neuron doctrine and its mathematical models

3.2.1 The neuron doctrine

Although contemporary researchers take the fact for granted, cell theory, the argument that cells are the basic unit of organization and reproduction of organisms, had not been formalized, documented and settled until the 19th century. The development of cell theory was catalyzed by the developments in microscopy and

[22] Formal languages are simply referred to as "languages" in the section on Turing computation. Each formal language is a set of string(s) made up of the symbol(s) in its alphabet (alternatively the empty set also qualifies as a language). Different formal languages require different amount of memory and computational capacity to be recognized by TMs or computationally weaker automaton. Chomsky's hierarchy distinguishes among such languages.

tissue preparation. That said, while the greater picture had been settled, the discussion whether the nervous system also consisted of discrete, individual neuronal cells or was rather a continuous nerve net remained highly disputed.

Two researchers are credited for advancing the discussion about the nervous system, namely Camillo Golgi and Ramón y Cajal. Golgi advocated the *reticular theory* which suggested that the nervous system was a single and continuous network. In the late 19th century, Golgi also developed the silver staining method that enabled researchers to visualize neurons with all of their branched dendrites and axons. Cajal was Golgi's contemporary and used his staining technique to study the brains of birds. The *neuron doctrine*, the claim that the nervous system is made up of discrete individual cells, was developed as a result of Cajal's comprehensive research. Both figures were jointly awarded the Nobel Prize in Physiology or Medicine in 1906.

The debate is regarded to be completely settled in favor of the neuron doctrine after the Second World War. The development of electron microscopy in the 1950s allowed for the closer examination of neuronal tissue. This examination enabled the formalization of the *synapse*, the structure between neurons that enables their electrochemical communication.[23]

3.2.2 The McCulloch-Pitts neuron and the first neural networks

Warren McCulloch and Walter Pitts were among the key members of the Cybernetics Group that participated in the Macy Conferences. McCulloch was a neurophysiologist and the founding president of the American Society for Cybernetics. Walter Pitts was a logician whose work was foundational to the field of

[23] A synapse consists of three main parts: the nerve terminal of the pre-synaptic neuron, the synaptic cleft and the dendrite of the post-synaptic neuron.

computational neuroscience. Their paths crossed at the University of Chicago in the 1940s and they collaborated with the intention of using logic to give an account of the nervous system.

McCulloch and Pitts published their seminal paper *A Logical Calculus of the Ideas Immanent in Nervous Activity* in 1943. The paper sought to integrate three concepts: propositional calculus, Turing machines and neuronal synapses (Boden, 2006). The authors proposed a model of nervous activity based on a mathematical model of the neuron as its building block. Widely known as the McCulloch-Pitts Neuron, this building block produces binary outputs: the neuron is either 'on' and generates a signal, or it is 'off' and does not generate a signal. Both scholars were convinced that "the activity of the neuron is an 'all-or-none' process" (McCulloch & Pitts, 1943, p. 119) and, hence, "in psychology, introspective, behavioristic or physiological, the fundamental relations are those of two-valued logic" (McCulloch & Pitts, 1943, p. 132).

Since McCulloch and Pitts based their model of nervous activity on two-valued logic, this model failed to capture some characteristics of the nervous system that were discovered later in the 1950s. Their artificial neuron's structure was modeled after the *action potential*, a unidirectional, robust, all-or-none signal generated by neurons to facilitate cell-to-cell communication. The action potential and its role in the nervous system owes to *chemical synapses*, junctions between neurons that facilitate interneuronal communication through the release of chemical structures called *neurotransmitters*. On the other hand, it has been since then documented that neurons also communicate through *electrical synapses*, a bidirectional mode of communication in which current flows directly from a neuron to another.[24]

Since the chemical synapse invokes a binary response in the post-synaptic neuron, "the activity of any neuron can be represented as a proposition" (McCulloch & Pitts, 1943, p. 118). In this framework, the physiological activities of and the relations between neurons are treated as equivalent to the relations among propositions. Hence, interneuronal communication could be modeled as a propositional calculus describable in terms of logical operations. Thus, a neuronal network made up of McCulloch-Pitts Neurons would essentially be equivalent to a Turing machine with finite memory.[25]

McCulloch and Pitts attempted to use logic to address neurological and psychological questions. They were not humble about the limits of their model: they wrote "if any number [any function] can be computed by an organism, it is computable by these definitions, and conversely" (McCulloch & Pitts, 1943, p. 132). Their model essentially embodied the argument that the mind is a logic machine and proposed a kinship between the mind and Turing machines.

---

[24] The sustained exclusion of electrical synapses from the theoretical framework of neurophysiology and the focus on chemical synapses (and thereby on two-valued logic) perhaps points to a greater problem regarding the field itself. "Indeed, the functional roles of electrical synapses in many systems described above, as well as others not listed, are sometimes superficially understood or have not yet been investigated in the context of the physiology governed by the circuits in which these synapses occur. This is perhaps indicative that the subject of electrical synapses in mammalian systems is still an emerging field, possibly providing at least one reason why recognition of the potential importance of these synapses and the ways in which they contribute to shape neuronal activity has not as yet widely permeated the general neuroscience community." (Nagy et al. 2018). For a brief but comprehensive history of electrical synapses, please see (Bennett, 1997).

[25] A network made up of McCulloch-Pitts Neurons will always be finite. In contrast, universal Turing machines are not finite (Boden, 2008).

3.3 Mind as Machine: GOFAI and Beyond

3.3.1 Origins

As discussed in the previous chapters, developments in the fields of cybernetics and computability theory already foreshadowed a research program describable as 'Artificial Intelligence'. This research program got underway in the summer of 1956 when prominent researchers of the time gathered in Dartmouth College for a two-month-long workshop on artificial intelligence.[26]

The proposal for the workshop was authored by John McCarthy, Claude Shannon, Marvin Minsky and Nathaniel Rochester. The proposal's opening paragraph is as follows:

> We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. (McCarthy et al. 1955)

This proposal is also the first official usage of the term *artificial intelligence*. The proposal goes on to list seven issues concerning the 'artificial intelligence problem'. These are:

- Automatic Computers [Computers operating on the basis of a pre-programmed algorithm]
- How Can a Computer be Programmed to Use a Language
- Neuron Nets
- Theory of the Size of a Calculation
- Self-Improvement
- Abstractions
- Randomness and Creativity (McCarthy et al. 1955)

[26] The participants were Claude Shannon, Marvin Minsky, John McCarthy, Nathaniel Rochester, Ray Solomonoff, Oliver Selfridge, Alan Newell, Herbert Simon, Julian Bigelow and John Holland.

In 2006, John McCarthy stated that he attempted to avoid using the terms ‘computer’ and ‘computational’ in the proposal as much as he could in deference to Norbert Wiener, who was promoting analog cybernetic devices rather than digital computers (Russell & Norvig, 2002, p. 17). As discussed in the previous chapter, over time digital devices’ commercial success overwhelmed the theoretical concern directed towards analog computation and simultaneously, digital computation became the sole primary ground for AI research.

3.3.2 GOFAI

The research program that was initiated at the Dartmouth conference is often called *GOFAI*, short for *good old-fashioned artificial intelligence*. The term’s coinage owes to John Haugeland’s 1985 book. The term is intended to capture the theoretical attitude assumed by pretty much all AI researchers until the 1990s. This theoretical attitude is the understanding that intelligence is symbol manipulation. Hence, AI researchers’ concern had been the construction of high-level representations of problems in terms of the symbols of a pre-defined language and the manipulation of these symbols according to a pre-determined set of rules to compute the appropriate response.

The symbol manipulation approach has been formally stated as the *physical symbol system hypothesis* in Newell and Simon’s 1976 paper titled *Computer Science as Empirical Inquiry: Symbols and Search*. Their central thesis is that “a physical symbol system has the necessary and sufficient means for intelligent action” (Newell and Simon, 1976) and they posit that “the symbolic behavior of man arises because he has the characteristics of a physical symbol system” (Newell and Simon, 1976).

Hence, according to this thesis, human thoughts are symbolic expressions, just like describable in terms of formal logic and digital computers.

For GOFAI researchers, intelligence is building up symbolic representations and manipulating those representations according to a rule governed system. GOFAI's line of thinking bears a kinship to and is most likely inspired by Chomsky's transformational grammar. This research program dominated the AI scene until the 1990s. During this time, there have been two *AI Winters*, Margaret Boden (2008) marks the first one as the UK's winter throughout the 1970s and the second one as the American winter in the late 1980s (p. 350 and 878).[27] Boden (2008) describes both periods as the result of researchers being overly optimistic and industrialists expecting too much and states that both parties ended up with disappointments (p. 878).

Between the two AI winters, researchers developed perhaps GOFAI's greatest contribution to daily life, namely *expert systems*. Expert system describes a computer system whose capacities are directed at a domain-specific problem. As long as the ontological primitives of the domain and the relations among them are well-defined, an expert system can be developed that tracks the domain and computes rule-governed responses. Upon the success of expert systems, researchers distinguished expert systems from *artificial general intelligence* to describe the intelligence exhibited across various domains.

3.3.3. What computers can't do

Stuart Dreyfus is a mathematician and his brother, Hubert Dreyfus, was a professional philosopher focusing on phenomenology. Both were employed at MIT

[27] AI Winter is a prevalent term used to describe the period(s) in the history of AI when there have been decreased interest and funding to the field.

during the 1960s and 70s. Both were hired as consultants by the RAND Corporation to study and evaluate contemporary research on artificial intelligence. As a result of their work at RAND, Hubert Dreyfus published "Alchemy and AI" in 1965 and the brothers co-authored *What Computers Can't Do* in 1972.

"Alchemy and AI" mocked the performance of contemporary AI programs and argued that since " 'higher' forms of intelligence are necessarily derived from 'lower' forms concerned with bodily action", [Dreyfus] posited that the artificial intelligence project is impossible (Boden, 2008, p. 839). In other words, Dreyfus argued that higher-order capacities at the level of symbol manipulation can only succeed when constantly supported by a plethora of sub-symbolic capacities deriving from having an embodied presence. Boden (2008) calls this work "the most widely read RAND memo of all time" (p. 846).

What Computers Can't Do is described as an extended, book-length version of Alchemy (Boden, 2008, p. 846). It is perhaps the most impactful AI critique of all time. Hubert Dreyfus leverages Heidegger's early philosophy to explain that humans have a vital body of knowledge as a virtue of their physical engagement with the world. Furthermore, these practical skills have primacy over humans' theoretical orientation to the world, namely the orientation that includes symbol manipulation, formalisms and well-defined relations, since these theoretical concepts are derivatives of humans' practical involvement with the world. AI research, however, attributes primacy to the theoretical orientation and treats the practical engagement as a derivative thereof. This is a natural result of AI's identifying symbol manipulation as the essential component of intelligence. Hence, through this discussion, Dreyfus suggests an alternative definition to intelligence.

Dreyfus follows through with the logical extension of his thesis and argues that common sense and relevance could not be accounted for by AI. At a later work, while discussing the Cyc project,[28] Dreyfus makes the following remark:

> In the 1960s AI researchers had been optimistic. They felt confident that they could represent the few million explicit facts about the world people knew and then use rules for finding which facts were relevant in any given situation. But in the late 1970s and early 1980s AI researchers reluctantly came to recognize that, in order to produce artificial intelligence, they would have to make explicit and organize the commonsense knowledge people share, and that was a huge task. (Dreyfus, 2008, p. 17)

The issue of identifying relevance is intertwined with meaning and semantics. Dreyfus posited that this would be impossible for AI programs to do, since they are only sensitive to changes in the structure of inputs. He emphasized this distinction in his later work as well: "Since the 1960s, AI researchers had been seeking to solve the problem of getting computers, which are syntactic engines sensitive only to the form or shape of their input, to behave like human beings who are sensitive to semantics or meaning" (Dreyfus, 2008, p. 16).

Dreyfus' critique can be paraphrased as 'not all knowledge can be formalized'. Humans are primarily engaged in the world through practical activities and the theoretical involvement including concepts and relations are incomplete derivatives of this engagement. However, this balance is different for expert systems, since there is a well-defined task, an enumerable set of components, well-defined relations and no shifting context. The lack of a shift of context and the presence of a

[28] Lead by Douglas Lenat, the Cyc project has been trying to develop, since 1984, a comprehensive ontology of the world along with a knowledge base that represents all of humanity's basic knowledge about the world. Dreyfus said in 2001 that "Lenat has now spent fifteen years and at least fifteen million dollars developing CYC, a commonsense knowledge database, in the attempt to understand commonsense requests for information" (Dreyfus, 2008, p. 17). Dreyfus recites an instance where the AI software fails to adequately identify instances of "someone relaxing" and argues that the AI software faces an insurmountable problem, since it seeks to account for all of the knowledge humans have by virtue of being embodied.

formalized system allows the agent to navigate through the problem without assigning any meaning.

3.3.4 Machine learning and the dataquake

As discussed in the previous sections, GOFAI focused on symbolic representations processed via digital computation based on a set of manually prescribed rules. As sophisticated as they are, these systems are still input/output systems; they process every input according to their set of rules and produce an output. Hence, the system can be described in terms of the input-output pairs it matches. The theoretical framework and the limits of this algorithmic approach have been outlined in section 2.1.

Following the post-war technological boom, daily life has become increasingly computerized. The proliferation of the internet, the integration of computers into businesses, the rise of personal computers and the mass adoption of smartphones created a world where every person creates a previously unimaginable amount of data. Ethem Alpaydın eloquently refers to this phenomenon as *dataquake* (Alpaydın, 2016, p. 15). The abundance of data towards the end of the 20th century gave prominence to a previously underutilized method of artificial intelligence, namely machine learning.

In GOFAI, an algorithm is engaged with digitally computing and thus automatically producing the output corresponding to any given input using an input-output relation that has been coded by a human programmer according to a finite set of pre-defined rules. In contrast, in machine learning an algorithm is used to automatically modify the internal relations of a system until it “learns” to exhibit a desired behavior with sufficient accuracy.[29] In many cases, the desired behavior can be given implicitly, in the form of a "training set" that consists of input-output pairs. In these cases, the system can be said to autonomously "extract" the input-output relation from the training data.

From e-mail filters to computer vision tasks and recommendation algorithms, machine learning has solved a significant amount of contemporary technological problems in ways that would not have been feasible to address with GOFAI.

3.3.5 *Im Westen nichts Neues*

Machine learning is today referred to as 'the new AI' (such is also the subtitle of Ethem Alpaydın's book). It has made tremendous use of the vast amounts of data available and, as stated, solved some tasks that would have been too complicated for GOFAI. Thereby, it has catalyzed perhaps the greatest commercial feats of the 21st century.

However, machine learning cannot be regarded as a remedy for Dreyfus' criticism about relevance and common sense. Machine learning algorithms simply imitate the relational patterns in the training set. While GOFAI aspired to produce prescriptive rules that will hold up in every imaginable instance, machine learning aspires to gather every imaginable instance so that the algorithm can extract the aspired prescriptive rules. Hence, the machine learning analogue of the Cyc project (see 3.3.3. footnote 28) would be an endeavor to feed the immense amount of data produced by all human activity, thus supposedly documenting all human experience so that eventually even common-sense behavior can be imitated. Needless to say, such a project is neither realistic nor destined for success, exactly for the same

[29] Here, the system refers to any structure ranging from a physical device to a mathematical construct.

reasons that have troubled the Cyc project: it seeks to compensate for the lack of embodied knowledge by further increasing the amount of detail (or data) included in the system.

The hypothetical Cyc analogue is representative of machine learning's 'leap of faith', namely generalization based on statistical inferences. Since training data must be finite, the machine learning algorithm has to learn a relational pattern between the training pairs and generalize it for all potential pairs.[30] Such a generalization is typically accompanied by the problems of *overfitting* and *underfitting*.[31] In such a framework, machine learning is unable to overcome the meaning and relevance problem that had troubled GOFAI. On top of these, machine learning also faces problems inherent to the method of statistical inference from finite data. Given its rule-governed framework, GOFAI did not encounter such a problem.

If asked about machine learning, Dreyfus would point us towards the fact that a machine learning algorithm would most likely identify a goat sitting on a human's lap as a dog, whereas a human with common sense and basic life experience would probably not do this.

[30] Had there been infinite training data, training would never come to an end.
[31] Both over- and underfitting are two faces of the same coin resulting from the difficulty of formulating a termination criterion for the machine learning algorithm. Too long engagement in tuning the system to the specific training data at hand would result in perfect adaptation to these data thus in loss of generality (overfitting), while too early termination would result in a too general representation lacking in discriminative power (underfitting). The names ascribed to these two phenomena imply the possibility of a perfectly timed termination, where the machine learning algorithm would produce the correct response to all unsampled inputs. The limiting resource in this case, however, is the sampled inputs' capacity to represent all inputs, including the unsampled ones. Hence, the issue is a mereological problem rather than a timing problem; the machine learning algorithm's success is contingent on the training set's sufficiency for representing all instances.

3.4 Region of validity of the underlying premise of cognitive science

Cognitive science expands upon the premise that the human mind is an information-processor, without however giving a precise definition of the term information. Throughout various ages, in Western thinking the meaning of the term information has often been appropriated to signify a conceptualization suitable for that era. While in Aristotle's conception it refers to the process of imposing a form upon amorphous matter, in the Middle Ages the notion was used to designate the content transferred during a process of instruction or education.[32,33] Eventually, in Shannon's definition it was presented as a signal structure encoding the content transferred during communication. While cognitive science does not explicitly subscribe to any definition of information, the field's proliferation follows the emergence of Shannon's Information Theory.[34] Furthermore, George Miller, the co-founder of the first institution dedicated to cognitive science, did explicitly subscribe to Shannon's definition and lead the efforts to reform the field of psychology.

Claude Shannon's Information Theory portrays information as a precisely defined and purely quantitative measure within a well-defined communication channel. Shannon accomplishes such precise measurements by evaluating communication only in terms of syntactical metrics and then treating communication as an optimization problem. This setup's complete disregard of semantics has already

---

[32] Originally, *informare* is a Latin term and means 'to impose a form upon'. However, the term can be traced back to Aristotle's conception of causation as the imposition of form upon matter (i.e., informing matter). In Aristotle's framework, motion (*kinesis*) begins with matter getting imposed a form upon. The motion stops when the form is completely imposed and thus the entity becomes complete (*teleoin*).

[33] This conception will be examined in Chapter Four.

[34] As part of the literature survey for this thesis, five cognitive science textbooks used at various universities have been checked in order to get a sense of how this question is treated in secondary literature (the textbooks are listed in the references). In most textbooks, information-processing is explained to be the same as computation. None of the textbooks define and discuss the term "information" by itself, independent from the term "information processor". The lack of clarity about the definition of information is not compensated for with thorough discussion of information-processor. Instead, information-processor is often trivially defined as follows: "[i]nformation processors must both represent and transform information" (Friedenberg & Silverman, 2011, p. 3).

been discussed in previous sections. Since the basic components of Shannon's framework are by definition insensitive to meaning, this framework can succeed in giving an account of the human mind if and only if it can reduce the semantic elements of the human experience to syntactical elements. Indeed, one of the authors that is extensively quoted in this work, namely Jean Pierre Dupuy, criticizes cybernetics and cognitive science for extending this approach beyond the range where insensitivity to meaning can be trivially assumed. Since this approach is the legacy of cognitive science's main premise, it becomes an existential problem for the field to insist on the claim that a phenomenon can be adequately represented by solely subscribing to syntactical definitions.

Shannon had developed his framework as a remedy for a purely technical, engineering problem. When it became a theoretical orientation and an overextended framework in just a few years, Shannon was also sensitive to these developments. Hence, he wrote a short essay titled "The Bandwagon" in 1956. In this essay, Shannon says "[t]he subject of information theory has certainly been sold, if not oversold" (Shannon, 1956, p. 3). He then warns that information theory had become a scientific bandwagon, mainly due to its connections to fashionable fields such as computing machines, cybernetics and automation (Shannon, 1956, p. 3). Shannon also explicitly questions his framework's relevance to psychology. He says: "workers in other fields should realize that the basic results of the subject are aimed in a very specific direction, a direction that is not necessarily relevant to such fields as psychology, economics, and other social sciences" (Shannon, 1956, p. 3). In his final words, Shannon invites everyone to raise their critical thresholds, engage in scientific rigor and maintain skepticism amid overwhelming popularity.

Given that contemporary textbooks do not consider it necessary to include a rigorous definition of information in their account of cognitive science, Shannon's reservations appear to have been justified. Moreover, this bandwagon arguably has brought forth a theoretical orientation, where mainstream cognitive science researchers expect to redefine everything related to human subjectivity as a well-defined information-processing problem and thereby reduce metaphysics to physics. Indeed, Patricia Churchland explicitly embraces and declares this approach in her book *Brain-Wise: Studies in Neurophilosophy*:

> [M]etaphysical questions are best *re*characterized as those questions where scientific and experimental progress is not yet sufficient to found a flourishing explanatory paradigm. This implies that 'metaphysical' is a label we apply to a *stage*—an immature stage, in fact—in a theory's scientific development, rather a distinct subject matter with distinct methods. Until rather recently, theories about the self, consciousness, and free will, for example, were at a very immature stage, since neuroscience and cognitive science were not sufficiently advanced to get very far in addressing these matters experimentally. Because of this relative immaturity, these topics may still be regarded as metaphysical, but when scientific success comes, that status will eventually be cast off as uninformative and burdensome. (Churchland, 2002, p. 39 – 40, emphasis Churchland).

Churchland's approach explicitly declares that any metaphysical problem is destined to be evaluated in terms of physics. Hence, Churchland boldly claims that cognitive science is destined to recast questions regarding self, consciousness and free-will in purely material terms. Given the field's main premise, those physical explanations are envisioned to be in terms of information-processing. As a result, Churchland foresees a physicalist framework accomplished by naturalizing all mental matters as neuronal, information-processing problems.

This physicalist orientation is the target of Jean Pierre Dupuy's critique of the field laid out in his book *The Mechanization of Mind: On the Origins of Cognitive*

Science.[35] As discussed in a previous chapter, Dupuy gives in his work an extensive account of the Macy Conferences and studies cybernetics' kinship to cognitive science in his book.

Dupuy's criticism can be summed up as follows: "[i]n inventing a type of transcendental inquiry that did away with the subject, cybernetics was to greatly assist the deconstruction of the metaphysics of the subject." (Dupuy, 2009, p. 107). Thus, the information-processor premise motivates approaching metaphysically loaded matters of subjectivity as "emergent effects produced by the functioning of subjectless processes… proceed[ing] from …a complex network of interactions among simple entities – formal neurons in the case of the individual quasisubject" (Dupuy, 2009, p. 160). Hence, cognitive science, through the information-processing premise predisposes the scientist to try to explain away any metaphysical conceptualization of subjectivity.

As Dupuy states, the information-processing perspective is also indebted to cognitive science's conceptualization of neurons. As discussed in section 3.2, cognitive science describes the neuron primarily in terms of the action potential. The idealized neuron, whose behavior can be expressed completely in terms of two-valued logic, created a very fruitful tool. However, modeling the nervous system in terms of these modular, atomistic building blocks failed to account for the complexity of the human nervous system. Dupuy makes a note of this divergence when recounting a discussion from one of the Macy conferences. He says "[a]t the seventh conference, it was the very notion of a model that became progressively more blurred – unsurprisingly, to the extent that it was usually supposed to involve a hierarchical relationship between the original and a reproduction" (Dupuy, 2009, p.

[35] This thesis problematizes the "physicalist" perspective but does not do so with the purpose of advocating any of its traditional alternatives, namely dualism and mentalism.

139). Dupuy notes that, over time, the Macy group became more concerned with what the model is capable of rather than developing the model to be equivalent to the phenomenon it seeks to represent. He concludes by saying "[t]he world and its representations thus found themselves flattened by logic. In this respect the cybernetic models anticipated poststructuralism: they were only models of themselves, or else of other models – mirrors of mirrors, speculums reflecting no reality beyond themselves" (Dupuy, 2009, p. 139).

Nevertheless, there exist some problem domains where the information-processor approach can be safely applied, namely domains that are insensitive to meaning and relevance. Indeed, all widely acknowledged accomplishments of cognitive science coincide with these domains. Perhaps the greatest feat of the information-processor approach has been about sensation and perception. David Marr's work on vision has created a generation of cognitive scientists studying the visual system. As part of this research, Marr has developed his famous *tri-level hypothesis*, his recipe for analyzing information-processing systems. These three levels are as follows:

- *Computational theory* corresponds to the goal of computation and an abstract definition of the task.
- *Representation and algorithm* is about how the input and the output are represented, and about the specification of the algorithm for the transformation from the input to the output.
- *Hardware implementation* is the actual physical realization of the system. (Alpaydın, 2016, p. 25)

By distinguishing these three levels, Marr emphasizes the fact that the relationship among these levels is not necessarily one-to-one for any given system; a given computational theory can be represented by various different representations and, likewise, a given representation and algorithm can have various hardware implementations or a computational theory can be carried out by various algorithms and implementations. Like all previously discussed frameworks, Marr's conception of vision is only sensitive to signal structures and does not evaluate the meaning or relevance of the signal. Yet, Marr's framework has achieved a high degree of success and no other modality has been deciphered to the extent that vision has been. To better understand the reason underlying this success, the unique status of vision in the human experience must be evaluated.

As Matt Cartmill outlines in his *visual predation hypothesis*, primates, including humans, have evolved their physical traits to better prey on small insects and other small creatures. This includes the evolution of the postorbital bar, having an opposable thumb, nails rather than claws, binocular vision and a neocortex that is primarily dedicated to vision rather than the other senses.[36] Hence, vision is the most foundational cognitive modality humans have. In the course of evolution, the lower levels of this capacity—such as color vision, edge detection, motion detection etc.—have gained a genetically well-preserved robust algorithmic structure embedded in humans' sense organs and neural architecture. Hence, these aspects are already inferred from the form of the sensory input. This fundament provides a secure basis for the higher level functioning of the visual capacity, where context-dependence, subjectivity, free and open-ended interpretation are at work. Needless to say, the functioning of these levels does not lend itself to modeling as an algorithmic process.

Moreover, these higher levels of vision can already be considered within the realm of cognition. The higher levels of vision share their properties with other mental processes such as attention allocation, contemplative thinking and decision-making that appear in the midst of a never-ending hermeneutic circle that relies on a

[36] The arched bone below the eye socket that protects the eye.

subject's viable engagement with the world.[37] The distinction of these capacities in their role in open-ended interpretations will be addressed comprehensively in the context of a Heideggerian framework in the next chapter.

Given the robust structure of the lower levels of vision and their insensitivity to relevance, it is not surprising that the information-processing framework has had most of its success with vision. Churchland seeks to extend this framework to metaphysical matters of subjectivity that are by definition evolutionarily recent and, therefore, not robust against perturbations emerging from a subject's engagement with the world. However, these metaphysical matters of subjectivity are dependent on the human ability to identify relevance and attribute meaning. Yet, as long as cognitive science builds upon the information-processor premise, it has to insist on the claim that all such capacities are robust and not sensitive to meaning like vision. With this attitude the field and its theoretical framework motivates the researcher to explain away dependencies on meaning and subjectivity as epiphenomena. Consequently, cognitive science, insofar as it is grounded by the information-processing premise, tries to reduce the human being to the status of an algorithmic machine. Perhaps as a progressive step, the field can distinguish modalities insensitive to meaning from those that are and thereby define boundary conditions for the applicability of the information-processor premise, while developing complementary models for modalities sensitive to meaning and relevance.

[37] As it is discussed in the next chapter, since Dasein is the entity where an understanding of being erupts into being, Dasein's engagement with entities contributes to Dasein's existential constitution by transforming Dasein's understanding of being. Likewise, Dasein's existential constitution predisposes Dasein to encounter entities in a delimited manner. This bidirectionally unfolding interpretation makes up the hermeneutic circle.

CHAPTER 4

HEIDEGGER AND COGNITIVE SCIENCE: *BEING AND TIME* AND BEYOND

It has been laid out in the previous chapter that AI researchers took up Heidegger's work in the context of Dreyfus' criticism of symbolic AI. While outlining his criticism, Dreyfus did not consider himself as listing the practical problems facing symbolic AI. Instead, he posited that symbolic AI researchers were "turning rationalist philosophy into a research program" (Dreyfus, 2012, p. 63) and considered symbolic AI to be the sum of "rationalism, representationalism, conceptualism, formalism and logical atomism" (Dreyfus, 2012, p. 63).[38] Hence, Dreyfus responded as he would to a philosophical line of thinking. Since the philosophical concern at hand deals with the attempt at a true characterization of the human mind, this discussion concerns the project of cognitive science as a whole.

Symbolic AI researchers subscribed to a Cartesian worldview, advocating that the world consists of a set of meaningless facts and that intelligence consists of assigning well-defined relations between well-defined entities, using symbols to represent them and tracking them over time. Dreyfus, in such a context, used Heidegger's work to argue that the intelligibility of people, their practices and institutions do not simply owe to an enumerated list of facts and relations we maintain. Instead, intelligibility derives from our human capacity to have a sense of pragmatic context and thereby assign meaning and relevance to entities. Dreyfus refers to this unreflective but engaged disposition as *skillful coping*. Furthermore,

[38] Dreyfus states that AI researchers "had taken over Hobbes' claim that reasoning was calculating, Descartes' mental representations, Leibniz's idea of a 'universal characteristic' (a set of primitives in which all knowledge could be expressed), Kant's claim that concepts were rules, Frege's formalization of such rules, and Russell's postulation of logical atoms as the building blocks of reality" (Dreyfus, 2012, p. 63).

Dreyfus emphasizes that our capacity to assign meaning and relevance owes to our embodied existence and the knowledge that derives from it.

Dreyfus' Heideggerian critique has influenced a generation of AI researchers. Indeed, there are cherished stories of renowned AI researchers like Terry Winograd teaching Heidegger's work in his computer science courses to Larry Page (Dreyfus, 2008, p. 21). Yet in such stories and in cognitive science circles, Heidegger is almost always considered solely with reference to *Being and Time*. Changes in Heidegger's approach and developments in his thinking after *Being and Time* are not taken into consideration by most cognitive scientists, most probably because these developments are deemed to be irrelevant to cognitive science's adoption of Heidegger.

Published in 1927, *Being and Time* is Heidegger's first major published work. Initially envisioned as two major parts each consisting of three divisions, *Being and Time* is an incomplete project. Only the first two divisions of the first part were published. As pointed out by Sheehan (2010), Heidegger described and adopted a transcendental framework for his project that served him well in the first two divisions, where he laid out how human beings projectively hold the world open for meaning-making. However, in division three where he had planned for a reversal of direction to give an account of world's relationship to human being, his transcendental framework collapsed (Sheehan, 2010, p. 89). Heidegger then adopted an approach he called *being-historical* (*seinsgeschichtlich*), emphasizing the historicity of meaning.

A Heideggerian account of cognitive science focused exclusively on the first two divisions of *Being and Time* fails to consider the ramifications of the change in his thinking for cognitive science. Moreover, they disregard what Heidegger's later approach may imply for cognitive science. Heidegger's later work and its implications about cognitive science remain understudied.

In this thesis I hope to develop a criticism of cognitive science based on a later work of Heidegger, namely his critique of modern technology. This chapter will briefly touch on *Being and Time*'s take on human being's relationship with equipment and its implications for cognitive science as well as the implications for the aforementioned technological domains. Then, it will attempt to develop a criticism of cognitive science based on Heidegger's essay "The Question Concerning Technology".

4.1 *Being and Time* and cognitive science

Being and Time is an inquiry into the meaning of being. Heidegger uses the term *Dasein* to refer to the experience of being that is peculiar to human beings. Although the term *Dasein* denotes the human being, it is also intended to carry a different connotation than terms such as 'self', 'soul' or 'subjectivity'. As a being that cares for and seeks an understanding of being, *Dasein* reflects the human being as the 'there' (*Da*) of 'being' (*Sein*). It is the place where the understanding of being occurs and the understanding of being emerges as being. One of Heidegger's main claims in *Being and Time* is that our primary relationship to the world is handy and practical. *Dasein*'s proximal encounter with entities reveals them primarily as equipment that can be employed to achieve other ends, instead of revealing them as abstract objects, i.e., substances that are to be represented on a geometrically ordered coordinate system. Such theorizing attitude, Heidegger argues, is a derivate of our practical encounter. This disposition makes Heidegger's approach anti-Cartesian in character. Indeed, Heidegger's position can be paraphrased as 'I am, and only therefore do I think'. As opposed to conceiving the human being as a free-floating self or a subject accessing a world of objects, Heidegger argues that Dasein is being-in-the-world.[39] Dasein's being and that of the world are inextricably linked. The world is a part of Dasein's fabric of existence. These two interconnected distinctions have been the backbone of most Heideggerian critiques of cognitive science. The upcoming sections will attempt to characterize the primacy attributed to practicality over the theorizing attitude in *Being and Time* and then consider this distinction in the context of the technological discussion laid out in the earlier chapters.

4.1.1 The Primacy of the *ready-to-hand* in *Being and Time*

In *Heidegger's Confrontation with Modernity*, Michael Zimmerman states: "*Being and Time* makes clear that instrumental activity is the basic way of being-in-the-world" (Zimmerman, 1990, p. 152). In his work, Heidegger attributes primacy to the activity of manipulating and producing things and states that "the sole alternative is to treat them abstractly: either as objects for ordinary curiosity or as objects for scientific curiosity" (Zimmerman, 1990, p. 152). Thus, in *Being and Time*, Heidegger posits that human beings engage with the world either practically or theoretically and that the former has primacy.

Heidegger's approach to human beings' engagement with entities is revealed through his analysis of the pre-industrial workshop and handicraft. From a craftsman's perspective, the workshop consists of a network of relationships that contain an interpretation of every tool and product. Every entity is encountered as equipment and the entity always occurs within an equipmental-referential context; in

[39] Heidegger's term *being-in-the-world* is intended to capture the activity of existing. The neologism is intended to facilitate a contrast with the line of thinking that gives a representational account of existing, i.e., objects being spatially located with respect to other objects. The term reinforces *Dasein*'s immersion in its world.

the daily craftsmanship experience, there is no such thing as an isolated tool. Moreover, the craftsman's interpretation of his equipment is not a cognitive activity that humans sometimes perform. Instead, interpretation is an ever-unfolding practice that is the result of Dasein having a *preontological understanding* of entities and having a world.[40,41]

Heidegger characterizes our practical relation to tools as *readiness-to-hand* (*Zuhandenheit*). In using a hammer, the craftsman treats the tool as an extension of his hand and does not notice the hammer as a self-standing object. The manipulability of the hammer reveals itself to the craftsman through the craftsman's using of it. While using the hammer, the craftsman does not take on the role of a detached observer and treat the tool as an object with properties analyzable by scientific inquiry. Instead, throughout this process, the tool disappears in favor of the work being done, and thus remains transparent as an extension of the craftsman's hand.

Tools cease to be transparent when, for some reason, they cannot be relied on for fruitful labor.[42] This situation invokes a different type of orientation towards entities; "when a breakdown of some sort occurs in the work activity, the work world suddenly becomes illuminated in a way that it is not when engaged in working" (Zimmerman, 1990, p. 139). Heidegger refers to this orientation as *presence-at-hand* (*Vorhandenheit*). This orientation sets up the user to examine the tool as an abstract object with reference to relevant facts and concepts.

---

[40] That is, an understanding of entities prior to and ripe for an ontological inquiry into the being of beings

[41] In this context, "world" is to be understood as the aforementioned equipmental-referential totality, i.e., the structure of reference relationships constituted by and for the human being that enables entities to reveal themselves as equipment (Heidegger, 1996).

[42] This does not necessarily have to be due to the malfunctioning of the tool. It could simply happen because of the tool being missing or being unfit for the task it is adopted for.

Cartesian philosophy has presumed that entities are primarily present-at-hand for us, i.e., the human being is essentially a thinking subject accessing a world of objects through theoretical contemplation and, therefore, entities can only become tools as a consequence of our theoretical orientation. Heidegger, in contrast, argued that this is the exact opposite of the true situation; Dasein is first and foremost engaged in everyday life and only by an act of abstraction can Dasein take on the role of a detached theorizing observer and begin to consider equipment as objects that are analyzable through scientific procedures as abstract entities.

4.1.2. "Erzählbar is not zählbar": Dreyfus' Heideggerian critique of AI and its implications for cognitive science

The technological feats outlined in chapter 2 are feats of general purpose modelling: Turing created a framework where logical machines of high computational complexity represent sets of relations among members of a list of entities, Shannon invented a formalism that quantifies the amount of novelty in a message and cybernetics created mechanisms that can be relentlessly employed for error-reduction. Each model can capture a large group of phenomena of interest by doing algebraic operations on mathematical constructs.

After the Second World War, these theories were adopted by a diverse set of scholars to develop a science of the human mind. As discussed in the previous chapter, a large part of this group considered the human mind as a physical symbol system, arguing that humans build up symbolic representations of the world and perform computations on these representations.

Dreyfus says his intention in emphasizing the primacy of *ready-to-hand* and crafting his critique around *Being and Time* was to convey the important insight that "being-in-the-world is not representational at all" (Dreyfus, 2012, p. 74). This implied that computational models of the mind that relied on formalisms and symbol manipulation were categorically wrong, since they foresaw an agent that theorized itself and the world in terms of well-defined concepts and comprehended objects as abstract entities. In Heideggerian terminology, these researchers had attributed primacy to the present-at-hand and considered the ready-to-hand to be derivative.

The discussion of the *frame problem* presents a domain where the approach attributing primacy to the present-at-hand and considering intelligence to be symbol manipulation run into problems. The technical frame problem, loosely stated, describes an issue that emerges when a system that has formal representations of its environment proceeds from one situation to the next.[43] How does the system know which representations to update and determine the ones that go unchanged? The frame, in this context, lays out all the relevant facts of a situation in need of revision. Initially originated as a technical difficulty that ought to be surmounted in the context of computer science, the same issue has been later problematized as a problem of epistemology: how do human beings know which beliefs about the world to revise following any given experience?

Without any reference to relevance and meaning, the philosophical frame problem appears to be insurmountable. Since a system of frames is not *in* a situation, Dreyfus states this approach is bound to create a regress of frames where there are frames for recognizing relevant facts for recognizing relevant frames for recognizing relevant facts *ad infinitum*. "The frame problem wasn't just a problem but was a sign

[43] McCarthy and Hayes formally define a situation as the complete state of the universe at an instant of time. (McCarthy and Hayes, 1969) They further note that, since the universe is too large for description, one should never completely describe it, instead only give facts about it. At this moment, there is an arbitrary renunciation of information and, arguably, the seeds of this perspective's own overcoming at the philosophical level (also see Fodor's problematization of informationally-encapsulated systems).

that something was seriously wrong with the whole approach" (Dreyfus, 2012, p. 64).

Contemporary Heidegger scholar Byung-Chul Han distinguishes what can be *recounted* (*Erzählung*) from what can be *counted* (*Zählung*). In resorting to well-defined relations among members of enumerable sets, evaluating entities on a purely quantitative basis and having disregard for meaning and relevance, cognitive science currently exists in the realm of what can be counted and remains distant to what can be recounted. In doing so, an integral part of the human mental experience remains inaccessible to cognitive science: "Numbers do not recount anything about the self. Counting is not recounting. A sense of self derives from giving an account. It is not counting, but recounting that leads to self-discovery or self-knowledge" (Han, 2017, p. 60). Han emphasizes this difference again when he distinguishes a narrative from a timeline. This difference for Han is intimately tied to the fact that humans forget, whereas memory as stored data persists in endless addition and accumulation: "Stored data admit counting, but they cannot be recounted. Storage and retrieval are fundamentally different from remembering, which is a narrative process. Likewise, autobiography constitutes a narrative: it is memorial writing. A timeline, on the other hand, recounts nothing" (Han, 2017, p. 67). The fundamental mistake Dreyfus mentioned stems from the fact that, in its current form, cognitive science seeks to reduce what can be recounted to what can be counted. Being-in-the-world, encountering entities as equipment, establishing a sense of relevance and meaning can only occur in the context of a narrative. In *Being and Time*, Heidegger deliberately emphasizes that equipmental relations "resist any sort of mathematical functionalization" (Heidegger, 1996, p. 122) and characterize our mode of being. Yet, cognitive science deliberately seeks to cast this narrative as something to be counted, not recounted. The underlying reason, according to Heidegger's essay on modern technology, stems from the character of modern technology and the relationship between technology and science.

4.2. Heidegger's critique of modern technology and its implications for cognitive science

Dreyfus appealed to Heidegger's account of handiness in *Being and Time* in order to criticize the account of mind produced within a Cartesian representational framework where our experience is appropriated as a subject accessing a world of objects. In contrast, in "The Question Concerning Technology" Heidegger states that humanity is now in a post-representational framework where both subject and object are absorbed into a technological-industrial complex as resources to be ordered.

Heidegger distinguishes modern technology from ancient technology in his essay "The Question Concerning Technology". He considers the beginning of modern technology to be machine-power technology, developed in the second half of the eighteenth century. Modern technology differs from ancient technology in the sense that the former seeks to extract and exploit its resource whereas the latter does not. Heidegger's example for ancient technology is an ancient windmill: it does draw energy from the wind but does not store it, whereas a modern wind turbine stores the energy produced at a moment when it is not needed to offer it on demand later on (Heidegger, 1977, p. 14). In his essay, Heidegger seeks to problematize modern technology.

Heidegger does not criticize particular items of technology or issues arguably caused by technology such as global warming or nuclear war. Rather, Heidegger conceptualizes technology as a mode of being and revealing.[44] The technological mode of revealing is uniquely dangerous, because "it drives out every other possibility of revealing" (Heidegger, 1977, p. 27). To understand the technological mode of revealing, Heidegger enquires into the essence of technology, which is itself not technological.

Heidegger's essay "The Question Concerning Technology" argues that technology employs science to fulfill its essence. By conceptualizing nature as a collection of calculable forces, science appropriates an understanding of technology as "applied physical science" (Heidegger, p. 23, 1977). Heidegger calls this conception of technology a "deceptive illusion" (Heidegger, p. 23, 1977) and seeks to present technology as a mode of revealing.

Next, this chapter will address these issues while also touching upon Heidegger's conception of productionist metaphysics, human essence, freedom and *destining* in this essay. Then, Heidegger's criticism will be considered in the context of cybernetics and cognitive science. It will be further argued that, in its current condition, cognitive science reflects the manner in which technological humanity finds itself under the threat of being reduced to the status of *standing-reserve* (*Bestand*), a mere resource whose only characteristic is its instrumental utility in advancing a task.

---

[44] The Ancient Greek word *aletheia* is traditionally translated as truth. Heidegger translates *aletheia* as unconcealment (*Unverborgenheit*). Heidegger considers the interplay of presencing and absencing (unconcealing and concealing) as the characterization of a primordial sense of truth. In this context, the Greek term *alétheuein* describes the act of unconcealing. Heidegger translates this term to German as *entbergen*. *Bergen* can be translated as to rescue, to shelter, to hide and salvage something and the prefix *ent-* connotes a change from an existing situation. In English translations, *entbergen* is translated as revealing. As a mode of being, technology bounds us in our revealing practices and how we encounter our world (Heidegger, 1977, p. 11).

4.2.1 The Janus-headed nature of technological humanity

Heidegger's essay "The Question Concerning Technology" declares that in the technological age, humanity finds itself in a rather peculiar position; humanity is threatened by the possibility that humanity itself will be "taken as standing reserve" (Heidegger, 1977, p. 27), and under this threat humanity "exalts [it]self to the posture of lord of the earth" (Heidegger, 1977, p. 27). To better understand the source of this peculiarity, one must grasp the essence of technology.

The essence of technology is itself not technological (Heidegger, 1977, p. 4). Heidegger calls this essence *enframing* (*Gestell*) and describes it as the "challenging claim which gathers man thither to order the self-revealing as standing reserve" (Heidegger, 1977, p. 19). Thus, the essence of technology, above all, challenges us to relate to our surroundings in a specific manner. This challenge dictates us to consider everything in the world as standing reserve, raw materials that do not have any intrinsic value or independent existence. Furthermore, we are challenged to further order this collection of raw materials constantly and restlessly.

The challenge to consider everything as raw material dictates a nihilistic conception of the world, since everything becomes levelled down to an equal plane where nothing is different.[45] After all, when there is no intrinsic aspect to it, every resource ultimately becomes interchangeable. As it will be elaborated in a later section, Heidegger considers this nihilism to be the product of the history of Western Philosophy since Plato. This nihilism homogenizes our interpretation of beings by regarding them as mere resources and thereby prepares the ground where humans themselves become considered as a mere resource. Indeed, Heidegger emphasizes

[45] Heidegger does not explicitly use the term "nihilistic" to describe the technological conception of beings. This characterization originates from Karl Löwith's – a student of Heidegger's – reading of his works. Löwith's reading is also echoed by the translator of the version of the essay quoted throughout this thesis, as indicated by the translator's remarks in his introduction.

this by saying: “If man is challenged… then does not man himself belong even more originally than nature within the standing-reserve? The current talk about human resources, the supply of patients for a clinic, gives evidence of this” (Heidegger, 1977, p. 18). Thus, humanity can now be regarded as a resource for the advancement of the clinic, just like electricity, beds and operating rooms.

The consideration of everything as ordered raw material also cultivates the ground for humanity to consider itself as that which brought order to the chaos of raw material. Hence, humanity comes to think of itself as the lord of the earth who is the creator of all and does not feel indebted to any other being for its creative capacities. Consequently, “the impression comes to prevail that everything man encounters exists only insofar as it is his construct” (Heidegger, 1977, p. 27). Heidegger, however, states that this impression is a mere illusion (Heidegger, 1977, p. 27). In order to understand the source of humanity’s peculiar situation, three issues have to be explained, namely Cartesian subject-object duality, the post-Cartesian framework of enframing and the authentic mode of revealing upheld by Heidegger. These issues will be discussed with reference to Michael Zimmerman’s related work.

4.2.2 Productionist metaphysics, the post-Cartesian framework and authentic production

In his book *Heidegger’s Confrontation with Modernity*, Zimmerman compiles a critique of *productionist metaphysics* based on Heidegger’s works and lectures. This term captures Heidegger’s conviction that being has been obscured and reduced to the status of a kind of entity throughout the history of philosophy, starting with Plato. As a result, any possibility of an ontological difference between being and entities is eradicated. Zimmerman explains further: “[p]roductionist metaphysics conceived of making in terms of ‘actualizing’ or ‘effecting’ a thing, in the sense of ‘causing’ it to be present” (Zimmerman, 1990, p. 223). Heidegger considers both the Cartesian representational framework and the post-representational world of enframing as successive periods within the history of productionist metaphysics. In contrast, he considers the work of the pre-Socratics as having an originary and non-reductionistic understanding of being and revealing (Zimmerman, 1990, p. 223).

In the medieval era, the productionist metaphysics framework fostered the conviction that “God became identified with the being of entities, i.e., the self-grounding Creator who produced all creatures” (Zimmerman, 1990, p. 171). In this framework, the cosmos revealed itself as a craftwork produced by God. The image of God as a primordial craftsperson is also mentioned in Heidegger’s essay. He says, “where everything that presences exhibits itself in the light of a cause-effect coherence, even God can, for representational thinking, lose all that is exalted and holy, the mysteriousness of his distance. In the light of causality, God can sink to the level of a cause, of *causa efficiens*. He then becomes, even in theology, the god of the philosophers, namely, of those who define the unconcealed and the concealed in terms of the causality of making” (Heidegger, 1977, p. 26). Thus, in the history of productionist metaphysics, prior to the Cartesian turn, entities exist only as the products of God’s work and, hence, their being is appropriated as forms God imposed on the primordial amorphous matter.

As Zimmerman (1990) explains, the medieval era is succeeded by the Cartesian turn which sets up humans to interpret the world as a picture (*Bild*), whose reality was determined against the standards of the cognizing subject (p. 191). Object (*Gegenstand*), in such a framework, is merely a thing placed before the subject, a representation (*Vorstellung*). Therefore, entities now exist as objects insofar as they are being cognized by a human subject. Consequently, "objective knowledge" is a means for humans to gain control of objects. Since the kind of knowledge that allows sufficient precision for control is mathematical representationalism, objects are now thought of as lumps of extended matter best represented in a geometrically ordered coordinate system.

Heidegger argues in his essay that we are now at a stage beyond the Cartesian turn. He says, "what is unconcealed no longer concerns man even as object, but does so… as standing-reserve" (Heidegger, 1977, p. 26). As noted above, the Cartesian turn allowed humans to interpret the world as a picture in which objects were placed against the subject for human domination. The German word for object, Gegenstand, shares roots with the word for standing-reserve, Bestand. This morphological kinship is most likely meant to emphasize the historical progression. What is placed before the human being is no longer an object in the Cartesian sense. Rather, it is an extracted and dominated resource that can be called upon at will. This way of ordering resources entices the human being to accept that "everything man encounters exists only insofar as it is his construct" (Heidegger, 1977, p. 27). His self-ascribed role is comparable to the medieval God of productionist metaphysics. This leads to the delusion that "man everywhere and always encounters only himself" (Heidegger, 1977, p. 27).

In contrast to the productionist metaphysics paradigm, Heidegger advocates an authentic mode of revealing that he traces back to pre-Socratic thinkers such as Heraclitus and Parmenides. To introduce the authentic mode of revealing, Heidegger traces the term 'technology' to its Ancient Greek root *technē*. He emphasizes the dual nature of the term technē: the term invokes both the term *poiēsis* and the term *episteme* (Heidegger, 1977, p. 13). Poiēsis means both poetizing and producing; it is a mode of disclosure that enables entities to shine forth and manifest themselves. Episteme means to know something, and to be an expert and entirely at home in it (Heidegger, 1977, p. 13). Heidegger explains the character of technē with the following paragraph:

> Whoever builds a house or ship or forges a sacrificial chalice reveals what is to be brought forth…This revealing gathers together in advance the aspect and the matter of ship or house, with a view to the finished thing envisioned as completed, and from this gathering determines the manner of its construction. Thus what is decisive in technē does not lie at all in making and manipulating, nor in the using of means, but rather in the aforementioned revealing. It is as revealing, and not as manufacturing that technē is a bringing forth. (Heidegger, 1977, p. 13)

Thus, an agent's mere forcing a form upon matter by brute force to manufacture a product does not qualify as authentic production. Instead, poiēsis as production corresponds to disclosing an entity for its own sake and letting it come forth on its own. This capacity for authentic production rests upon episteme, which includes foresight, a deep understanding and appreciation of the materials, artisanship, and engagement with the product for its own sake. Likewise, poiēsis, understood as poetizing, discloses a world, reveals in advance what has not yet shone forth – comparable to the artisan's foresight with the artifact. Heidegger's conception of authentic production is expanded upon in a later section with the discussion of thinking and *logos*.

In contrast to Heidegger's notion of authentic production, enframing challenges us to conceive everything in terms of calculative rationality. Under the reign of enframing, the production process is a mere steering of resources planned to serve a purpose in the technological-industrial complex. 'Disclosing an entity for its own sake' or 'letting an entity shine forth' is incompatible with the challenge of enframing. In order to envision a path towards authentic production from the reign of enframing, one needs to consider Heidegger's notions of freedom, human essence and *destining*.

4.2.3 The way out

Heidegger's understanding of technology sharply contrasts with the common views of his time. He enumerates these definitions when he says, "[t]he current conception of technology, according to which it is a means and a human activity, can therefore be called the instrumental and anthropological definition of technology" (Heidegger, 1977, p. 5). Although correct, these definitions cover up the essence of modern technology by presenting technology as a tool human beings can manipulate and master to achieve self-determined ends. In contrast, as Heidegger's definition of enframing suggests, it is the essence of technology that challenges us to approach the world as a mere resource by disclosing a background of meaning that appropriates specific places and contexts for tools. This essence, Heidegger claims, is a threat to the human being's freedom and, therefore, is the "supreme danger" (Heidegger, 1977, p. 26).

Heidegger's understanding of freedom also contrasts with the conceptions of his time. Heidegger highlights this contrast when he says "[t]he essence of freedom is *originally* not connected with the will or even with the causality of human willing" (Heidegger, 1977, p. 25, emphasis Heidegger). Thus, he does not consider freedom as the ability to make unimpeded choices or the nature of the will. Instead, he puts forth the following definition: "Freedom is that which conceals in a way that opens to light, in whose clearing there shimmers that veil that covers what comes to presence of all truth and lets the veil appear as what veils" (Heidegger, 1977, p. 25). Thus, for Heidegger, freedom lies with recognizing that one's mode of being and revealing practices are contingent on a specific interpretation of Being. Recognition of that contingency allows one to inhabit one's contemporaneous mode of being as "an ordaining of destining, as is every way of revealing" (Heidegger, 1977, p. 24). As a result, one is able to establish a free relation to the technological mode of revealing.

Prior to discussing Heidegger's conception of freedom further, to better understand how modes of revealing hold sway over humanity, another term has to be introduced, namely *destining* (*Geschick*). Heidegger defines *destining* as "that sending-that-gathers [*versammeldes Schicken*] which first starts man upon a way of revealing" (Heidegger, 1977, p. 24). Hence, destining ordains humanity the modes of revealing that have reigned throughout history. Enframing is one of these modes of being, as is poiesis, the mode of being contrasted with enframing throughout the essay.

Heidegger considers the freedom of human beings in relation to their belonging to destining. He says, "man becomes truly free only insofar as he belongs to the realm of destining and so becomes one who listens and hears [*Hörender*], and not one who is constrained to obey [*Höriger*]" (Heidegger, 1977, p. 25). Thus, as portrayed by the metaphor of the 'veil', Heidegger argues human beings can be free, when they listen to and hear the call of *destining* as a gathering for them to inhabit the contemporary mode of revealing. If they consider the call of destining as enframing the only possible mode of revealing, they are constrained to obey and, therefore, will no longer be free.

This capacity to hear and listen to the call of destining defines human essence, namely their "needed belonging to revealing" (Heidegger, 1977, p. 26). When the call of destining as revealing is understood as a constraint for human beings, revealing is only understood in terms of regulating and securing the standing-reserve within a larger technological-industrial complex. When a human being "does not apprehend enframing as a claim… he fails to see himself as the one spoken to… hence also fails… to hear in what respect he ek-sists, from out of his essence" (Heidegger, 1977, p. 27). Like the aforementioned contemporaneous definitions of technology, when human beings consider technology as a means or their own activity, they fail to understand that technology is the result of a call of destining that gathers them to interpret the world according to the technological mode of being. Instead, human beings mistakenly believe that technology is the result of their struggle to further transform the products of their own making with the purpose of improving the technological-industrial complex. Because of this instrumental understanding of technology, human beings do not grasp the technological mode of revealing *as* an ordaining of destining. As a result, they are not free and are alien to their essence.

4.2.4 Technology and science

Heidegger further argues in the essay that technology employs science to fulfill its essence. This is made possible by enframing challenging modern science to conceptualize nature according to its needs. Since enframing needs the human being to order the self-revealing as standing-reserve, nature needs to be modelled as a set of orderable resources. As a result, enframing challenges science to represent nature as "identifiable through calculation and orderable as a system of information" (Heidegger, 1977, p. 23).[46]

[46] The above quotation is taken from the (Heidegger 1977) translation by Lovitt, nevertheless I suggest that a translation as "a system composed of pieces of information" better reflects the German original "System von Informationen".

Modern science, as a result, becomes bounded by rigid mathematical representationalism. Heidegger posits that “[m]odern science’s way of representing pursues and entraps nature as a calculable coherence of forces” (Heidegger, 1977, p. 21). He makes his case in terms of physics, the paradigmatic science since modernity. He says, “[b]ecause physics…as pure theory, sets nature up to exhibit itself as a coherence of forces calculable in advance, it therefore orders its experiments precisely for the purpose of asking whether and how nature reports itself when set up this way” (Heidegger, 1977, p. 21). Of course, what can be calculated and predicted in advance can also be steered for further ordering as standing-reserve. Hence, considered in this light, science can be regarded as a mere means for technology to fulfill its essence.

Heidegger further expands upon science’s status in the technological age in his 1966 *Spiegel* interview and 1967 lecture. His lecture, as cited by Zimmerman, builds upon Nietzsche’s following remark: “[t]he victory of *science* is not what distinguishes our 19th century, but instead the victory of scientific *method* over science” (as cited in Zimmerman, 1990, p. 200, emphasis Nietzsche). Rather than being subservient to science, method takes the place of science’s intellectual concerns.[47] As a result, ‘real’ gets equated with ‘calculable in mathematical terms’. Since method concerns itself with what is identifiable through calculation, science now deals only with what is orderable as standing-reserve. This is evident in Heidegger’s following remark: “method is the victorious challenging of the world upon a thoroughgoing availability for man” (as cited in Zimmerman, 1990, p. 200). The challenging is victorious because method assumes the calculability of its object as a premise. Thus, scientific enterprise begins to flourish on the ground of

[47] These concerns and the way of concerning perhaps will become apparent if we consider a past era in which science is “natural philosophy”.

calculability. What can be calculated can be ordered as standing reserve. As a result, any given scientific success is more means for technology fulfilling its essence.

4.2.5 Technology, cybernetics and *logos*

In the same lecture, Heidegger states that this victory of method achieves its uttermost possibility in cybernetics. The cybernetic approach claims that all world processes are steering processes; all such processes are based on information exchanges between a finite set of actors whose governing principles of behavior can be quantified. Thus, according to cybernetics, all world processes are calculable. Hence, cybernetics is mathematical representationalism in its highest form, since it perceives the world "orderable as a system of information" (Heidegger, 1977, p. 23). With the help of informational feedback, one process can be steered through another. In this spirit, Heidegger reminds the audience that the etymological root of the word 'cybernetics' is *kybernetes*, the Greek word for 'steersman'.

Heidegger also acknowledges in his *Spiegel* interview that, in the technological era, cybernetics has taken the role of philosophy. This can also be considered as an extension of the modern disposition that necessitates a quantified grounding for every claim. Thus, having lost its capacity to order and inform the sciences, "[p]hilosophy dissolves into psychology, logic and political science" (Heidegger, 1976, p. 212).

Michael Zimmerman considers the cybernetic siege of philosophy in relation to the following quote from *Was heißt Denken?*:

> In the West, thought about thinking has flourished as 'logic'. Logic has gathered special knowledge concerning a special kind of thinking. This knowledge concerning logic has been made scientifically fruitful only quite recently, in a special science that calls itself 'logistics'. It is the most specialized of all specialized sciences. In many places, above all, the Anglo-Saxon countries,

> logistics is today considered the only possible *Gestalt* of strict philosophy, because its results and procedures *yield an assured profit for the construction of the technological universe*. In America and elsewhere, logistics as the only proper philosophy of the future is thus beginning today to seize power over the spirit. (as cited in Zimmerman, 1990, p. 200, emphasis Zimmerman)

Logic deals with valid inferences. However, such inferences are only possible in light of specific understandings of validity, soundness, consistency and completeness. These understandings make possible the formal systems we know today. Yet, the formal system's contingency on the constructed definitions of valid, sound, consistent and complete is rarely if ever recognized.

Boolean algebra, as an extension of logic, maps truth values on variables that are defined in a formal system. Conventionally, truth value 1 denotes 'true', whereas 0 denotes 'false'. In modern technology, "Boolean algebra can be materialized in the form of electric circuits and relay switches" (Dupuy, 2009, p. 52). Thus, with the help of electric circuits and relay switches, automata can be set up to behave according to a predetermined set of rules. The more sophisticated the logical system is, the more sophisticated the circuits can be.[48] As a result, the utilitarian and mercantile needs of technology are met. Hence, the philosophical enterprise in the form of logic becomes a source of greater efficiency and higher profits for the technological man.

As logic and logistics increased their impact in societal organization and economics, human being's image began to be conceived in these terms. This created the conditions where the conception of mental activity and Boolean circuits

[48] The circuit does not necessarily have to be electrical. Once the logical structure is in place, it can be implemented by whatever means that is deemed appropriate, e.g., hydraulic, mechanical etc. This also means that this issue is not specific to electric circuits. Instead, it generalizes to enframing's challenge to instrumentalize logic. It was the case that the cited work encountered this issue in terms of electric circuits.

converged in a manner which likened the human being to a logic machine, whose mental capacities can be extracted and ordered along with other resources.

In contrast, Heidegger considers a conception of thinking based on logic to be inadequate. In his 1946 “Letter on Humanism”, where he reflects on the essence of thinking, he declares that in order to experience the essence of thinking, “we must free ourselves from the technical interpretation of thinking” (Heidegger, 1993, p. 218). The technical interpretation of thinking conceives thinking as *theōreia* and conceptualizes knowing as ‘theoretical behavior’.[49] Heidegger posits such conception of thinking to be the result of a reactive attempt to ascribe credibility to thinking as opposed to acting or doing, and to justify philosophy’s existence before the sciences. In doing so, philosophy conceives thinking in terms of logic.

While criticizing logic’s dominance in thinking, Heidegger advises his reader not to reject logic, but to conceive thinking in a more originary manner. He says: “To think against ‘logic’ does not mean to break a lance for the illogical but simply to trace in thought the logos and its essence, which appeared in the dawn of thinking, that is, to exert ourselves for the first time in preparing for such reflection” (Heidegger, 1993, p. 251). Thus, freeing oneself from the technical interpretation of thinking is only possible when we appreciate thinking as an activity of *logos*, the ground on which logic flourishes. Of course, such a conception of thinking will only be fully appreciated in the context of Heidegger’s aforementioned interpretations of truth, unconcealment and revealing.

Heidegger seeks to instill a conception of thinking based on *logos* in his technology essay as well. While discussing Aristotle’s four causes through the example of a silver chalice and the silversmith’s role, Heidegger notes that

[49] *Theōreia* is an Ancient Greek term. The term’s literal translation is “to look upon”.

conventionally, the silversmith is regarded as the efficient cause due to the physical labor that brings about the finished chalice.[50] Heidegger rejects such a conception of efficient cause. In fact, Heidegger translates Aristotle's term *aition* not as "cause", but as "that to which something is indebted" (Heidegger, 1977, p. 7). This distinction is a testament to the notion of 'authentic production' that Heidegger seeks to instill and contrast with manufacturing for the advancement of *enframing*. For Heidegger, what makes the silversmith the efficient cause is the fact that the silversmith considers carefully (*überlegen*) and thereby gathers together the first three causes. Heidegger notes the kinship between *überlegen* and the Greek terms *logos* and *legein*, which is itself rooted in *apophainesthai*, to bring forward into appearance. The instance of carefully considering (*überlegen*) gathers the material, formal and final cause to "start it on its way, namely, into its complete arrival" (Heidegger, 1977, p. 9) and thereby prepares the ground for its unconcealment. Thus, logos is appropriated as a mode of disclosure, a way to make things manifest in their being.

Heidegger's consideration of the Rhine river in three different contexts reaffirms his emphasis on logos and its contrast with the technological mode of being. In the first conception, the hydroelectric plant built into the Rhine challenges the river forth as an energy source and makes it into a water power supplier. The river now "derives from out of the essence of the power station" (Heidegger, 1977, p. 16). The second conception presents 'The Rhine' as uttered in Hölderlin's hymn by the same name. Hölderlin's poetic disclosure, just like the silversmith's careful consideration, admits the Rhine into being and preserves its being. Heidegger introduces the third conception, the objectification of Rhine as a tourist destination:

[50] Aristotle proposed to give a fourfold account of how an entity comes to be. An entity has a material cause (the material out of which it is made), a formal cause (the shape into which the material enters), a final cause (the end for which it exists) and an efficient cause (that brings forth the entity itself into presence) (Heidegger, 1977).

"But it will be replied, the Rhine is still a river in the landscape, is it not? Perhaps. But how? In no other way than as an object on call for inspection by a tour group ordered there by the vacation industry" (Heidegger, 1977, p. 16). Thus, the Rhine – even when it is "dammed up into the power plant" (Heidegger, 1977, p. 16) – is encountered as a self-standing entity. However, this encounter is only due to being involved in a culture that has inherited poetic disclosure of the river and an opportunistic tourism industry that uses the Rhine as an orderable resource for its further advancement. At the end of his essay, Heidegger quotes Hölderlin once again "poetically dwells man upon this earth" (Heidegger, 1977, p. 34) and suggests that poetic disclosure, along with authentic production, could possibly show a way out of the danger of *enframing*.

4.2.6 Cognitive science and *enframing*

Turing and the cyberneticians adopted computational approaches to the mind. This enabled them to reduce the human to an automaton that simply transforms inputs to outputs. Arguably, this could be considered as part of the threat of man being taken as standing reserve under the reign of enframing. As discussed in the previous chapter, cognitive science has been dominated by the eliminative materialist position since its inception. The eliminative materialist line of thinking in cognitive science reduces mental experience to a set of calculable physical interactions among neurons while viewing metaphysical notions as indicators of an immature stage in the scientific process. Heidegger argues that such a conceptualization of science is a vehicle for technology to fulfill its essence.

Cognitive science's eliminative materialist position promotes experimenting with animals whose brains are evolutionarily conserved and anatomically similar to the human brain, since their response would be representative of a human being's response to a particular situation. This certainly contradicts Heidegger's position that appropriated the experience of being peculiar to human beings as *Dasein*. Indeed, in a slightly later work in *The Fundamental Concepts of Metaphysics*, Heidegger states that humans are unique in having a world (*Welt*) and contrasts this with having an environment (*Umwelt*) which marks the essence of animality (Heidegger, 1995). It is this notion of having a world that enables us to respond to what entities are, as opposed to just responding to entities. Matthew Ratcliffe (2012) grounds this notion of having a world in our human capacity to "understand possibilities as possibilities" (p. 148). We understand ourselves in terms of possibilities, remain sensitive to what is possible and anticipate an open-ended range of occurrences while pursuing possibilities. In contrast, having an environment (*Umwelt*) means that "throughout the course of its life the animal is confined to its environmental world, immured as it were within a fixed sphere that is incapable of further expansion or contraction" (Heidegger, 1995, p. 193). As an example, Heidegger considers the bee: "The worker bee is familiar with the blossoms it frequents, along with their colour and scent, but it does not know the stamens of these blossoms as stamens, it knows nothing about the roots of the plant and it cannot know anything about the number of stamens or leaves, for example" (Heidegger, 1995, p. 193). Heidegger posits that humans respond to entities as entities, whereas animals do not. The distinction between world and environment claims that the mental experience of humans and that of animals are categorically different. This distinction calls for an alternative approach to the science of mind and questions the capacity of experiments with other mammalian brains to inform us about human brains.

Such an alternative approach would above all diminish the importance of rodent models in cognitive science. Almost half of all contemporary neuroscience-related research uses a rodent model as representative of the human brain and behavior due to their anatomical similarity (Ellenbroek and Youn, 2016). Some of these studies seek to explore prospective treatments for complicated disorders such as anorexia nervosa, depression and alcoholism by attempting to replicate the 'relevant symptoms' in a rodent in a chamber. Such conditions impact one's mental experience globally as an extension of their understanding of being and their intersubjective relations. Yet, this approach reduces such mental experience to a mere neural correlate and an enumerable number of symptoms while considering symptom suppression to be identical to a viable treatment.

The biological reductionism of contemporary cognitive science paves the way for human beings to be perceived as standing reserve by presenting them as resources that can be exposed and exploited. As expressed by the information processing premise of cognitive science, the human being is presented as an information processing resource and is ordered to respond to systematized pieces of information. Within such systems, entities are disclosed as elements of the system. As a result, the human being responds to the entity in ways designed by the system. Consequently, the human being, as represented by enframing –i.e., as an information processing resource- can only have an environment (Umwelt) rather than a world (Welt). Here, the environment is what a living being is exposed to, while the world is what *Dasein* discloses.

Since technology employs science to fulfill its essence, enframing challenges cognitive science to represent the human being as an information processing resource. This conception eradicates Dasein as the fundamental character of human being, since having an environment (Umwelt) suggests that being is no longer an issue for the human being.[51] As stated, this is achieved by reducing mental experience to calculable interactions among a finite enumerable set of constituents. This enumerable set ascribes a fixed ontology to the human mind, comparable to a software engineer modelling an expert system. The act of fixing suggests that one has already gained control of the object; its capabilities under different conditions have already been documented under scientific investigation. Hence, an encountered thing is no longer an object that stands over against the subject in the Cartesian sense. It is a resource whose capacities and shortcomings have been documented. As discussed, this has been made possible by an overwhelming interest in scientific method and its adoption of mathematical representationalism.

Cognitive science marks humanity's attempt to fix the human being into objectness. Grounded upon the notion that thinking is computation, it seeks to document human beings as information processing resources by analyzing their capacities and shortcomings. Cognitive science sees them as a coherence of forces calculable in advance. It dismisses any human aspect that resists quantification and objectification. This is possible because cognitive science understands the human being as a logic machine, converting inputs to outputs. Hence, the discipline sets out to formalize and calculate the mechanism that governs the process by which this conversion happens.

In this thesis, I claim that we are not obliged to the reductionistic conceptualization of cognitive science resulting from the productionist metaphysics paradigm. The first step for developing an alternative approach to mental experience in line with Heidegger's understanding of poiēsis is to follow through on his

[51] Here, Heidegger's initial definition of *Dasein* is referenced as "the being which is concerned in its being about its being" (Heidegger, 1996, p. 40)

translation of aition as 'that to which something is indebted', rather than 'caused by'. This can be achieved by abandoning the representation of the human as a worldless automaton whose mental experience is caused by neuronal interactions, and instead, consider the human as a being whose mental experience is indebted to its biological composition. This renders the next step possible, namely conceptualizing the human being as Dasein and suggesting that the understanding of being and worldliness (i.e., having a world) have an impact on mental experience. The fact that worldliness resists quantification or reification should not entail that cognitive science denies its existence. On the contrary, the scientific intractability of worldliness can ground cognitive science and delimit it to a healthy domain. This would then open up cognitive science's theoretical horizon to accept metaphysical explanations to account for what it cannot adequately address rather than considering metaphysics as an immature stage.

In its current form the scope of cognitive science is confined to mental activities that are stabilized as repeatable patterns, and thus can be addressed by formal logic. This leaves out the real work of cognition, namely, the coming into being of these regularities as a result of Dasein's engagement with its world. To put it in Heideggerian terminology; the cognitive science's present conception corresponds to participation in the stabilized regularities rather than their disclosure. This crucial shortcoming of cognitive science can be overcome by admitting worldliness into the discipline's theoretical framework, and thereby expanding its understanding of cognition.

CHAPTER 5

DISCUSSION AND CONCLUSION

The digital revolution has come a long way since the Macy conferences. Electronic devices with high computational powers have transformed our way of living and have come to define the contemporary age. The digital revolution has also created an economic revolution. Technology companies such as Apple, Amazon, Facebook, Microsoft, and Alphabet (Google's parent company) are now the most valuable companies in the world. There have been two instances, noteworthy from a Heideggerian perspective, of how some of these companies have responded to the digital revolution's impact on humans.

The first instance happened when Apple's CEO Tim Cook delivered MIT's 2017 commencement address. Throughout the speech Cook repeatedly advises graduates to search for a purpose and a mission for serving humanity. In his speech, he declares that "I'm not worried about artificial intelligence giving computers the ability to think like humans, I'm more concerned about people thinking like computers: without values or compassion, without concern for consequences" (MIT, 2017, 9:32). From a Heideggerian perspective, his remark could be interpreted as a warning against the 'supreme danger' where humans do not grasp enframing as a claim and encounter the world only as a collection of orderable resources that ought to be extracted and stored for later use, thus further ordering the technological-industrial complex without seeking any ethical justification. The Heideggerian view would regard Cook as advising students to belong to the realm of destining instead of being challenged by enframing to advance technology without any regard for implications.

The second instance happened when Facebook's former vice president of engineering Regina Dugan was the keynote speaker of Facebook's developer conference in 2017. In her speech Dugan shares her research team's aspiration to develop a "silent speech system" (Inverse, 2017, 12:47), an interface that allows one to communicate with electronic devices through one's own brain activity. To make her case, Dugan claims that the human brain can produce 1 terabyte of information per second (an issue we will examine shortly), while speech can convey merely 40-60 bytes per second. Therefore, she claims that voice interfaces serve as a bottleneck for human-machine communication. Indeed, her main premise is that speech is a "compression algorithm", yet a lossy one (another graver issue to be examined in the continuation of this discussion) (Inverse, 2017, 5:24). To eliminate the bottleneck created by such inefficient compression, Dugan promises to develop an interface that decodes speech directly from brains to type hundred words per minute in an electronic device, compared to the twenty words per minute mark achieved by the average person typing on a smartphone. She celebrates the fact that there has already been success with implanted electrodes on decoding speech from neural activity.[52] However, she warns that her product must excel with only non-invasive sensors, since, she warns, otherwise the product would not be economically feasible to scale up in production.

Regina Dugan's way of calculating the 1 terabyte of information the human brain produces in a second is noteworthy: she notes that the average human brain has around 86 billion neurons. She assumes that at any given moment one percent of

[52] It should be noted that these successful experiments work only with a predetermined set of words and phrases uttered in predetermined contexts that are controlled laboratory environments. The electrode signal analysis seems to be helpless when such a narrow set does not exist or when the speaker chooses to make up a new word, as humans occasionally do.

those neurons activate at 1 kHz.[53] Dugan's perspective is the same as the one laid out in Chapter 2; she thinks that neurons – as the sole ground of mental activity – are discrete entities that can be adequately characterized by two-valued logic. Although Dugan does not share details about her calculation, she appears to claim that every action potential in the human brain generates 1 bit of information. Of course, measuring information in terms of bits implies that a purely syntactical account of mental activity will produce a correct account of the human mind.[54] Furthermore, Dugan believes this project will be able to decode semantic information from neural activity in the long run. Semantics is expected to be extracted from the syntax of neural signals. Thereby, Dugan declares that people will be able to share their thoughts independent of the language – English, Spanish or Mandarin all languages will become the same.

From a Heideggerian perspective, Dugan's research project is an attempt to further order Facebook's supply of human users as standing reserve (Bestand) to extract more data out of them. The company's business model is primarily 'targeted advertising', achieved by leveraging the accumulated personal data of the users for advertisers to accurately target their desired audience. Hence, Facebook's commercial success is completely dependent on users interacting with their electronic devices and generating data. If successful, Dugan's project would surely increase exponentially the data Facebook users generate.

---

[53] If we consider every action potential to be corresponding to one bit, this calculation actually adds up to 0,1 terabyte per second. Dugan does not provide any further detail about her calculation. Any further attempt to make the calculation work would require making greater assumptions on Dugan's behalf.

[54] It should be noted again that Shannon's information theory is explicitly formulated for a communication environment, where a sender and receiver who share a listing of all possible events transfer messages announcing the occurrence of the actualised event. These restrictions and assumptions do not resemble the practice of human being's contemplation and speech.

Despite Dugan's departure from the company, the research project continued. Facebook published an update about the research project in March 2020 (Tech@facebook, 2020). They hope that their technology can produce a device that recognizes a handful of commands such as 'home', 'select' and 'delete' within a decade. This is fundamentally different from the grandiose mission of developing a silent speech system. Developing an algorithm to implement a command chosen from an enumerable set of commands is at least as old as Turing computation. The single novelty in this case is a device that distinguishes a handful of neural patterns to register the choice on the user's behalf.

As it has been laid out in Chapter 2 through discussions of Turing computation, Shannon information and cybernetics, technological developments followed a single theme: identify a phenomenon's constituents as discrete entities, compile them as an enumerable list, and formalize the relations between them. This computational framework conceived of information as a purely quantitative concept measured in bits without any reference to semantics. Chapter 3 revealed that cognitive science became institutionalized as the application of this framework to the human mind. This application conceived humans as physical symbol systems and characterized mental activity by representing neurons with two-valued logic. By attempting to characterize all mental activity as fundamentally computational, this framework set up the researcher to explain away higher order human capacities that are sensitive to meaning and relevance. Philosophically, emergence of the eliminative materialist position consolidated this approach.

The limits of the computational framework have also been laid out in Chapter 2: Turing computation is bound by countable infinity, well-defined relations, and algorithms. In contrast, from a Heideggerian perspective based on *Being and Time*, one would characterize our mode of relating to the world as something that cannot be counted (*gezählt*) but only recounted (*erzählt*). Our embodied presence in the world engages in open-ended relationships while being guided by a sense of relevance and meaning. As a result, we foster a narrative that characterizes *Dasein*'s mode of being. Our theoretical conception and formalisms about the world arise as a derivative of this narrative. While Heidegger posits that the character of this narrative resists formalization, cognitive science aspires to formalize it by reverse engineering.

Heidegger's analysis in his technology essay states that modern technology has imposed mathematical representationalism on science to fulfill its essence. Cognitive science could be regarded as a specialized domain within this paradigm. Considering the aforementioned five companies that have leveraged digital technologies to create a world of digital hyperconnectivity where they seek to fulfill all of humans' needs while positioning humans as data generators whose natural disposition is to process information, cognitive science can easily be credited for the rise of digital technologies. Man's image has been reconstructed as an information processor in the information age. Initially, the term 'information' was conceived within Shannon's prescribed boundaries. However, in attempting to account for capacities such as attention, memory, contemplative thinking and language, cognitive science has been using the term information outside of Shannon's intended scope. As Sybille Krämer states, the term has become "not so much a scientific concept as it is a mythical one" (Krämer, 2012, p. 17). In its current form, cognitive science lacks a rigorous definition of information and yet the term pervades every instance of cognitive science.

In topics such as vision and expert systems, the information-processor approach has achieved a high degree of commercial success. The success of the computational framework in these topics owes to the fact that the functionalities needed in these domains lend themselves to a treatment on a purely syntactical basis with a quantitative conception of information. However, this success comes at the expense of a recognition of human being's higher-order capacities, since they cannot be accounted for with the information-processor approach. In this context, thinking is considered to be a form of computation and conceived in logical terms.

While characterizing the technological mode of being and its danger, Heidegger seeks to instill a conception of thinking based on logos and to uphold poetic disclosure as a way of revealing. While stating that thinking can be conceived in terms of logic, Heidegger states that such a conception remains alien to the essence of thinking. He likens evaluating the human being's cognitive capacities by a standard benchmarked against logic to "the procedure of trying to evaluate the essence and powers of a fish by seeing how long it can live on dry land" (Heidegger, 1993, p. 219). Cognitive science – insofar as it equates thinking with computation – strands thinking on dry land.

A rather brief examination of Heidegger's work in the context of cognitive science and technological development suggests that while the human mind is not a logic machine, it can perform algorithmically. Yet the technological mode of being challenges human beings to conceive themselves in purely information-processing terms, as an orderable reserve in a world of orderable reserves. While this approach may be fruitful for some of human beings' capacities, it closes them up to their essence. This computational framework prepares the ground for cognitive science to consider mental processes as mere symbol manipulations devoid of subjectivity and reduce logos to logic.

Given its main premise that ‘mind is an information processor’, the immediate task of cognitive science should lie in developing a rigorous definition of the term information. In doing so, it can then propose a region of validity for its information-processor premise. The region of validity will also delimit cognitive science to what it can methodologically address, while enabling the discipline to admit concepts it cannot objectify or quantify into existence. In such a framework, the discipline can harbor metaphysical conceptualizations, instead of dismissing them on the basis of being immature. Outside the said region of validity, as far as those aspects of subjectivity are concerned that are sensitive to meaning and relevance, cognitive science has to seek an alternative conception of information that is less exact, yet more sensitive to meaning, and capable of representing the multidimensionality of human thought.

REFERENCES

Alpaydin, E. (2016). *Machine learning: The new AI*. Cambridge, MA: The MIT Press.

Bennett, M. V. (1997). Gap junctions as electrical synapses. *Journal of neurocytology*, *26*(6), 349-366.

Boden, M. A. (2008). *Mind as machine: A history of cognitive science*. New York: Oxford University Press.

Cartmill, M. (1992). New views on primate origins. *Evolutionary anthropology: Issues, news, and reviews*, *1*(3), 105-111.

Churchland, P. S. (2002). *Brain-wise: Studies in neurophilosophy*. Cambridge: The MIT Press.

Dreyfus, H. (1965). *Alchemy and AI*. Santa-Monica: RAND Corporation.

Dreyfus, H. (1972). *What computers can't do: The limits of artificial intelligence*. New York: Harper & Row

Dreyfus, H. L., & Hubert, L. (1991). *Being-in-the-world: A commentary on Heidegger's Being and Time, Division I*. Cambridge: The MIT Press.

Dreyfus, H. L. (2008). *On the internet*. London: Routledge.

Dreyfus, H. L. (2012). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. In J. Kiverstein & M. Wheeler (eds.). *Heidegger and cognitive science* (pp. 62-104). Hampshire: Palgrave Macmillan.

Dupuy, J. P. (2009). *On the origins of cognitive science*. Cambridge: The MIT Press.

Ellenbroek, B., & Youn, J. (2016). Rodent models in neuroscience research: Is it a rat race? *Disease models & mechanisms*, *9*(10), 1079-1087.

Fodor, Jerry A. "Modules, frames, fridgeons, sleeping dogs, and the music of the spheres." (1987). In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 25–36). Cambridge, MA: The MIT Press.

Friedman, M. (1994). *Handbook of perception and cognition*. E. C. Carterette (Ed.). New York, NY: Academic Press.

Friedenberg, J., & Silverman, G. (2011). *Cognitive science: An introduction to the study of mind*. Thousand Oaks, CA: Sage Publications.

Han, Byung-Chul. (2017). Psychopolitics: Neoliberalism and New Technologies of Power. London: Verso.

Haugeland, J. (1985). Artificial intelligence: The very idea. Cambridge, MA: The MIT press

Heidegger, M. (1976, May). Nur noch ein Gott kann uns retten. *Der Spiegel* 30. pp. 193-219. trans. by W. Richardson as "Only a God Can Save Us" In *Heidegger: The Man and the Thinker* (2017), ed. T. Sheehan, pp. 45-67. New York, NY: Routledge.

Heidegger, M. (1977). *The question concerning technology and other essays* (W. Lovitt, Trans). New York, NY: Harper & Row.

Heidegger, M. (1995). *The fundamental concepts of metaphysics*. Bloomington: Indiana University Press.

Heidegger, M. (1996). *Being and time: A translation of Sein und Zeit* (J. Stambaugh, Trans.). Albany, NY: SUNY press.

Inverse. (2017, April 20). *Regina Dugan's keynote at Facebook F8 2017 | Inverse* [Video]. YouTube. https://www.youtube.com/watch?v=kCDWKdmwhUI

Kolak, D., Hirstein, W., Mandik, P., & Waskan, J. (2006). *Cognitive science: An introduction to mind and brain*. New York, NY: Routledge.

Krämer, S. (2012). Information Measurement and information technology: A myth of the twentieth century. In Mitcham, C., & Huning, A. (eds.). *Philosophy and technology II: Information technology and computers in theory and practice* (pp. 17-28). Dordrecht, Holland: Reidel Publishing Company.

Li, M., & Vitányi, P. (2008). *An introduction to Kolmogorov complexity and its applications* (Vol. 3). New York, NY: Springer.

Massachusetts Institute of Technology (MIT). (2017, June 10). *Tim Cook's MIT Commencement Address 2017* [Video]. YouTube. https://www.youtube.com/watch?v=ckjkz8zuMMs

McCarthy, John, & Hayes, P. Some philosophical problems from the standpoint of artificial intelligence. *Readings in artificial intelligence*, 1969, 431-450.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine*, *27*(4), 12-12.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115-133.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81-97.

Miller, G. A. (1983). Informavores. In F. Machlup & U. Mansfield (eds.), *The Study of Information: Interdisciplinary Messages* (pp. 111-113). New York, NY: Wiley Interscience.

Nagy, J. I., Pereda, A. E., & Rash, J. E. (2018). Electrical synapses in mammalian CNS: Past eras, present focus and future directions. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1860*(1), 102-123.

Newell, A., & Simon, H. A. (1976). Computer Science as Empirical Inquiry: Symbols. *Communications of the ACM*, *19*(3), 11-126.

Núñez, R., Allen, M., Gao, R., Rigoli, C. M., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behaviour, 3*(8), 782-791.

Piccinini, G., & Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics*, *37*(1), 1-38.

Ratcliffe, M. (2012). There can be no cognitive science of Dasein. In J. Kiverstein & M. Wheeler (eds.). *Heidegger and cognitive science*. pp. 135-156. Hampshire: Palgrave Macmillan.

Roochnik, D. (2012). *Retrieving Aristotle in an age of crisis*. Albany, NY: SUNY Press.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, *10*(1), 18-24.

Russell, S., & Norvig, P. (2002). *Artificial intelligence: A modern approach*. Upper Saddle River, NJ: Pearson

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379-423.

Shannon, C. E. (1956). The bandwagon. *IRE Transactions on Information Theory*, *2*(1), 3.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*, Urbana, IL: University of Illinois Press.

Sheehan, T. (2014). The turn. In B. W. Davis (ed). *Martin Heidegger: key concepts*. pp. 82-101. Durham: Acumen.

Sipser, M. (2012). *Introduction to the theory of computation*. Boston, MA: Cengage Learning.

Tech@facebook. (2020). Imagining a new interface: Hands-free communication without saying a word. Retrieved March 30 2020, from https://tech.fb.com/imagining-a-new-interface-hands-free-communication-without-saying-a-word/

Thagard, P. (2005). *Mind: Introduction to cognitive science*. Cambridge, MA: MIT Press.

Turing, A. (1950). Computing machinery and intelligence-AM Turing. *Mind*, *59*(236), 433-460.

Zimmerman, M. E. (1990). *Heidegger's confrontation with modernity: Technology, politics, and art*. Bloomington: Indiana University Press.