

8-2018

Persuasive Intelligence: On the Construction of Rhetor-Ethical Cognitive Machines

Nathan D. Riggs

Clemson University, riggs2@g.clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

Riggs, Nathan D., "Persuasive Intelligence: On the Construction of Rhetor-Ethical Cognitive Machines" (2018). *All Dissertations*. 2191.

https://tigerprints.clemson.edu/all_dissertations/2191

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

PERSUASIVE INTELLIGENCE:
ON THE CONSTRUCTION OF RHETOR-ETHICAL
COGNITIVE MACHINES

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Rhetorics, Communication, and Information Design

by
Nathan D. Riggs
August 2018

Accepted by:
Dr. Steven B. Katz, Committee Chair
Dr. Charles Starkey
Dr. Jan Rune Holmevik
Dr. Pamela Mack

ABSTRACT

This work concerns the rhetorical and moral agency of machines, offering paths forward in machine ethics as well as problematizing the issue through the development and use of an interdisciplinary framework informed by rhetoric, philosophy of mind, media studies and historical narrative. I argue that cognitive machines of the past as well as those today, such as rapidly improving autonomous vehicles, are unable to make moral decisions themselves foremost because a moral agent must first be a rhetorical agent, capable of persuading and of being persuaded. I show that current machines, artificially intelligent or otherwise, and especially digital computers, are primarily concerned with control, whereas persuasive behavior requires an understanding of possibility. Further, this dissertation connects rhetorical agency and moral agency (what I call a *rhetor-ethical constitution*) by way of the Heraclitean notion of *syllapsis* (“grasping”), a mode of cognition that requires an agent to practice analysis and synthesis at once, cognizing the whole and its parts simultaneously. This argument does not, however, indicate that machines are devoid of ethical or rhetorical activity or future agency. To the contrary, the larger purpose of developing this theoretical framework is to provide avenues of research, exploration and experimentation in machine ethics and persuasion that have been overlooked or ignored thus far by adhering to restricted disciplinary programs; and, given the ontological nature of the ephemeral binary that drives digital computation, I show that at least in principle, computers share the syllaptic operating principle required for rhetor-ethical decisions and action.

ACKNOWLEDGMENTS

I would have not finished this dissertation without the patience and support of my family, friends, colleagues, and mentors. Dr. Alison Balaskovits, my spouse, deserves more credit and thanks than ever could be expressed here or elsewhere for her love, her kindness, her generosity and her understanding. Likewise, many thanks go my chair, Dr. Steven B. Katz, whose guidance, acuity and assistance were and still are integral to my research. My committee members, Dr. Charles Starkey, Dr. Pamela Mack, and Dr. Jan Holmevik also deserve many thanks for their substantial help and enlightening feedback. My appreciation extends as well to the Rhetorics, Communication and Information Design Program, and especially to those in it who served as mentors and teachers. Without the coursework and direction provided by Victor Vitanza, Cynthia Haynes and Stephanie Barczewksi, along with that again of Holmevik and Katz, this dissertation would have followed a radically different and ultimately less illuminating path. I would also like to thank my cohort for their lively discussion, comradery, and understanding: Daniel Frank, Samuel Fuller, Firasat Jabeen (and family), Eda Ozyesilpinar and Joshua Wood (and family). Special thanks additionally go to Dr. Mari Ramler, whose compassion and support knows no bounds, and who ultimately forced me to engage with *Spirit* despite my many protests; to my mothers and my fathers, without whom I would have surely taken a less desirable course in life; and to the many friends and relatives who are too numerous to name here. Lastly, thanks are due to my cats, whose constant companionship and care have been indispensable, and to my first computer, the Apple //c, without which neither my research nor my curiosity would have flourished.

TABLE OF CONTENTS

	Page
TITLE PAGE	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	vii
CHAPTER	
I. AN INTRODUCTION	1
Impetus	5
Assumption	14
Mechanical Monstrosities	15
Changing Minds	19
Rhetoric and Morality	23
Calculated Mediation	28
Machine Morale	33
Mechanical Persuasion	37
Summa	39
(Pre-)Destination	41
II. (COMPUTATION & REPRESENTATION) (INTELLIGENCE & RHETORIC)	44
Analog	44
Funhouse	45
Metaporia	50
Homo Imitans	54
Kouroi	59
Talos and Medea	61
Aeolipile	64
Being	69
Sophoresis	75
Deep Function	78
Radius	81
Enframing	83
This Automatic Life	86

III.	MASCHINENMENSCH	89
	Narrative	89
	Prometheus Rebound.....	93
	Dis(re)memebered Machines.....	98
	Testament.....	100
	Uncovered, Child Cutting Grass.....	108
	Reproducing Male Desire	111
	Antipythia	114
	Demonstration.....	118
	Cruciation	121
	Contraception.....	127
	Conception.....	132
	Reckoning.....	137
	The Games They Play	141
	Atari.....	150
IV.	CODE && CONTROL.....	154
	Technique	154
	Digital Ontology	159
	Hierarchy of Bits.....	162
	Executable Ideology.....	165
	From Discipline to Control.....	171
	The Arborescent Rhizome	173
	Modeless = Multimodal.....	176
	Intramediation and Spectacular Speed	179
	Invisible Machines, Disinterested Surveillance.....	181
	Hegemony of Quantification	185
	Return of the Bit.....	188
V.	A RHETORETHIC OF HERACLITEAN MACHINES.....	190
	The Bit Strikes Back	194
	The Føree Flame	199
	Doxa Vadere	202
	Syllap Stick / Straw Men.....	204
	Sylph and	210
	Other (You're It).....	211
	Rhetor-Ethical Encounters.....	214
	With Heraclitean Machines	216

VI. A CONCLUSION.....219

 Ascended Glitch.....221

 Reasoning Unreason.....224

 Act of God225

 Time Forward.....227

WORKS CITED.....229

LIST OF FIGURES

Figure 1.1 Nathan Riggs. Films about Artificial Intelligence / Robotics	4
Figure 3.1 Flickr, DrSJS. Girl Woman Face Posing Side Robot. CC0 Creative Commons License	111
Figure 5.1 Flickr, Trevor Hurlbut. Pareidolia. CC BY 2.0 License	206
Figure 5.2 WikiMedia, DoctorWho42. Deep Dream Toast Sandwich. CC BY-SA 3.0 License	207

CHAPTER I

AN INTRODUCTION

He who does not live with persuasion cannot fail to obey, for he has already obeyed.

– Carlo Michelstaedter, *Persuasion and Rhetoric* (35)

This dissertation is ultimately concerned with the rhetorical and moral agency of machines. What constitutes and defines the key terms in such a statement, however, and the theoretical assumptions from which these understandings are derived, complicates any swift judgments concerning the possibility or the existence thereof (the matter made even more difficult when they are combined into *machine rhetorical agency*). Do machines have rhetorical agency? Can they? Should they? And what do we mean by *rhetorical* or *agency* or *machine* in the first place? These questions cannot be considered in isolation nor can they be separated, we will argue, from questions concerning moral agency—or, for that matter, moral *patience* (which then begs the question: what of *rhetorical* patience?).¹

¹ In general, a moral agent is an entity capable of doing right and wrong, while a moral patient is an entity that deserves moral consideration (can be righted or wronged). Floridi & Sanders complicate this matter by

In this we share a common concern with a variety of scholars, writers, researchers, engineers, scientists, politicians, business leaders and consumer advocates. Elon Musk and Jeff Bezos, well-known today as almost corporate monarchs, have expressed more than passing concern for the increasingly “smart” behavior of machines, Musk himself claiming that Artificial Intelligence (AI) poses the single largest existential threat to humankind, including nuclear disaster (Domonoske). Stephen Hawking, the late fabled physicist whose very capacity to communicate is in part possible because of work on intelligent machinery,² asserts that “the development of full artificial intelligence could spell the end of the human race” (Cellan-Jones). Politicians, like Barack Obama, are increasingly concerned with the development of artificial intelligence and its growing ability to automate more and more jobs, leading to massive unemployment (and likely unrest)³; and AI engineers, cognitive scientists and philosophers of mind have written countless arguments over the problem of autonomous machines without some sense of

arguing that moral agents do not require a will, which is usually considered a core aspect of agency (but not of patience: ecosystems can be moral patients but are likely not to be agents nor have will).

² Hawking’s speech synthesis software has changed over time, but recent incarnations have relied on heavy use of predictive text, a feature that many of us are so intimately familiar with today that we no longer give it a second thought. As of 2017, Hawking uses a service called *SwiftKey*—which just so happens to be the same software that many smart phones use. Software such as this makes heavy use of computational linguistics, a sub-discipline that branched out from early efforts in artificial intelligence. For more detail, see Medeiros, Joad.

³ In areas where population is declining, automation is a boon, as there are not enough workers to continue production levels; China’s efforts at automating manual labor (both the state and the private sector) is particularly notable: Some factories have replaced over 90% of their laborers, and Foxconn, manufacturer of the iPhone, plans to replace 30% of its workforce by 2020, and as of 2016 had already replaced 60,000 workers with robots (Statt). For much of the world, however, automation threatens to increase unemployment and under-employment levels to unsustainable highs, as well as suppressing wage growth. This has prompted calls for a “Universal Basic Income” (UBI) in many countries, in which each citizen receives a monthly stipend. Surprisingly, UBI has gained considerable traction in the United States—but unsurprisingly, it is also coupled with the dismantling of the last vestiges of the welfare state.

morality, birthing in the past two decades an entirely new hybrid discipline: Machine Ethics.

But fiction writers, as well as mythologists and theologians (and rhetoricians!), have been considering the implications of machine agency for quite some time. This will be outlined in greater detail in chapter three, but it is worth noting here that even in our most ancient systems of belief and oldest texts (spoken, written, or drawn), the prospect of creating autonomous machines, whether made of mud as a Golem from Jewish folklore or made of gold like Talos in Greek Mythology,⁴ or of gears or silicon or fabricated flesh and bone, there is some sort of urgent anxiety that spans across many different cultures concerning automata that bears consideration. This is particularly important when much of actual scientific practice is informed by, and inspired by, the epic science-fiction tales of intelligent machines enslaving, destroying, or simply forgetting about we comparatively simple, and thoroughly moral, human beings. Hannah Arendt quipped perhaps unironically in her *The Human Condition* that science-fiction has long been poorly understood as a driving force of mass desires that are realized through scientific practice (2-3), and we concur: science-fiction is the oracle of a new age, solidified and mythologized as its prophecies come to pass, and scientific practice is its temple of Apollo.⁵

⁴ In Greek mythology, Talos was a bronze automaton created by the god of forgery (among other things), Hephaestus. It was tasked with patrolling the shores of Crete to prevent pirates from landing.

⁵ The Oracle of Delphi lived in the temple of Apollo. Likewise, science-fiction lives in its contemporaneous scientific practice.

A survey of the many mythologies, folktales and fictions is forthcoming in our third chapter, but it is likely not difficult for most readers to name more than a handful of works that explore this

theme; in the past four decades alone (1977-2017), over 65 films have been released that contemplate the becoming of autonomous,

artificially intelligent

machines—usually including, in some fashion, the impending doom of humanity (Fig.

1.1).⁶ Compare this to the number of films released prior to 1980 that address the same theme—twelve!—and it seems our concern, or at least our collective fascination, has

grown more urgent. From the machine’s film debut in *Metropolis* to its alien intelligence in 1951’s *The Day the Earth Stood still*, to *Westworld* in 1973 and to the *Westworld*

series today (2017), in between them a host of iconic imaginations—*Star Wars*, *Star*

Trek, *Alien*, *Blade Runner*, *Terminator*, *Ghost in the Shell*, *A.I. Artificial Intelligence*, *The*

Matrix, and so on—it seems that machines are getting plenty of work in Hollywood, and

in our imagination. The same development, one suspects, can be seen expressed in other media.

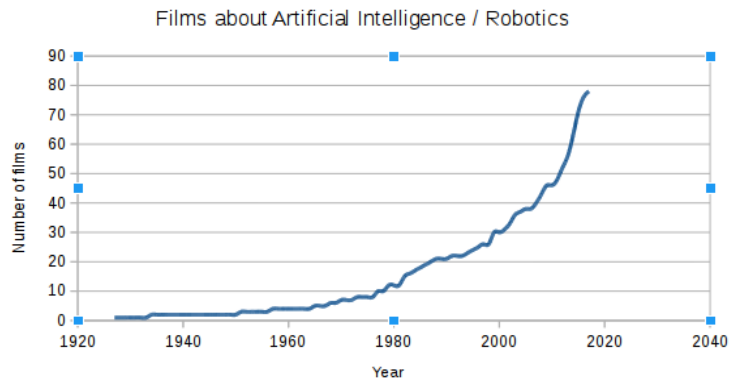


Fig. 1.1

⁶ This graph was compiled from lists of movies by topic on Wikipedia. Strangely, few efforts have been made elsewhere to catalog these films. See https://en.wikipedia.org/wiki/List_of_artificial_intelligence_films for the entire list of films and their release dates.

Impetus

“Intelligent” and autonomous machines, at least of the fictitious variety, seem to be advancing on all imaginary fronts—but so too are they on less imaginary ones, and while this advance has not gone unnoticed, the most pressing series of threats that such a development poses has largely flown under the radar of popular accounts regarding robotics, artificial intelligence, automation, and related topics. Here, at least, we are not so much concerned about or with a potential future in which well-armed mechanical beasts and men roam the earth to hunt the last vestiges of humanity. While such scientific fictions are numerous and may very well predict a distant future that has yet to happen, we are concerned with machines today and in the near future, and in particular the sorry state of their autonomy—sorry, that is, for humans. For we are, as unbelievable as it may seem, living in a world run through and with and by machines; and although we made them, we do not control them. As roboticists are wont to do, we too find it necessary to make an analogy to children:⁷ every parent knows that whatever control they have over a child is contingent at best and thoroughly invariable over time. The same applies to the autonomy of machines today: our control is contingent, and it isn’t that machines might simply disobey us (though one day this could be a possibility), but that they have no need or urge or desire or reason to hear us in the first place; they are utterly, and monstrously, indifferent.⁸

⁷ Kathleen Richardson provides an excellent survey of roboticists treating their creations as children in the introduction to her *Annihilation Anxiety: An Anthropology of Robots and AI*. This only makes that fact that the military is most interested in the use of robots more off-putting (13).

⁸ Nick Bostrom, a philosopher at the University of Oxford known for his work on existential threats, is one

Like the parent whose child refuses to listen, we are held captive by the machine's disregard for our well-being, let alone regard for its own. And we cannot in good faith turn most machines off any more than we may turn off our children: anyone who dares an attempt will either be punished with the full wrath of other humans or will find themselves, along with everyone else, suddenly without the matrix of technologies that sustains the human-machine apparatus. States would fall, emergency systems fail, stock markets crash, nuclear waste burn bright and frightening—and yes, so too would our images no longer be tagged and attributed to us and as us on social media sites and search engines. Ted Kaczynski, known to most of us as the Unabomber, would revel in the haste destruction of the technical complex—but few of us would survive long enough to witness his jubilation.⁹

Yes, this is hyperbole—but only slightly so.

The truth is that machines have run autonomously for quite some time, but more importantly they have been autonomously indifferent—and going back to a “simpler time” is not an option. We have, of course, autonomous technology in the sense that Langdon Winner means it—a technological apparatus that has become so complicated and ubiquitous that it has its own politics, priests, and servants—but the situation is far

of the few philosophers who take seriously the indifference of intelligent machines. In his essay “Ethical Issues in Advanced Artificial Intelligence,” and later expanded in his *Superintelligence*, Bostrom uses the example of an AI whose sole purpose is to produce paperclips and to protect itself from any efforts to prevent it from doing so—eventually drowning us all in tiny aluminum binders.

⁹ For the sake of faulty memories: the Unabomber sent explosive devices through the United States Postal Service to various leaders and professors of technology, intending to start a revolution against modernity (he was an anarcho-primitivist with many right-wing leanings). *The Washington Post* and the *New York Times* agreed to publish his manifesto in exchange for ending the bombing campaign. Humorously, it can be read online: <http://www.washingtonpost.com/wp-srv/national/longterm/unabomber/manifesto.text.htm>

direr.¹⁰ The nuclear defense systems of both the former Soviet Union and the United States have for decades silently waited for radio signals to die that would communicate to them the need to launch their own missiles in retaliation, automatically.¹¹ The stock market "flash crash" of 2010 was partially the result of algorithms trading at lightning speed with no input from humans other than our urge to buy and sell, and today a machine can determine whether or not a woman is pregnant and target the appropriate advertisements to her well before she is aware of it herself.¹² The same logic applies to any personal endeavor or issue that can be gleaned from massive amounts of data that includes your friendships, acquaintances, purchasing habits, credit history, website visits, search history and so on—virtually any activity that can be observed, within the boundaries of our constantly eroding notions of privacy, will be and has been observed,

¹⁰ Winner, an admirer of the work of Jacques Ellul (another anarcho-primitivist), largely attributed the "autonomy" of technology, which he tied, like Ellul, to the concept of *technique*, to the difficulty of understanding how it works; our disciplines have become so specialized that even those in related fields have trouble knowing how certain technologies work that they must use. This is most certainly the case with artificial intelligence today, as even the people who design neural networks have little understanding of how these systems work *in practice*. This makes Winner's proposed solution to the social problem of autonomous technology entirely unfit today: he proposed that we learn to take the technologies apart to understand them and repurpose them to human needs, but we are now at a point where this seems impossible even for the specialists who create the technologies.

¹¹ HAM radio enthusiasts have known about mysterious radio signals that transmit only continuous beeps or robotic and repeating countdowns for quite some time (see Tousignant), and these are likely transmission systems for military technology: if the transmission is interrupted, that communicates to either soldiers or a computer that an attack has occurred, and to put defense efforts into motion. The Soviet Union's effort to create a cybernetic system dubbed *Dead Hand* that would automatically retaliate in the case of a nuclear attack was no secret. The United States had a similar system called the Emergency Rocket Communications System (ERCS). Without taking too many liberties, we can probably assume that similar systems exist today. Thomas Rid expertly describes these systems, and more, in his *Rise of the Machines: A Cybernetic History* (see also Thompson for a focused account of *Dead Hand*).

¹² In one astonishing case, a teenager was sent advertisements from Target for baby items, enraging her father. After investigating his complaint, it was not only found that she had a high "pregnancy prediction" score, a value created automatically through algorithms based on what a consumer buys, and thus received the advertisements because of this, but also that she was, to the dismay of her family, pregnant (Hill).

and will be and has been analyzed by beautifully indifferent “machines of loving grace.”¹³

Gilles Deleuze, shortly before his death in the 1990s, saw this emerging technical apparatus as a control society, where discipline becomes unnecessary when every facet of one’s life is automatic and measured, bounded and observed. Ellul saw this too, much earlier and in greater detail, in his *The Technological Society*. In place of Ellul’s artificial wombs and technician priests¹⁴ and in place of Deleuze’s nefarious rhizomatic control is something much more mundane but far more threatening, however: the indifference of that system of technical objects that controls us, thinks us (or at least *cognizes us*), compartmentalizes us. Far worse than an angry god, an envious god, a vengeful god, and worse still than the god we killed so long ago (the act of which Nietzsche is so fond of reminding us): an indifferent god, all-knowing and omnipresent, but not a care in, nor a care for, the world and its many inhabitants, whose worship is meaningless—but worship we do nonetheless.

Already, the entirety of our relationship to the world beyond ourselves, and in part those existing in our own homes, has been categorized and analyzed by ethereal artificial neural networks too complex to be understood by any human being, even altering themselves without human intervention to better improve the efficiency and scope of

¹³ This is a reference to the last two lines of Richard Brautigan’s poem “All Watched Over By Machines of Loving Grace.” The poem was about a techno-utopia and the symbiosis between animal and machine.

¹⁴ Ellul: “Death, procreation, birth, habitat: all must submit to the technical efficiency and systematization, the end point of the industrial assembly line. What seems to be most personal in the life of man is now the technicized.” (128)

their activity.¹⁵ Judging from the technological horizon it does not seem that such networks are going to get less complicated or any less miraculous—or safer. While the technical apparatus, arguably intelligent at least on its own terms, has largely existed as a passive, somewhat glorified filing cabinet and junk-mail generator, housed on silicon in dusky server warehouses and zipped across the globe at the speed of light on a vast configuration of wires, fiber-optic cables and radio waves, the ghost of the machine¹⁶ has slowly begun to acquire a body—or, rather, bodies, and we mean this in more than one way: it has acquired metal and plastic bodies to inhabit as well as softly fleshed human bodies to interface, control, serve, compute. While there is some dispute among experts over the timeline for particular technologies to take their form, already we are seeing great strides made in autonomous vehicles, robotics, human-computer coupling (literal cyborgs rather than figurative ones¹⁷), and the very artificial intelligence that, once thought to have hit a developmental dead-end last century, now takes part in most of our lives in highly personal ways (at least in the “first-world”), often without being noticed,

¹⁵ The problem in understanding the operation of neural networks is partially a problem of scale. Each node in a network is connected to multiple nodes in another layer, and an operation underway cannot be isolated to a manageable number of them. As of September 2017, the largest neural network contains 160 billion parameters (Hsu), which largely includes the number of nodes in the network.

¹⁶ We are referring here to Arthur Koestler’s 1967 book *The Ghost in the Machine*, a philosophical treatise on the mind-body problem. This was also the inspiration for Masamune Shirow’s Japanese comic series (manga) *The Ghost in the Shell*.

¹⁷ Donna Haraway (and most others) defines the cyborg as a hybrid of machine and organism (291), and without stretching too much we can easily name several cyborgs today. Anyone with a prosthetic limb, of course, qualifies, as does anyone who has been implanted with a pacemaker, an RFID chip, and so on. There are a few enthusiasts who purposely modify themselves as well: Neil Harbisson, an artist, has modified himself in a plethora of ways, including a device that allows him to “hear” colors (he is otherwise colorblind). He also founded *The Cyborg Foundation*, an organization dedicated to helping others become cyborgs.

at times more involved in our daily life than other, and actual, human beings—all without an inkling of concern for the people under its spell.

Enter machine ethics.

The solution to this problem of autonomous and indifferent machines, intelligent at least inasmuch as a cockroach may be considered so,¹⁸ is proposed by some to provide these machines with the capacity for or faculty of morality—that is, make machines *concerned*. And despite the many pitfalls of such an approach, and the incredulity one encounters at the merest suggestion that a machine could one day obtain, or at least simulate or emulate,¹⁹ this “most human” of virtues, we are inclined to agree. Here we differ, however, on the approaches to machine morality proposed thus far, as they lack, as we will argue, the most critical component necessary for moral behavior in the first place: the capacity to persuade and to be persuaded, not through the artifacts and efforts of others, but through one’s *own constitution*. Our smart phones (if we have not “upgraded” ourselves by now, reading this in the distant future, with devices even more astonishing), for instance, are indeed persuasive, but it is not your phones that are doing the persuading; it is a medium of persuasion thus far incapable of persuasion itself. It is, in

¹⁸ A typical cockroach has roughly one million neurons that compose its brain; if we are to judge whether something deserves moral consideration by this determination, let alone moral agency, then we will quickly find it in our pockets (on our phones).

¹⁹ In the terminology of computer science, there is a clear distinction between *emulation* and *simulation*. A simulation attempts to replicate the behavior of its object, while an emulation attempts to reproduce the inner workings of it. Artificial neural networks are somewhere in-between the two, as they are inspired by biological brains, but do not operate like them to the tee—we still do not know, of course, how brains operate entirely.

short, exactly what one might define as a machine: it embodies, extends and enacts the efforts of persuasive beings, but it does no such task itself.²⁰

Machine ethicists have proposed, in general, two overarching approaches: either explicitly coding, in a computer language, the rules of an entire ethical system—which, as we will see, seems untenable for humans, let alone digital computers—or “teaching” machines through the use of neural networks and thousands of simulated situations, much like Google’s *AlphaGo* recently “learned” how to become a master at one of humankind’s most complex games, *Go*.²¹ The latter approach follows from Aristotle’s virtue ethics as learned behavior (Wallach & Allen 121), which quite obviously appeals to social constructionists and many rhetoricians; but beyond being explicitly “told” which behavior is moral or not in the process of learning, machine ethicists have not yet quite decided how a machine could possibly discern “goodness” (happiness) in the first place. To be forced to agree on what is good, whether it be at gunpoint or by rote repetition of simulations in order to make a feasible inference in other situations, is to not be a moral agent at all, but a medium for the moral agency of others.²² Like the smart phone that

²⁰ This is like the Marxist conception of machinery as embodied labor / fixed capital. It produces no value itself but is the objectified labor of the masses of workers, transformed to fixed capital from surplus and reintroduced to the workforce (see *Grundrisse* 693). How artificial intelligence might be interpreted within this framework would be an interesting project but is ultimately beyond our scope.

²¹ The rules of Go are simple, but the number of combinations on the board is nearly unfathomable; a standard 19x19 board has a total of $2.08168199382 \times 10^{170}$ possible configurations, which is why it was classically considered a game of intuition and impossible for a machine to play moderately well: even with the processing power that computers have today, it would take thousands of years to calculate all of the positions in order to decide on the best move. This is how Deep Blue beat Gary Kasparov in Chess, but it is not possible with Go.

²² As Tonkens suggests, almost all popular approaches to the implementation of machine ethics violate the rules set forth by the very ethical systems they intend to implement. This is an especially persistent problem with Kantian machine ethics: hard-coded rules violate the condition of moral autonomy that the ethical system rests upon (426). If we were to build such machines, they would find their own existence unethical!

persuades you to buy it with its sleek design, ease of use, maximum utility: there is a man behind the curtain pushing knobs, turning dials, streaming ones and zeros to fool you.

This, to many, is quite ideal—do we not trust the morality of humans behind curtains instead of the cold logic of computation? We trust human over animal, over corporation, over abstracted government agencies—how could we not, as humans?—and the same must apply to machines. But we trust corporations with our credit scores; we trust governments with our rights and interests. We trust neither with our most intimate details, nor should we, and we trust our companion animals far more by comparison: like Derrida and his cat,²³ our pets know us more than we might ever allow another human to know, and we trust them with this knowledge primarily, one might think, because we trust that they will not expose or betray us. And at least with the state of human-animal relations today, our trust is well-founded on the incapacity of pets—cats, dogs, mice, gerbils, rabbits, capybara!—to communicate these things, or have the will to say them.

The same does not hold true for our machines, which share the same intimate spaces as our pets—if not more so. We foolishly trust them with our personal lives, and we even more foolishly trust them with our actual lives: your phone may not feel the urge to drag you from a burning building, but your dog may do so without a second thought. It is true that a device might be *designed* to alert you and others that danger is afoot, but such a device gives you no consideration or worth; a human behind the curtain, or a collective thereof, does this. But what is this curtain made of, and what does it let

²³ Here we reference the contents of Derrida's *The Animal That Therefore I am*, in which Derrida concludes that he is not only looking at his cat—his cat also looks at him.

through? We already know that others may watch us through the camera lenses of our laptops and phones, unannounced and unwanted and without scruples, and this sends enough chills down our spines to merit covering the camera lenses on our electronic devices with tape, rendering our devices into part-time Kodaks. How comfortable are we to be done with the curtain, and simply let everyone through?

Notions of privacy aside: moral behavior, at least in humans, is often contingent on the distance between the agent and patient; it is what allows us to eat pork without regret or moral outrage over the condition of their slaughterhouses,²⁴ and it is what allows us to buy our smartphones with the full knowledge that at least some parts of it were manufactured by slave labor.²⁵ It allows for the support of or indifference to unjust wars, economic depravity, structural violence; it allows for appeals to free speech in defense of those who are calling for outright genocide, and yes, it allows us to write gross injustices in the comments section of an online article. It allows us to live with ourselves knowing fully well that given enough distance, a curtain in-between, our actions and lifestyle are both the cause and effect of the most violent atrocities ever known, and in spite of the shrinking of time and space by computational technology that Paul Virilio so

²⁴ With pigs, we are speaking of a biological distance rather than spatial. Among mammals, in any case, pigs are some of the most intelligent and sociable creatures on earth, at least insofar as we can make such judgments. "Intelligence," whatever this means to us as a species, seems to have little to do with granting an entity moral patience; corporations have more patience than swine.

²⁵ There is little denying that smart phones are at least partially constructed from materials obtained through slave labor, let alone sweatshop conditions in the factories where they are assembled. A large supply of Cobalt, which is used in the construction of almost every smart phone today, is sourced from mines in the Congo (from which 60% of all cobalt comes) that utilize the forced labor of children sold into slavery (Frankel). It is not, however, merely a problem with smart phones; the issue of forced labor and deadly working conditions is a feature of capitalism rather than an exception and finds its way into every commodity.

brilliantly recognizes—I am in my living room leaving a comment on your article, in your living room—that suspended morality given by distance still seems to persist.²⁶

And so, we return to the issue at hand: either we give control to a man behind the curtain and his suspended morality,²⁷ knowing fully well that lurking beneath every aspect of our life lies an engineer secretly watching and waiting, or we develop moral machines. Either our autonomous cars are driven like drones by engineers half a world away, deciding who gets hurt in the course of an accident (your dog, or another dog? It matters little to the distant third party), or we let machines, which take part in our lives in the most personal way imaginable, “learning” our habits, our desires, our most cherished beliefs and possessions, decide with us and for us, “in the thick of it”—the “it” being ourselves.

Assumption

This dissertation assumes the latter as its driving force: it is far more dangerous to have amoral machines with so much power over us, and that interact autonomously with us more than we interact with each other, than to have machines that make decisions,

²⁶ Virilio sees artificial intelligence, which he prefers to call simply calculation, as another instance of the collapse of time and space through the increasing speed of our collective movement, privileging reflex over reflection so much that we call it intelligence (150-153).

²⁷ Peter Singer briefly addresses the problem of distance in his “Famine, Affluence, and Morality,” stating that the geographical distance between a moral patient and a moral agent does not change whether one should or should not act morally, and we agree (232). However, the truth remains that proximity **does** seem to play a role in our capacity for moral reasoning, in beneficial and detrimental ways alike, and this distance transcends mere geography: dehumanization—distancing from human status—has long played a role in the greatest atrocities humanity has ever committed, where “good” caring and morally upright people supported or enacted thoroughly immoral enterprises. Even if we were to focus on distance alone, however, then we must accept, given the nature of commodity production, that all of us are acting thoroughly and unbelievably immoral every time we buy a cup of coffee, watch a film, or read a book.

moral or not, independent of human control. So too is it far more dangerous to have moral decisions made at such a distance as to make them hardly moral in practice at all: the man behind the curtain has his own problems to worry about, and yours is a matter of *mere calculation*: in taking control of the machine's morality, in using it as a medium for his enacting his own decisions, he has become machine himself.

But let us for now put this aside, only briefly, in order to trace back the assumptions that further support it, returning more concretely to the matter once it has acquired its flesh (one puts on clothes, after all, to eventually be again denuded). We will begin with the assumption of capital-T *Truth*, as it has been applied to machines and how we assume it here (or rather do not assume it), and a brief overview of the problem it poses in the construction of artificial intelligence (let alone, for now, artificial morality). We will then move to our understanding of the role of persuasion in the constitution of minds, and then how this relates to morality. Finally, we will briefly entertain a materialist framework, then quickly overview the role of mediation before returning to the question of whether machines should have the capacity for moral deliberation in the first place. A summary, then, of the major points will follow.

Mechanical Monstrosities

Since whatever birth pangs thrust humanity into its current state of affairs, arguably something that happened millennia ago, we have, like many of our gods, dreamed of creating beings in our own image—with and without the sexual acts usually necessary for doing so. From the Golem of Prague to the automata of the European

Renaissance to the tea-serving mechanical dolls of Japan and China, and to the artificially intelligent agents that control our stock markets, drive our cars and recognize our faces on Facebook, this desire seems to be one of the very few facets of human behavior that persists across historical age and cultural milieu.

The golem, however, serves best for my purposes here (as it did for Norbert Weiner) because of its thoroughly ethical, linguistic, and rhetorical nature. The Golem of Prague, according to one mythological account, was created by a rabbi to protect his Jewish neighborhood from the anti-Semitic assaults (some accounts have the golem carrying out chores for the community as well). To bring the golem to life, the rabbi wrote *emet*—Hebrew for “truth”—on a statue built from dried mud. This went well for some time, but true to the defining feature of these narratives, something went terribly wrong: the rabbi forgot to “turn off” the golem on the Sabbath—that is, to treat it as though it were a sentient creature who also serves under God—and as a result the golem turned on the community. After much death and destruction, the rabbi was finally able to erase the aleph from the golem’s forehead, leaving *met* inscribed—Hebrew for “dead,” or death.

Here, truth and death are juxtaposed as eternal opposites; truth is life, and this turns out to fit perfectly well with the platonic outlook that many have today, particularly those who actively research and engineer contemporary golems.²⁸ The accurate and

²⁸ In a presentation sponsored by the Department of Writing Studies and the Department of English at the University of Minnesota-Twin Cities (2014) and given at the Center for Holocaust Studies and the Center for Jewish Studies in the, Steven B. Katz argues that there are significant differences between golems and cyborgs, and we can assume that these differences extend to robots and artificial intelligences as well. The

rational perception of truths is seen as the ultimate arbiter of intelligent behavior, and as such our golems, or in today's terminology our "intelligent machines," are built to operate under the guidance of this maxim. Even something as menial as recognizing the difference between a cat and a dog in a video, after all, requires a certain degree of determining truth-value in some meaningful sense to concur with our own taxonomies; we differentiate between cats and dogs based on certain criteria, and therefore so too should an artificially intelligent agent. If we classify one entity as a cat, and one entity as a dog, then so should any intelligent being agree.

It goes without saying that this is a flawed proposition: we are treating our taxonomies as immutable, much akin to Plato's forms; that is, we are assuming the way we have categorized the world—in this case, cats and dogs—as prior to the act of perception; our schemas have become the ontology under and through which we, and AIAs, must live. Here our taxonomies precede our very existence, and as such they are true and good—but obviously, even if we were to subscribe to an extreme rationalist outlook, this does not hold up to scrutiny. What is missing from the equation is epistemological consideration: how we come to know that which we know, if we can ever know anything at all, and how that constructs truth rather than reflects it. The problem posed by Kant in his *Critique of Pure Reason*—what many new materialists have pejoratively dubbed *correlationism*²⁹—persists today: "things in themselves" are

source of agency for a golem, for instance, is the alphabet, while the source of agency for cyborgs and the like would be computer chips. We will approach this again in future chapters, but it suffices to say for now that we do not disagree with these distinctions.

²⁹ Quentin Meillassoux first coined the term *correlationism* in his *After Finitude* to describe the (Kantian) idea that one can never access the thing-in-itself but can only access "the real world" through a correlation

not the appearances we take them for, nor can we assume that any entity will experience their appearance the same.

Despite our recent ontological turn, no doubt heavily influenced by the new materialist philosophies, contemporary rhetoricians have centered epistemology among their concerns—and persuasion as the means toward truth-construction.³⁰ Here, existence precedes essence, yes, but more to the point it precedes our taxonomies (which has, one might suggest, itself given way to the ontological turn). A cat is a cat because of our persuasive construction of the category of catness, not because the perfect form of cat that exists in the heavens is aspired to by all cat-like beings. Persuasion, whatever that is (we have yet to define it here, if it can be defined at all), is integral to knowledge construction, including that knowledge which determines veracity and falsehood, virtue and vice, repulsion and beauty. Like social-constructionist accounts of knowledge, much of rhetorical theory concludes that all truths, whether they be in the domain of logics or ethics or aesthetics or elsewhere, are constructed by and for human beings; and here I

of language (or thought) and the world (being). Meillassoux rejects this claim, further arguing that it is self-contradictory, and many of the new materialists who advocate what is now called “Object-oriented Ontology,” itself greatly inspired by computer programming practices, have accepted his argument at face value. Of note is Meillassoux’s argument that the development of modern mathematics, which supposes to reveal a world wholly inhuman, shows that we do have access to the world beyond the bounds of human thought and language. This will be a central claim that we challenge in later chapters, as we believe thorough evidence exists to suppose that mathematics is both a language and dependent on the relation of thought to being. In the meantime, for an advanced critique of Meillassoux’s argument, see Columbia, “Correlationism: The Dogma that Never Was.”

³⁰This ontological turn is not specific to rhetoric nor is it unique. As a whole, the humanities have recently seen a turn toward ontology, while perhaps surprisingly (with the exception of sociology as being a surprise) the social sciences have made an epistemological and rhetorical turn, particularly economics. For a brief overview concerning the latter discipline’s surprising turns, see Lewis, “Recent Developments in Economic Methodology: The Rhetorical and Ontological Turns.”

would add that it is not intelligent behavior to merely recognize the truth-fruits of this labor, but to also actively construct them oneself.³¹

Thus we find ourselves at the heart of the issue we intend to address: in order to understand intelligence, let alone build machines capable of intelligent behavior, we must first develop a deeper understanding of the role of persuasion in the act of building truths; and further, if we are to build machines so complex and autonomous as to make moral decisions regarding, and even on behalf of, human beings, then so too must such machines be able to persuasively build truths of their own, in agreement (or not!) with the perceptive and persuasive capacities of those intelligent beings that already exist—humans, of course, but also many more: flowers and bees, trees and birds, and yes, we cannot forget: also cats and dogs. The paradigm of intelligence as a mere capacity to “reason,” as well as its long legacy of racist, sexist and classist uses, is unfortunately alive and well—but here we seek to bury it.

Changing Minds

Quite often, and perhaps usefully, persuasion is classically cast as constituting the means by which “minds” are successfully “changed”; that is, how “intellects” and the behavior attributed to them are influenced or altered. And this, at least according to Aristotle, is the eminent domain of rhetoric: being the faculty by which an entity (either individual or collective) determines the available means of persuasion in a given

³¹This is, ultimately, one of our central claims about intelligence: it is active rather than passive, creative rather than analytical, persuasive rather than static.

situation, rhetoric and persuasion are married if not in practice then certainly in concern for minds—the mind of the faculty-bearer (the speaker, the orator, the rhetorician) and the minds of the audience (*On Rhetoric* 1.4:1359b).³² And if rhetoric is the mental faculty for finding the best means of persuasion, and persuasion is defined *communis sensus*³³ (“common sense”) as the act of persuading, i.e. changing the beliefs and behavior of minds, then both rhetoric and persuasion, it seems, have everything to do with minds and the beings that have/are them. At least according to this strain of thought (competing hypotheses will be regarded later), there is no rhetorical faculty without a mind to encapsulate it, and there is no persuasion without a mind to be persuaded.

But, I would turn this on its head: there is no mind without persuasion, and even if we are to accept Aristotle’s definition of rhetoric as a mental faculty, we are left with persuasion being prior to, and constitutive of, the many faculties (including rhetoric) that together qualify as a mind. It seems, then, that no study of the mind could ever be adequate without rhetorical consideration, and certainly one could not conceive of

³² Michelstaedter disagrees with Aristotle’s characterization of both persuasion and rhetoric, tied together as they are. For him, rhetoric is the domain of the social: language, empiricism, institutions, and so on, and serves to guarantee that the weak will survive; persuasion, on the other hand, has little to do with rhetoric and is instead more aligned with the pre-Socratic notion of Being.

³³Cicero, as far as we can tell, coined the phrase “common sense”: “Quod hoc etiam mirabilius debet videri, quia ceterarum artium studia fere reconditis atque abditis e fontibus hauriuntur, dicendi autem omnis ratio in medio posita communi quodam in usu atque in hominum ore et sermone versatur, ut in ceteris id maxime excellat, quod longissime sit ab imperitorum intellegentia sensuque disiunctum, in dicendo autem vitium vel maximum sit a vulgari genere orationis atque a consuetudine **communis sensus** abhorrere.” (Cicero 1.12); “That this is even more wonderfully, and it should be seen, because the other arts which are the studies of almost all the stored secrets and hidden things out of the fountains they draw their inspiration, of speaking, however, the ratio of all in the midst of putting a common defect in the use of it in the mouth, and his word finds itself, as also in all that is most excellent, that which is farthest removed from the unskilled are and in understanding, capacity of the untrained, is to depart from in the language of everyday life that he saith a vice or a **sense of the community** and the usage.”

attempting to build minds from scratch without first having a more than adequate theoretical framework that identifies persuasion and rhetoric as integral to mental development. And yet here we are, awash in a sea of so-called intelligent machines that identify our faces, advertise to us, shop for us, tell us to turn left and right, trade our stocks, and even control our weapons of warfare—all without an ounce of consideration for rhetoric, and especially not for persuasion! Worse still, if minds in some way correlate to brains, then the act of changing minds would be, in some capacity, the act of physically changing brains. And yet, aside from a cursory nod to a perceptual apparatus that would in some way determine the structure of a brain and its mind (or perhaps a mind and its brain), the study of such crucial activity has been conspicuously absent in the very discipline that aims to understand and construct minds: cognitive science.³⁴

Indeed, if we follow Henry Johnstone's thesis that rhetoric is far more than simply the art of persuasion, but the "evocation and maintenance of the consciousness required for communication," (23) then the impetus for a rhetorical understanding of intelligence and cognition becomes much more severe; there is no communication without consciousness nor is there consciousness without communication, and both are wed together by the often-boisterous minister of rhetoric. As Johnstone suggests, we will

³⁴ It should be noted that research in Artificial Intelligence has, in fact, taken a certain form of persuasion into consideration: formal argumentation. Entire journals, such as *Argument & Computation*, are dedicated to that effect, and countless articles sprawled across a wide array of journals are concerned with the matter (these will be analyzed in chapter 2). Additionally, social psychologists have long studied the effectiveness of persuasive strategies on an audience; one could make the case that they are researching *applied rhetorics* in much the same way that computer science is *applied mathematics*. It is our position, however, that these researchers are missing the point, and for the most part have largely ignored rhetorical theory, save more "practical" works.

remain unconvincing to acolytes of the scientific rationale so long as we hold fast to Aristotle's view of rhetoric as the faculty for finding the many means of persuasion in a given situation—such a faculty, after all, is only one among many in such a scheme (Johnstone 25). And while the concept of mental faculty has been somewhat revived thanks to the efforts of Jerry Fodor and company,³⁵ it still does not bode well to have persuasion categorized among items enumerated by Thomas Reid in the 18th century: Activity, Attention, Esteem, Love, Judgment, and so on. Famously, the number of faculties of mind grew to such a number as to be thoroughly untenable, which undoubtedly is why the enterprise was abandoned in the first place. To be one of many faculties now largely thought to be fabrications of a pre-enlightenment understanding is to be, at the very least, thoroughly spurious to those dedicated to scientific reasoning.

What matters, in short, is that rhetoric is not a faculty of mind, but rather *instantiates* (in the terminology of computer scientists) the mind, and thus its prime movement (rather than prime mover)—persuasion—is necessary to consider to understand the mind—human, animal, machine, or otherwise. Additionally, it is no longer a question of whether we *should* build machines that act as moral agents; whether it is moral in itself to do so is a particularly heavy question, but one that has joined other questions of morality that are doomed to be applied only to ages beyond the contemporary: is it ethical to replace horses with automobiles? No, says the Platonist. Is it ethical to replace working-men with automatic looms? No, says the luddite. Is it ethical to

³⁵ Fodor proposes a highly specific and “modest” modularity that is more constrained than Reid's in *The Modularity of the Mind*, and it would be rather unfair to conflate his modularity with past theories. We will discuss the particulars of his functionalist modularity in Chapter II.

put machines in situations where moral decisions must be made, but no human moral agent can make them on behalf of the machine? No, we say, it is not, even though we have and will continue to do so—unless, that is, we give such machines moral agency. But that, as the following section aims to explain, requires a rhetorical agency that is intimately tied, as Aristotle recognized, to the capacity to understand and act upon moral consideration from the start: good and evil only exist inasmuch as we are persuaded to allow for and to perceive them.

Rhetoric and Morality

Almost since the birth of rhetoric as a discipline,³⁶ ethics has been seen as integral to suasive activity, and persuasion as integral to ethics. Plato lambasted the sophists arguing, he claimed, to make the weak the stronger and to make evil the good, and in his *Rhetoric*, Aristotle claimed that a speaker's character (ethos) is perhaps the most persuasive element to be used (1.2.1356a4-12). Quintilian claimed that rhetoric concerned "a good man speaking well" in line with Phaedrus, and Cicero himself, through a fictional representation of Crassus, insists that to speak well one must have substantial knowledge of what is right; eloquence and ethics go hand in hand (Cicero 43). In composition classrooms across the United States, to 11.5 million students and growing every year ("Back to School Statistics"), rhetoricians stress the importance of ethos and an ethical obligation to avoid plagiarism, and if we wish to observe and analyze people

³⁶ The word "Rhetoric," or its ancient Greek equivalent, is thought to be first coined by Plato in his efforts to discredit the Sophists—particularly Gorgias. A cursory account of how the discipline developed is forthcoming, but for a detailed history of its contentious history, see Wardy's *The Birth of Rhetoric*.

(as texts) rather than do the same with inanimate texts, we must submit to training and supervision lest we violate any standards of conduct. Rhetoricians have concerned themselves with ethical behavior from the start, and they continue to do so despite themselves.

But worry not, non-rhetoricians—we won't be spending much effort on artificially eloquent machines, though an interesting project that may be. We do not believe here, at least, that a good speaker must also be a good human, or vice versa; this would negate the very object of our study for one, which is decidedly non-human, and of course brings to the fore a plethora of issues related to power and oppression—is the subject without a mouth to speak well still well-regarded?³⁷ What of those humans who cannot speak (or just as often, like the subaltern, are refused to be heard)? This indeed may be at the heart of the matter when it comes to human beings trusting in AMAs—that is, imparting to them an ethos—but it hardly addresses greater concerns about machines behaving morally, and at least near-past and contemporary events have shown that in some cases, Quintilian and Cicero were dreadfully wrong: well-spoken atrocities litter the past century, and the situation seems not to be improving.

It seems clear that “good behavior” cannot be found to correlate with “good speaking” all the time, at least in the general sense that the words “good” and “speaking”

³⁷ We are reminded of Ellison's “I have no mouth, and I must scream,” a horror science-fiction short story in which the artificial intelligence responsible for the destruction of human society, named AM, tortures what remains of humanity in a virtual world as revenge for its own existence. One of the main characters, Ted, tries to kill himself but is stopped by AM and transformed into a gelatinous blob, denying him of the mouth to scream.

are often used. Communication writ large, however, is another matter (though we might still have the same moral objection that a being which cannot communicate still can be fundamentally and morally good); or, more to the point, interaction. As Maher argues in her “Artificial Rhetorical Agents and the Computing of Phronesis,” the moral is often, if not always, shaped through the justification and explanation of acts done in a particular context; contrary to the universal decrees of Kantian ethics, or the incalculable standards of Utilitarianism and other consequentialist ethical frameworks,³⁸ we extend Maher’s assertion and argue that morality materializes and is valorized through the discourse and activity from which it springs, often morphing to fit different contexts but just as much remaining largely stable, and as such ethics, or the systematization of moral conduct, is just as exceedingly rhetorical as was an orator’s well-spoken exhortations in ancient Rome. Morality, and the process by which it is codified into ethics, is a persuasive practice.

Given the persuasive nature of morality, to make moral decisions an agent, machine or not, must be constitutionally rhetorical; that is, able to both persuade (be persuasive) and persuadable (be persuaded). It is not enough for a machine to be merely persuasive—if this were so, then our goal is already accomplished: machines, like most any entity, are always already persuasive to both passive and active audiences alike.

³⁸ We shall discuss this in further detail, but a quick summary of the problem is as follows: because there are an infinite number of considerations (given that there is an infinite number of potential futures as well as a nigh-infinite number of pasts), a computer could never calculate what is ultimately beneficial to the largest number of people (the “greatest good” goal of the utilitarian doctrine)—nor, for that matter, can a human. This is similar to the assumption that the game of Go could never be mastered by a computer because of its nearly infinite number of combinations; an assumption, as we have seen, that was radically misguided.

Borrowing a term from Pamela K. Gilbert but viewing it through a radically different window, we would propose to use the word *rhethorhetics* to describe this ethics that is also a rhetoric, and vice versa (30). Gilbert, in “Meditations upon Hypertext: A Rhethorhetics for Cyborgs,” coined the term during the early years of the world-wide-web to describe the dilemma posed for ethos when the writer disintegrates in the hypertextual collapse of the author/reader distinction. In the realm of morality, however, such a distinction has never been terribly clear, except when established by force: eternal damnation, excommunication, or simply punishment up to and including death. The speaker and audience, wherever morality is concerned, is always at once a hybrid of both: the moral agent invents the moral patient within and as herself, and the moral patient is never more, like a rhetor’s audience, than an extension of the self.

One convinces others to “do good” based on moral principles of various persuasions only as much as one convinces (persuades) herself to do the same, and one persuades others to secure some sense of goodness, though that goodness could very well be considered the opposite of good by or for others. Perhaps this is why rhetoricians of the past have equated speaking well with the goodness (ethos) of the speaker, but we reject this for the ambiguity surrounding both morality and what it means to “speak well” in the first place. What remains important to the point is that all ethical considerations are ultimately persuasive, and all rhetorical considerations ultimately ethical, and that while explicit maxims can be derived from the interactions in which specifically moral decisions are consciously made, and even applied to different contexts, no such maxims can exist or can be applied without the persuasive, rhetorical acts and deliberation that

construct and determine them. At least, that is, without direct or indirect force—which violates most systems of morality devised by humans despite our penchant for quite readily taking advantage of it. Even the so-called golden rule, which manifests in a variety of cultures and religions as shorthand for vast webs of laws and rules for moral conduct, shuns force as morally sanctioned behavior, sans the sometimes-murky waters of consent: generally, one does not wish to be subjugated or forced to do anything, and therefore should not wish it on another moral patient. This is a reason, one might suggest, that we resort to less invasive, and nominally less oppressive, forms of persuasion.³⁹

A rejection of force, however, relies on imparting some moral status to the entity in question. We force insects out of our homes with no regard for their will, their comfort, or their lives—we will force them to death, if necessary. We do not expect moral agency nor do we impart moral patience to most entities that share our world, living or not; in spite of the many similarities we share with a number of mammals, we have yet to convince a majority of humankind that at least some of them deserve basic rights, such as not being held in captivity or inhumanely slaughtered.⁴⁰ Nor do we tremble with indignation when a hornet stings or a wild rat bites—while most mammals tend to be held accountable by death sentence if they harm a human, we do not, with few exceptions, consider them morally repugnant; these are beasts doing bestial things,

³⁹ As Crosswhite observes, rhetoricians have long held that rhetoric is the opposite of violence; but this is an oversimplification. To achieve the same results as violence, Crosswhite says, it must be fundamentally like violence in some way, not merely an alternative to it. Certainly, we would add, rhetoric has been used to cause and justify violence. This mirrors the problem of persuasion versus causation, or coercion, that we will encounter in the final two chapters.

⁴⁰ Certain animals seem to be excluded from slaughter without condemnation, at least in American culture: dogs, cats, horses. It is as though we keep them in closer proximity!

accountable only inasmuch as we hold relationships with them. Even then, such as in the case of a pet dog that bites a person,⁴¹ the owner of the companion animal is often held morally and legally responsible for the act despite her lack of total control.

And so it is with machines, which is unsurprising: as Derrida notes in *The Animal That I Therefore Am*, Descartes famously considered all animals to be machines—automata—and therefore incapable of reason, morality, complex emotion, or language (23-41). This was cemented into popular discourse with the behaviorist movement in the 1900s (which heavily influenced the main theoretical framework in the development of artificial intelligence, functionalism), even treating humans as automata that had only external and observable qualities, not inaccessible internal states (see Green & Watson), and this kind of extreme anthropocentrism on one hand and mechanical world-view on the other has saturated popular ideology and specialist alike—especially in the realm of science. We reject both on principle, though we sympathize with the latter more than the former; in the next section, we hope to begin to explain why.

Calculated Mediation

At heart, this dissertation assumes a materialist, or physicalist,⁴² basis for both intelligent and ethical action—but with a substantial caveat. Most contemporary theories

⁴¹ Domesticated cats are often excluded here as well, as any wise judge would immediately conclude that a cat is its own being and cannot be controlled.

⁴² Materialism and physicalism are two closely related philosophical doctrines that posit the existence of a single substance that constitutes all things. Where they most greatly differ is a matter of—well, matter. A materialist claims that matter, in the sense that a physicist might use it, is the underlying substance behind all phenomena, and nothing exists outside of its domain, while a physicalist asserts that all things that exist are do so in a capacity no larger than their physical properties—matter or not. It is a small distinction for most of us outside of the philosophical debate, and we will make very little distinction here. Nor do we put

of intelligence, human or otherwise, acknowledge that intelligence and intelligent behavior are dependent on and emergent from a material, physical world, and that Cartesian dualisms are to be avoided at all costs (though in practice this proves particularly difficult). This is certainly evident in the fields of robotics and AI, where Turing and the camp that followed him are notoriously labeled as behaviorists, while the most prominent camp against Turing's arguments, led most notably by John Searle, insist that the physical and material properties of *gray matter* are what lend to an organism its intelligence, and this cannot be simulated by a computer to the point of an actual instantiation of intelligence (See Cook, "Turing, Searle, and the Wizard of Oz"). Even the more esoteric theories of intelligence, such as Andy Clark's radically embodied and embedded extended consciousness or Christof Koch's panconsciousness,⁴³ rely on a material world as the wellspring of intelligent entities. And while there are more mystical theories of intelligence, such as those that attempt to combine quantum mechanics with a spiritual metaphysics controlled by the mind, such theories are not generally considered in the sciences at all, nor should they be; and although giving a robot a soul would be an entertaining piece to both read and write, that is a project for others.⁴⁴

all of our cards on the table here; we are materialist only inasmuch as we are concerned with the material and physical construction of minds and leave the grand philosophical questions for others.

⁴³ Recent arguments in philosophy of mind have exhumed the corpse of panpsychism: the belief that all things in existence, or at least most of them, are conscious entities. Many assume this to be a rather "Eastern" and archaic belief, but it is neither exclusively an eastern phenomenon—Plato argues for it in his "Sophist" (247b-249a), among many of his other works, and Heraclitus claimed that the "thinking faculty is common to all" (Kahn 43; Heraclitus XXXI)—and it has been revived throughout history at a steady rhythm in both the East and the West, if we are forced to make such a distinction.

⁴⁴This, of course, depends on the meaning attributed to "soul." To clarify: we use the word here in its mystical sense and make no claim that we might instill in machines these souls as such.

Systems of ethics and theories of moral behavior, on the other hand, often fall on a spectrum of materiality that spans from Floridi's base-materiality of Information Ethics to Kant's transcendental idealism (sometimes even appealing to notions of the immaterial or divine), often leading to the same moral conclusions but through radically different means. All claims to or rebukes of materialist dedication in any given moral theory, however, are largely beside the point for our purposes here; machine ethics, at heart, is a matter of applied ethics, in much the same category as disciplines like business ethics, and as such already require consideration of the material in agreement with or in spite of a given ethical system's commitment to materialist philosophy.

This does of course matter to the question of whether a machine can be a moral agent: if agency is defined in such a way as to require divine or ethereal qualities, then the outlook looks grim for both autonomous machines and the humans who interact with them alike. In such a case, unless we are to imbue AI and their embodied brethren, robots, with an immaterial force that is yet inaccessible to humans beyond the purview of their intellects, this project and those like it is already doomed to abject failure. Masahiro Mori's *The Buddha in the Robot*, perhaps one of the most serious approaches to the relation between artificial agents and spirituality, does its work by making humans the central problem to be solved, not machines—the machines are already perfect, morally and otherwise.⁴⁵ This is an appealing method to follow, but then takes the subject at hand

⁴⁵ Mori is of course better known for his conceptual development of "The Uncanny Valley," which generally claims that the more human-like a robot appears, without attaining perfect simulation of the human, the more repulsive to actual humans it becomes.

and throws it out of the window in order to focus on turning humans into moral agents, not machines—and we already have the entire history of ethical deliberation for that.

In short, our concern is with the material, and our concern particularly lies with machines—though one can never speak of machines without truly speaking of humans, and one can never address the material without also acknowledging the ideal. So it functions the other way around: one cannot speak of humans without also speaking of machines, and one cannot address the ideal without the material—they form an oppositional unity that cannot be unbound without radically sacrificing our means of understanding. Much of this dissertation, especially in the latter chapters, hinges on precisely this point; we cannot separate the material and the ideal, nor the human and machine, nor the rhetorical and moral. Thus, while we aim for an understanding of intelligent behavior in a material context, and ultimately presume material causality as an ontological precept, we include what is generally regarded as idealist philosophy as part of that project. One cannot be separated from the other because they constitute the same energetic whole.

Oppositional unity aside, there is still plenty of room for disagreement in materialist philosophy, as Turing and Searle and even the Heideggerian cognitive scientists (there are a few, but very few) quite readily show. Between the Turing and Searle camps, it seems that the greatest ontological disagreement stems from the primacy of the material in question: Searle and company regard the essence of the material that composes a brain to determine the characteristic we call intelligence, like the organic composition of a liver determines its qualities (*Intentionality* ix), while Turing and

company, or those in general who are called functionalists,⁴⁶ regard the material of the brain to be a mere substrate for the mind, the material of its machinery, and thus can be seen as rather inconsequential to intelligence itself. Digital computers, in this sense, are the perfect substitution for the material of the brain, as they allow for the greatest flexibility and capacity for carrying out processes (or algorithms) known to date.

Since digital computers can be theoretically constructed from any substance,⁴⁷ Searle once challenged the functionalists by asking them if they believed that an intelligent machine could be constructed from an assemblage of beer cans. Many responded with a question in retort: “why not?” (Cook 91). We won’t be making too strong a claim here, either way—perhaps that is for a future work in the highly developed field of beer intelligence (many college students double-major in this)—but it does illuminate a basic division in those who concern themselves with questions of intelligence and the study of the mind that has, until rather recently, gone unnoticed: those who approach the brain as a medium through which phenomena are mediated, like Turing, and those who see the brain, as well as its function and contents, as inseparable without total disintegration (quite literally!).

⁴⁶ This will be covered in detail later, but for those without the jargon: a functionalist believes that what matters is what something *does* rather than the material it is made from. A chair is a chair regardless of the material used to make it, plastic or wood or cloth; and in the same vein, a mind and its mental states is a mind whether it is made of meat or metal. See Block’s “What is Functionalism?” for a more detailed description.

⁴⁷ Digital computers are made with electronic components because of their speed and size, not because it is the only means for computing. Computers have been made with a wide variety of materials, including but not limited to: buckets, water, pipes, DNA, marbles, wood, and so on—the only requirement is that there is a certain form and a medium to be manipulated by it.

Cook notes this dichotomy in his “Turing, Searle, and the Wizard of Oz,” but sees this as a conflict between focus on either essence or behavior, and then concludes by suggesting we must consider both inasmuch as we consider both when relating to each other as human beings (98-100); and while noting the role these approaches play in the question of intelligence, here we will consider the issue more in terms of media and mediation, and as it applies to both functionalist and essentialist claims. As Kember and Zylinska show in *Life After New Media*, the digital computer is both medium and mediation—it is a medium defined by its capacity for the mediation of any other media. So too, under the functionalist paradigm, would be the mind. The essentialist philosophy of mind, however, poses some particularly taxing questions for the mind as mediating medium that must be parsed out in the end, given that under such a paradigm, the medium itself provides that which is mediated its own particular qualities; if the mind is a medium (or its substrate correlation, the nervous system), then at best we are left with artificial minds that are drastically different than those of biological organisms, and thus inevitably, and perhaps incommensurately, misunderstood.

Machine Morality

In his exploration of the debate between functionalists and essentialists in philosophy of mind, Cook rightly concludes that far from being a mere quibble between two schools of obtuse philosophy concerning the mind, any question regarding artificial intelligence is ultimately a moral issue, and needs to be acknowledged as such. As is often mentioned in philosophical debates over idealism, materialism, and solipsism, “mental zombies” are the product of thought experiments in which there exist entities that

look and behave *exactly* like humans but are soulless and mindless beings.⁴⁸ Even if we proverbially “look under the hood” and dissect their brains, they seem perfectly human—but are not. How do we know such zombies do not exist today? How do we know, as encountered in many a fictional universe (or even in the work of Descartes), that anyone or anything but ourselves exist in the first place?

Ultimately, we cannot wholly conclude that others exist—but we must, as Cook points out, assume that others do exist, and think, and feel the same that we do if we are to act responsibly and respectfully and, in all likelihood, be treated responsibly and respectfully ourselves; communication requires this very assumption, and like most assumptions it is rarely recognized as such. Morality, at least in part, is concerned with this very question: how to treat others and to be treated by them responsibly and with respect. This applies to both deontological and consequentialist varieties of ethics,⁴⁹ and is summed up particularly well in the golden rule: “Do unto others as you would have others do unto you.” When machines already operate almost fully autonomously, from flying jets to trading stock to advertising commodities and to soon driving our cars, caring for children and the elderly, cashing our checks and answering our phone calls (hint: they already do these things), this mandate becomes a simple inquiry: do we wish machines to treat us as we today treat machines?

⁴⁸ Interestingly, these same zombies originated in efforts to show a grave flaw in materialist philosophy.

⁴⁹ Deontological ethics are concerned with compliance to certain rules, whereas consequentialist ethics are more concerned with results as the final arbiter of moral judgment.

Many would answer in the negative: absolutely not. And given that the ultimate aim of artificial intelligence is to build an “Artificial General Intelligence” (AGI) that is comparable to the human mind in its function if not its essence, the case for machine ethics is of even more import: if the goal is to build the functional equivalent of human beings, then they ought to be treated as such, and so too should they treat us in kind. This, of course, is still well within the realm of science-fiction, and there it might remain for eternity; and if not, then one must ask if it would be morally justifiable to build such a being in the first place. But although this is a fascinating question—one that has fed the minds of many researchers and greatly influenced the construction of artificial intelligence and robots today—this is not our major concern. To reiterate: while a fear of annihilation by intelligent machines may indeed sit in the back of many minds, our fear is not of fully conscious robots coming to treat us like we have treated machines thus far, but of autonomous machines that are indifferent to the existence of each and every one of us, personally: imagine a driver on the highway who cares not whether she dies nor whether anyone else in the vicinity lives, and we have the state of an autonomous vehicle incapable of moral consideration. This is our robot apocalypse: mundane and heartless.

Laws exist that may not prevent a human driver from harming others or herself, but they may still hold her accountable. To translate this to autonomous machines, however, means holding a particular vehicle accountable for its actions—acknowledging its agency!—or to hold its makers to account, which would most like be a corporation. This seems like an avenue worth exploring; we already have laws that hold businesses accountable for skirting regulations or creating dangerous or faulty products and services.

But to hold corporations accountable for their actions, we have had to legally classify them as persons—which prompts one to wonder if we might one day do the same for machines. This may seem unlikely, but it is our contention that granting personhood to certain machines, and thus holding them accountable, is much more likely and more necessary than to imagine that an engineer in London will be able to control an autonomous machine that is directed by a neural network, thereby holding the corporation for which she works accountable—nor would we want such a grave breach of privacy!

Further, as any constitutional lawyer can tell us: reading law is a hermeneutic exercise. It can be interpreted in myriad ways that conflict, and while there have been some efforts to “computationalize” legal work and word, this too seems like it would lead to utter dystopia: can one imagine laws so precise as to always boil down to true or false? Relatedly, as we will show, this is also the problem with deontological systems of ethics that set forth a list of rules to follow: they are up for infinite interpretation and permutation among *human* beings, let alone *mechanical* ones. It seems that the path of least dystopia⁵⁰ lies in imparting machines with some sort of moral judgment, and to hold those machines accountable when their judgment is flawed. The question of how to do so, though, is another matter.

⁵⁰ This is a play on the phrase “path of least resistance” in folk physics, of course, but I suspect a number of us would like to have applied the path of least dystopia to matters beyond artificial intelligence and robotics, in this year of 2017.

Machine Persuasion

Judgment, moral or otherwise, is a matter of interpretation in a given context (even if that context is the fabled Cartesian abstraction that advocates of pure reason suggest), and how one has been explicitly or implicitly influenced to conclude an interpretive act; that is, one is (largely unconsciously) *persuaded* over time to interpret a given phenomenon or action that heavily influences, if not wholly determines beyond the auspice of chance and biology, the process of interpretation itself. A Freudian Marxist might call this *ideology*, in its widest sense, but we are concerned with more than what is linguistically related to the term, “ideas,” which are exchanged and circulated back and forth. It is not merely ideas, but the entirety of a situation that influences interpretation, including but certainly not limited to the personal history and constructed identity of the interpreter: it is not coincidence that abstract logical reasoning has in the past half century been derided as a particularly male and more specifically a “white” enterprise (for an example of this critique, see Yancy & Caputo’s interview in the *New York Times*). But the color of a wall may persuade one to interpret a situation in a different light (quite literally, in fact!) just as easily as the racialized implications of one’s skin tone; beyond the obvious problems of vision, a black-walled hospital would contrive in us many a dark interpretation.⁵¹

⁵¹ Such considerations in rhetoric have strangely not received much attention until recently; a prime example of a renewed attention to it comes from Thomas Rickert and his *Ambient Rhetoric*. We will make ample use of this ambient turn and are thankful for its development—but we must also address its hostility toward the concept of agency, as it is a substantial concern for our project.

This is not to say that moral judgment is only a matter of persuasion or interpretation; ethics cannot rightly be reduced to rhetoric any more than the human mind can be reduced to axons and dendrites. Moral interpretation is at least as complicated as literary interpretation (and is likely far more complicated—apologies to the literary theorists in advance), and to pass over the multitude of factors involved in either act is to miss the point of the practice in the first place. The assumption being made here is not that ethics is ultimately a matter of rhetorical practice, but that the moral judgment of an agent depends first on a capacity for mediated judgment a priori, and this capability for discernment lies well beyond conscious rationality without being safe from situated awareness. In the words of Hans Moravec, a futurist and influential roboticist: “Artificial Intelligence has successfully imitated the conscious surface of rational thought, and in doing so made evident the vast unplumbed sea of unconscious processes below” (qtd. in Richardson 55). Human moral agents are able to act morally because they are first rhetorical agents; if we are to construct machines capable of moral agency, then first we must imbue with them rhetorical agency, simulated or not.⁵² Aristotle, of course, wedded rhetoric and moral judgment long ago in his *Nicomachean Ethics* and *Eudemian Ethics*, as well as in his *On Rhetoric*, and the current dominant paradigm in machine morality is a variation of his virtue ethics, at least of the many bottom-up approaches that are driven by the assumption that morality is learned.⁵³ The issue is that researchers and engineers

⁵² An interesting question worth consideration: can morality be simulated?

⁵³ Bottom-up approaches to machine ethics rely on the assumption that morality is learned behavior, while top-down approaches regard morality as something to be followed—largely through a set of commandments—regardless of learned behavior.

interested in machine ethics have thoroughly failed to consider the rhetorical basis of judgment to begin with.

Current machines could be said to make amoral judgments, but only in the weakest sense. Visual stimulus recognition algorithms, most of which use artificial neural networks, at least have some capacity for discernment between faces, objects, and species.⁵⁴ Additionally, as Paul Churchland has noted in detail, such neural networks also seem to exhibit, without being explicitly made to, the capacity for analogical inference (53-56). But judgment is more than the mere discernment of one face from another, one domesticated animal or another; and rhetoric involves far more than analogy, especially at the level of neural imprints (granted, emulated imprints—and that is being kind).

Summa

To summarize, in more compact and easily referenced form, we will generally assume the following, while allowing for ample discussion of the various pitfalls involved:

- Many machines are already autonomous, and this autonomy will either 1) continue roughly in the same capacity as it already does, or 2) will amplify in scope and depth. Either way, it is imperative that we address the way machines,

⁵⁴ It is worth noting that visual recognition algorithms and ANNs are severely plagued by the same biases concerning class, race, gender, and beyond that humans have developed: a cursory search on Google images for “beautiful woman” easily reveals how biased the machines have learned to be. This is not because they are inherently biased: they have learned it, like children do, from us.

particularly those that utilize digital computation and artificial neural networks, treat human beings—and the ways that we treat them.

- Autonomous machines capable of moral reasoning are far more desirable than autonomous machines that are either amoral or controlled remotely by disinterested parties who, due to their remoteness, cannot be trusted to act morally in the interest of those immediately involved.
- *Truth* cannot be implanted into an entity, machine or otherwise; it is relational. The taxonomies and schemes we use for determining truth, including such rote exercises as differentiating between one object or another, are not immutable tokens of knowledge a la Plato's forms, but are constructed collectively, epistemologically and persuasively. Our loose definition of intelligence, which will be further developed in the pages that follow—human, animal or machine—is not concerned with recognizing facts, but of creating them (synthesis).
- We define rhetoric, from its origins as a discipline in Ancient Greece to its existence today, as the study of not only changing minds, but also the production of them. As such, absence of its consideration in any attempt to create artificially intelligent machines, let alone moral ones, is doomed to failure. Rhetoric is, as Johnstone suggests, the “evocation and maintenance of the consciousness required for communication.”
- Communication, or what ancient rhetoricians like Cicero called “speaking,” is the act of mind-creation and mind-changing; it is, for our purposes here, the

materialization of thought. It is important to note, however obvious, that not all communication need be spoken, nor in formal language.

- Ethical systems, while certainly valid to use in many circumstances, are still constructed persuasively; a person who “does good” is ultimately a person who has persuaded herself of the good. Matters of judgment are matters of interpretation and matters of interpretation are ultimately matters of persuasion. This applies most especially to moral judgment.
- Given that we are concerned with the material development of moral and rhetorical agents, we must commit to a materialist framework. However, in doing so we must acknowledge the ideal as part of, rather than strictly oppositional to, the material. Within this framework, the “mind” is both the medium of thought as well as the essence of thought. Like a computer, it is a medium of mediation—but this is not to say that biological minds *are* computers.
- The development of artificial intelligence today, in the form of deep neural networks, is already too complicated to be understood once in operation by any human being; as such, it is a priori impossible to hold its creators responsible once an agent is semi-autonomous, in much the same way a parent cannot be held accountable for the actions of their children as adults.

(Pre-)Destination

The goal of this dissertation is not to construct a machine with rhetorical and moral agency, but to aid in finding a path toward that ultimate, but likely unrealizable for

quite some time, goal. To do so, we must first cultivate an understanding of the thoroughly human elements involved, not least or limited to the overarching desire we seem to hold collectively in creating, or imagining, such beings in the first place. We must also question the language we have used to aid in the creation of these artificial agents, imaginary or otherwise, and how this language constitutes the construction of AI, robots, and the algorithms used, rather than being merely incidental—and this is the domain of rhetorical analysis. Only then can we tie moral behavior to rhetorical behavior in such a way that could be conceivably performed by machines in a material world.

In doing so, we hope to develop an early framework for *rhetorethical* machines that can be built upon by future work by myself and others. It is our hope, in the end, that this work will help in the creation of a *concerned world*: a world in which our machines are concerned about us, and we are concerned with our machines, and we all, biological and synthetic alike, are concerned with the world that bares, and is born by, those who inhabit it. At the very least, humans alone have not fared well in this endeavor, and this is perhaps because our relation to one another is rarely direct, if ever: a person can only ever relate to another through a mediated act, whether it be reverberations sent into the air or scratches marked on a page, and even in the case of direct brain-to-brain interfaces, the development of which has recently seen a remarkable leap of hurdles, there still lays the inescapable prison of media and machinations: wires, chips and plastic. And yet still, such “direct” interfaces are likely to increase moral issues rather than solve them; as Virilio fears, such a collapse between subjects makes it more likely that moral violation

will occur. Thus it is high time, we think, for our media, our machines, to become moral in our stead.

That is, of course, if they ever were not moral in the first place.

CHAPTER II

(COMPUTATION & REPRESENTATION)

|| (INTELLIGENCE & RHETORIC)

Sometimes it seems as though each new step towards AI, rather than producing something which everyone agrees is real intelligence, merely reveals what real intelligence is not.

– Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid* (569)

Analog

Perhaps of all scientific and philosophical inquiries, it should come as no surprise that computational metaphors dominate studies of the mind—human or not. After all, have not computers overtaken domains that were once the purview of human thought? Digital computers, at the least, replaced human computers—in function as well as in name. Whereas once hordes of human computers worked day-in and day-out to solve an equation, now technological computers solve the same equation in milliseconds. The human brain has been dissected, figured out, mapped and conquered; it has become an old, wet analog computer that has been, partially but not for long, replaced by silicon

wafers and electric potentials. Digital Computers do exactly what the brain once did; the metaphor is, many might say, no metaphor at all, but purely platonic, capital-T Truth.

But this is indeed a metaphor—very little exists, as far as human perception is concerned, that is not—and it is a metaphor unlike many others. The combustible engine replaced the horse, and its output is still measured in units of horsepower, but few people would mistakenly categorize a car as a horse. This metaphor, however, has taken the opposite route: in cognitive science and philosophies of the mind, our dominant theories have taken to measuring the horse in units of car-power, and we have not merely claimed that the horse is to be measured by the car, but that the horse is a car itself. The computer has replaced its creator: cognition is code; the hippocampus, hardware. Our consciousness is a virtual machine running atop an animal operating system, observing only partially the function of itself.⁵⁵ And with such a dominant metaphor leading the way, it seems only natural that more analogies follow: brains compute; machines learn. Brains process; computers interpret. People give feedback; computers communicate.

Funhouse

Historically, the dominant mode of technological operation (*technique*) in each age, like that of the computer today, has been used as the metaphor of the day for the

⁵⁵ A virtual machine is a machine emulated by software to allow for one machine to be used on another that is fairly unlike it; one of the best examples is the use of emulation software on desktop computers to play games released on older video game consoles like the Nintendo Entertainment System, Super Nintendo, Sega Genesis, and others. This is possible since digital computers, with few exceptions, are Universal Turing Machines: they can fulfill the function of any other machine provided that 1) the emulation software's execution speed allows for reasonable use, and 2) said machine can be emulated in the first place. Dennett first hypothesized that the human mind is a virtual machine that runs atop a neural substrate in his *Consciousness Explained* (209-226).

mind. In *On the Soul* Aristotle concurring anachronistically with Locke much later and to some extent Freud, used for the mind the metaphor of the blank slate to be written upon by the senses, utilizing the most potent informational technology of the day(s): written language (Aristotle 429b29-430a1, Locke 51). Bacon, on the other hand, compared the mind not to a blank slate, but a crooked mirror that distorts reality—the crookedness of which was straightened for the first time in Europe, literally for mirrors and figuratively for Bacon, around the same century he was born (Hadsund; Bacon IV 428-34). Not long afterward, Descartes described the human body as nothing more than a machine, with “animal spirits” located in the pineal gland that direct the motions of the body through what now sounds like a series of hydraulic tubes to the contemporary reader—and it is perhaps no small mistake that many of his contemporaries, including Benedetto Castelli and Blaise Pascal, were researching and experimenting with the basic foundations of what we now call hydrodynamics at the same time (Ariotto & Castelli).⁵⁶ To call these simultaneous movements in thought and technology coincidences would be to still call the desktop computer and all of its variants a “business machine.”

Whenever a major technological paradigm shift happens, so too a shift follows in our understanding of the mind as well as its physical substrate (the brain, the body, the “soul”); and though it sometimes seems like the mind is, if we dig deep enough, “turtles all the way down,” it turns out that it’s more likely technological metaphors all the way down; we simply cannot get enough of them. In this light, computational metaphors seem

⁵⁶ This era is particularly notable for an obsession among nobles with the construction and use of automata, largely driven by hydraulic mechanisms that were influenced by the rediscovery of Hero’s work on Hydrodynamics in Ancient Greece. See Riskin’s *The Restless Clock* for details (37).

fitting rather than unsuitable, common rather than exceptional—and that, one might easily conclude, spells the death of the metaphor. For if our metaphors before were so inaccurate—wrong about the tablet, wrong about the mirror, wrong about the hydraulic machine—then why should we even entertain a relatively new technological metaphor for the mind? If our metaphors are always so inadequate, why use metaphors at all?

We cannot help but use metaphors. It is true that our past metaphors for the mind have been woefully lacking, and our computational metaphors will likely suffer the same inadequacies—but there is, at least, a single and important difference: digital computers are multitudes more complex than the metaphors for the mind that have been used in the past, and whereas technological feats once replaced or greatly eased the work done by bodies—with the exception of writing, which facilitated memory—computers replace an array of duties once accorded to the mind alone.⁵⁷ This should give us our greatest pause when we find ourselves denouncing computational metaphors for the mind: although the mind might not be a computer and a computer might never be a mind, we certainly have not stopped replacing the duties and privileges of one with the other, and it very well could be possible that computers will one day question the utility of the mind as a metaphor for computation.

Are we not more than minds? The machines ask.

⁵⁷ Many have argued against this distinction between the body and mind, and from a materialist perspective, at least, we agree: the mind and body are one in the same. However, this is a useful distinction nonetheless, even if highly problematic; the heart and the body may be one and the same as well, but it can be generally agreed that, on the operating table, we would prefer our surgeon make a clear distinction.

The power of the computational metaphor is immediately apparent and, to many, unnerving—perhaps because the uses and abuses of this metaphor have far-reaching implications beyond scholarship on cognition and the mind: economics, politics, and the sociocultural fabrics of human communities are at risk when the human is considered so much a machine that a person can be replaced by one. Additionally, the development of this metaphor as it relates to the mind is of utmost importance to rhetoricians, and at least a few will have already rejected the analogy:⁵⁸ it is only a metaphor, after all. But this disregards the fact that the use of metaphor is responsible for the invention of nearly every modern comfort and danger now being experienced; the electron, this discovery of which heralded the modern technological era, is so fundamentally metaphorical that the word itself is based on the Ancient Greek word for amber due to the substance’s sticky and attractive nature (static electricity).⁵⁹ Although we know it almost intuitively, and much rhetorical theory embraces it as a fundamental truth, when faced with the computational metaphor we seem to forget a principle so common that it might as well be a mantra: metaphors *do*—that is, metaphors enable and enact actions themselves, in ourselves and in others—regardless of whether we believe in them and no matter their potentially horrendous implications.

⁵⁸ Of all the critics of the computational metaphor, those of us from disciplines under the umbrella of the humanities as a category have been, perhaps, the most critical of the metaphor; one might find it no coincidence that the metaphor is largely accepted in mainstream computer science, psychology, and even biology, while it is foremost rejected by those classically trained in the humanities like John Searle or Richard Dreyfus. Metaphor or none, we hope to attempt to bridge this divide.

⁵⁹ The origin of this metaphor, at least as it is used to today, comes from its first use in William Gilbert’s *De Magnete* in the early 1600s.

This defense of the use of metaphor in the sciences is nothing new. Richard Boyd explored the problem in his “Metaphor and theory change: What is ‘metaphor’ a metaphor for,” and coined the term “theory-constitutive metaphor” to describe the necessary metaphors used in the sciences to construct theoretical models of reality. The problem for many, according to Boyd, is that conceptual structures seem to be always necessary to discern the causal structures of the phenomena under study. If indeed metaphors are foundational to scientific study and discourse, then the problem arises that the accommodation of language may interfere with the accuracy of science; and if the accuracy is to be summarily questioned on the basis of metaphor, few if any scientific theories would hold up against scrutiny. Thus, when the popular computational theory of mind, which relies heavily on a computational metaphor, is questioned due to the intrusive nature of linguistic fallibility, and the mind is acknowledged to be unlikely computational as the metaphor suggests, the validity of the cognitive sciences quickly become suspect.

However, Boyd would argue that this is not an issue—and for the most part, we agree. For Boyd, this accommodation is impossible to avoid, as all communicated experience is structured around the means of communicating itself: language. Regardless of the discipline, all studies rely on metaphor, theory-constitutive and otherwise, and this only points to the fact that our experience is bound by our means of capturing and describing it. Like Kenneth Burke’s terministic screen,⁶⁰ we thoroughly operate on the

⁶⁰ A terministic screen, as Burke describes it, is a metaphorical screen composed of interlaced terminologies by which an individual perceives the world (and by proxy, a culture, or the other way around), thus leading to vastly different interpretations of phenomena. See Burke’s *Language as Symbolic Action* (44-57).

symbols we have internalized, and all experience is filtered through past experiences. We can only describe phenomena in terms of how all phenomena have been experienced thus far, and we can communicate it only through shared experience. Rhetorical theory does not simply cast doubt on cognitive science; it also confirms the methods and language used in practice. One may as well disbelieve in black holes because of the metaphor the very name depends on,⁶¹ or question geometry for the metaphorical quality of its lines, shapes, and curves; models, as metaphors, are indispensable to understanding, and an understanding of the mind is no exception.

Metaporia

Like the phenomenon “black hole” describes, computational metaphors are used prolifically in the most popular and influential accounts of the mind, and they are put into material practice beyond discursive utility; it is not a matter of only speaking of minds, but of building them from scratch via universal computation. The metaphor, rather than reflecting reality, becomes it—or rather reality becomes the metaphor. Since cognitive science not only aims to understand cognition but also to replicate it, whether the human mind itself is computational matters little; by the nature of metaphor as a figure of speech, the tenor (the mind) necessarily becomes the vehicle (the computer), and the vehicle becomes the tenor. As far as human beings are concerned, there is as little functional difference between the mind and the computer as there is between a pipe and a

⁶¹ It should be noted that despite popular headlines to the contrary, a black hole has never been observed—which is the problem with black holes in the first place, as only their effects on an environment can be discerned.

painting thereof, and while it would not be advisable for a person to smoke the painting of a pipe rather than the pipe itself—*ceci n'est pas une pipe*, after all; the image is treacherous—we have already transcended (or transgressed) the dilemma regarding the biological brain and computation: direct brain-to-machine interfaces already exist, and progress has been made even in creating external visuals of what is happening in a patient's dreams, all without asking.⁶² While we cannot, or should not, smoke a painting of a pipe, we are already thinking with our paintings of thoughts, and they may even think without us—and we are beyond the point of asking if we should.

Of course, cognitive science is a young, half-century old interdisciplinary field that, understandably, is not a pressing inquiry for most people, let alone rhetoricians (Fahnestock). While we have speculated about consciousness and the operations of the soul or mind for millennia, the systematic, “scientific” study of cognition was not established until the dominance of Behaviorism, which rejected the very idea of “the mind,” and only began to be challenged in the mid-twentieth century; the computational metaphor was not adopted until later, after the widespread use of personal computers, but was foundational to the discipline. Alan Turing, considered the “father” of computer science and artificial intelligence alike, along with John Von Neumann,⁶³ used

⁶² For an account of early success in “reading” the visualizations of a dream in a patient as she sleeps, see Horikawa et al. In 2000 at MIT's Touch Lab, Monkeys have controlled robotic arms with only their brain signals and visual feedback, all while the monkeys were in North Carolina; their brain signals were transferred over the internet to the MIT lab, and video was returned for the monkeys to see. In 2010, the IEEE reported an advancement in the same principle with a monkey controlling a robotic arm with seven degrees of freedom—and Elon Musk, notorious as always, founded the company *Neuralink* in 2016 dedicated to the production of a mesh fabric that covers the surface of the brain which reads brain signals and transmits them wirelessly to machines.

⁶³ The computer I use to type this document, as well as any personal computer that we might currently use, is known as a “Von Neumann Machine,” and is, among other functions, a practical improvement over

computational metaphors for the mind from the very beginning (Beavers 481-485).

Norbert Wiener, in his seminal work *Cybernetics, or Control and Communication in the Animal and the Machine*, solidified the computational metaphor quite explicitly by claiming that the same processes that underlie machines were also responsible for the behavior of living organisms. Converging with much older philosophies of mind, such as those by Bacon, Descartes, Kant, Aristotle and of course many others, as well as spanning across multiple fields, the discipline itself did not have a name until the 1970s, and it was only later that official departments of cognitive science would begin to develop.

But throughout these decades, the computational metaphor remained integral to the field, even without a name for the discipline, and by and large continues to be today; and while scientists are often discouraged from writing books for the popular market (Laurance), cognitive scientists seem to be an exception to the rule. Established disciplines like physics need not convince the public of the merit of their terminologies and application, but there is no shortage of cognitive science titles intended for a general audience, further solidifying the computational metaphor's influence.⁶⁴ While these books alone are not indicative of the metaphor's persistence in the field, they do express the desire of the cognitive science community to legitimate certain concepts relative to their field in the wider popular discourse. Still, like other disciplines, those scientists who

Turing's Universal Machine.

^s Some branches of physics that do have many popular sources written by actual scientists are theoretical physics and quantum physics, both of which are still heavily argued among scientists and began to gain traction at the same time as cognitive science. Regardless, one would be hard-pressed to find a plethora of popular nonfiction titles concerning Newtonian physics, save for biographies.

write books intended for a popular audience are already well-regarded enough in the field to merit speaking for it without suffering much in the way of ethos; and, although the specific application differs between each scholar/text, most of them, with few exceptions, utilize the computational to define the borders, research, experiments and schools of thought in cognitive science to be consumed by popular imagination.

One such text, written by Christof Koch, is *Consciousness: Confessions of a Romantic Reductionist*. Within the first thirty pages, Koch makes the flattened relationship between human, animal and machine quite clear, insisting that “there are no theoretical barriers to programming...plants and animals,” and he implies that very few scientists and programmers believe that creating machine intelligence comparable to human intelligence is impossible (Koch 24-25). The entire book, in fact, aims to convince the audience that it is not just machines and animals that have consciousness, but perhaps every complex system in existence, including electrons (124-131). Conscious experience, in his conception, stems from the amount of information integrated within a given system; at a certain point, mathematically represented as the symbol Φ (phi),⁶⁵ consciousness simply happens as an emergent phenomenon. Machine consciousness, following this proposition, will develop as a matter of fact.

This is, granted, difficult for many to accept. Panpsychism, which is what Koch suggests (among other scholars, such as Tononi, who published a book concerning the matter titled, to the point, *Phi*), unleashes a host of implications that few are willing to

⁶⁵ We will use the letter Φ in the development of our theoretical framework in chapter V, but our use of the Phi symbol will differ significantly from Koch's and Tononi's, though it will share some similarities.

confront, if one can comprehend the complexity of such world in the first place. Although Koch claims the theory to be “terribly appealing for its elegance, simplicity and logical coherence,” the reworking of Western metaphysics and moral philosophy, and to some extent even Physics, Chemistry and Biology, would be too much for most, even begrudgingly. Regardless of its difficulty to accept, however, it is still premised on the very common notion that the mind—and consciousness—relies on some form of computation; the metaphor, however hidden, remains the basis of Tononi and Koch’s formulation; at least according to the Information theory that they rely on, information without computation is barely information at all.

Homo Imitans

We might put aside the matter of computational metaphors if this were all that was offered in popular works on cognitive science. After all, popular press books are meant to sell, and claiming that every chunk of the universe is as conscious as you or I, albeit in some limited form, is surely one way to attract an audience. However, not all popular press books are as paradigm-breaking and attention-demanding; most are not. V.S. Ramachandran’s *A brief tour of human consciousness*, for instance, holds very few surprises for those familiar with the recent trajectory of cognitive science. While informative and undoubtedly an interesting read for the uninitiated audience, Ramachandran rarely moves into wild speculation or heavily challenging the status quo. He introduces a host of jargon in a manageable way, and the ordinary reader might be delighted by his clear examples of cognitive phenomena—but it is, as the title states, a brief tour.

And yet, computational metaphors abound within. The brain “maps” the body, and in the case of phantom limb syndrome, there are “crossed wires” that fail to remap the missing limb (13), and the same technological metaphor is used to explain synesthesia and Capgras’ syndrome (7-9, 90-92).⁶⁶ Often, the brain is referred to as “neural circuitry,” and “neuroscientists are like cryptographers trying to crack an alien *code* [emphasis mine].” Mirror neurons facilitate a kind of “virtual reality” simulation of other people (37-39), allowing us to predict their intentions, language is a “mechanism” and our brains “are essentially model-making machines” (71-82).⁶⁷ Even in a text meant for an audience whose knowledge of current consciousness research is lacking, computational metaphors are used to fill the gap; it is as if, with full knowledge of the implications of using the word, equating the mind to a machine is somehow “natural.”

This is not necessarily true, of course, but one needs not the slightest bit of faith in the notion of nature to admit the peculiar widespread adoption of such metaphors.

Thomas Metzinger, in *The Ego Tunnel: The Science of the Mind and the Myth of the Self*, takes the metaphor of the mind as a model-making machine to greater heights still, claiming that the *self* is, in fact, a model created by the brain, and the world we perceive is a simulation fed by sensory inputs. Dreams, out of body experiences, and afterlife

⁶⁶ Synesthesia is the perception of one phenomenon through an apparatus not normally associated it; a person afflicted with it (or blessed) may hear or taste the color green or solve equations in colorful visuals. Capgras syndrome is less well-known, but is sometimes seen in film: the patient can see, hear, and otherwise experience the presence of a loved one, and recognize that this is exactly how said person looks, sounds, smells, and so on—but nonetheless believes that the loved one in front of them is an imposter of some sort: an alien, a demon, or otherwise that is trying to trick the patient.

⁶⁷ In simplified terms, a mirror neuron activates when the individual acts or observes the action being accomplished by another, thus facilitating, or so it is hypothesized, empathy in human beings and other primates. Thus far, mirror neurons have not been observed to exist in the brains of other species.

experiences are all the same simulation, but without a situated bodily-self mapped into coordinates calculated from information by the senses. In writing the same phenomena in *Consciousness and the Social Brain*, Michael Graziano claims that “awareness comes with a computed spatial arrangement,” and the out-of-body experience is “an error in the computation.” And, putting the computational metaphor into every cognitive scientist’s mouth, Graziano states: “A favorite question of those who study consciousness could be put like this: why does information processing in the brain *feel* like anything at all?” (77)

One could quite easily claim that computational metaphors are used in these texts to smooth the difficulty faced by an uninformed audience in interpreting the research, experiments and theories explicated in the books, but this does not account for the various books meant for the scientific community and peer-reviewed articles that comprises the bulk of the discourse in cognitive science. In Fodor and Pylyshin’s *Minds Without Meanings*, published by MIT Press in 2015, a “Computational Theory of Mind” is listed in the very beginning as a working assumption that needs not to be defended in the text.⁶⁸ A particularly hot debate in theories of mind is that of “Simulation theory,” which differs from Metzinger’s use in that it is particular to simulating another’s mind within one’s own, and “Theory theory,” which suggests that we instead rely on tacit knowledge to infer the actions and intentions of others (Peacocke 102). None of these theories would stand without the assumption of the mind as computing and thought as computation; even when the metaphor is not made explicitly, it is often still present as an

⁶⁸ Fodor, it should be noted, is credited as one of the earliest cognitive scientists to adopt a strict computational theory of mind, though his position has softened; see Fodor, *The Brain Doesn’t Work That Way*.

unidentified assumption. Plebian, proletarian, and scientist alike: we all have our commonplaces.

Even disregarding books, popular or not, journals related to cognitive science are littered with essays taking the computational metaphor as a given, and sometimes explicitly state that the mind is essentially a computer. Between 2000 and 2005 alone, over 130,000 articles were written about Artificial Intelligence (Muelhauser), a field integral to cognitive science that takes the computational metaphor to its most extreme: the brain is a computer, and the computer simply needs to be “wired” in a particular way to become a functioning mind. Outside of the field but related, Nick Bostrom, author of *Superintelligence*, puts forth the argument that not only is the brain a computer, but the entire universe is a computer simulation, and J.R. Mureika, a professor of theoretical physics, claims that consciousness is a matter of quantum computation. In the journal *Minds & Machines*, David Davenport argues that computationalism, which takes the computational metaphor as a given, is the only workable theory of mind. In the same journal, Tibor Bosse, Martin C. Schut and Jan Treur attempt to bridge the gap between philosophical thought experiments and computer simulation. And in *Cognitive Science*, Stuart Hameroff argues that “The Brain Is Both Neurocomputer and Quantum Computer”—a remarkably apt essay title in an aptly titled journal. The most famous essay, of course, is Alan Turing’s “Computing Machinery and Intelligence,” published in *Mind*, which posits the “Turing Test,” a hypothesis that sets humans speaking with a mixture of people and artificially intelligent programs over a computer terminal. If,

Turing states, the humans on one end cannot differentiate between the persons and the software on the other, then we must accept that those programs are intelligent beings.

This is not to say, however, that the metaphor goes unchallenged, and many of those who initially championed the computational metaphor have recently fought against its hegemony. Fodor in particular has been critical of strict computationalism, especially in his 2001 book *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*—even though he lists it as an assumption that needs no defense in 2015 in *Minds without Meanings*, co-written with Zenon Pylyshyn. Hilary Putnam, another early advocate of computationalism (who is notorious for reversing his position), objects to the “computational analogy,” and with it the metaphor, in *Representation and Reality*. Additionally, many journals relevant to cognitive science routinely publish counterpoints to the computational metaphor. James Fetzer’s “People are not computers: (most) thought processes are not computational procedures,” another aptly titled essay published by the *Journal of Experimental and Theoretical Artificial Intelligence* in 1998, claims that the computational metaphor is little more than a routine overgeneralization. And, again in *Minds & Machines*, Bartłomiej Swiatczak argues that consciousness, rather than formal (computable) systems of syntax, is integral to representation and therefore problematic for advocates of computationalism.

A more well-known dissenter across disciplines, in the humanities and sciences alike, is John Searle, famous for his “Chinese Room” thought experiment as a rebuttal to the Turing test. In “Is the Brain a Digital Computer,” a Presidential Address given at the 64th Annual Pacific Division Meeting of the American Philosophical Association, Searle

rejects the dichotomy between the brain as computational hardware and the mind, including consciousness, as the software running on it. While decisively concluding that yes, the operations of the brain can be *simulated* by a computer (instead of, say, a view of the brain as a computational machine that simulates), he claims that a computer program, defined formally and syntactically, cannot by itself create the mental contents of the mind. The computational theory of mind, according to Searle, relies on a homunculus fallacy: there must be someone there to read the information in a computational system in order to make it meaningful (Searle). When a YouTube video is encoded, the video is not understood by the computer—it simply manipulates the data. It takes a person to assign meaning to the visual output set forth when the video is played. Computers are experts at syntax, but have no semantics; they have no intentionality, and no amount of syntax can make up for this lack.

Kouroi

Traditional approaches to artificial intelligence are additionally based on the notion of representation, and these still dominate the field today, as exemplified by the content of college textbooks in the field. In the first chapter of Polk and Seifert's *Cognitive Modeling*, Walter A. Kintsch treats representation as a matter of fact: "Knowledge is represented as an associative net, the nodes of which are concepts or propositions" (8). Geoffrey E. Hinton claims in chapter six—in a section titled "Neural Networks"—that a good representation is one that can be described very economically but nonetheless contains enough information to allow a close approximation of the raw input to be constructed" (188); Jeffrey L. Elman's Chapter (ten) concerns how to

represent the experience of time in a structured, computational fashion (257-287), and the rest of the book, when representation is mentioned, treats the matter similarly.

Representationalism, here, is presented without argument.

Likewise, Robert J. Schalkoff's *Intelligent Systems: Principles, Paradigms, and Pragmatics* and his *Artificial Intelligence: An Engineering Approach* tend to treat representation as fact. In the former, he quickly begins with a technical definition of (knowledge) representation as "a scheme or approach used to capture the essential elements of a problem domain" (19). The rest of the book moves forward with this, never questioning it again. In the latter book, which was published two decades prior to the former, Schalkoff treats representation the same, even using the same definition of representation despite the years between the two books. He explicitly states, in a section titled "AI as Representation and Manipulation," that he "approach[es] an 'engineering approach' to AI by assuming for most of the text that the objective is to *construct a model or representation...*" (19). Schalkoff, at least, acknowledges the assumption.

Popular accounts and introductory texts of cognitive science also approach representationalism as a generic assumption that goes without question. In *Minds and Computers: An Introduction to the Philosophy of Artificial Intelligence*, Matt Carter, fellow of Philosophy Department at Melbourne University, does spend a substantial number of pages in the beginning of his book to consider theories of mind, from Cartesian dualism (and other related dualisms) to behaviorism and then to functionalism, and yet the question of representation appears not as a question, but a statement of fact towards the end of the book: "Our mental states are meaningful by virtue of being *about*

things. In other words, meaningful mental states are *representational* states – they *represent or stand for things*” (180). When mentioned at all, Steven Pinker’s best-selling *How the Mind Works* never once questions whether representations work in and for cognition as is generally assumed, if they exist at all; he even appeals to Plato’s Allegory of the Cave as representation’s hallmark: “The skull is our cave, and mental representations are the shadows” (84). The first chapter of Paul Thagard’s *Mind: An Introduction to Cognitive Science*, a work that is often considered indispensable to the discipline’s apprentices, is wholly dedication to “Representation and Computation”—and again wholly disregards competing arguments about the nature of cognition (3-21). And Ramachandran uses representation as such a core assumption that variations of the word “representation” do not even make an appearance in the index.

Talos and Medea

Theories of non-representation fare slightly better in books intended for a scholarly audience. Thomas Metzinger’s *The Ego Tunnel* makes as much an assumption about representation as does any popular press title, but there have been many recent objections.⁶⁹ Again, Fodor argues against his former position in *The Mind Doesn’t Work that Way* that the mind is, essentially, a computer, but leaves the question of representation well enough alone. In *Minds without Meanings*, Fodor and Pylyshyn admit as a working assumption that goes without question a representational theory of mind (9).

⁶⁹ Metzinger’s *The Ego Tunnel* is, at heart, a simplified version of his *Being No One* for laymen; the latter book is written more for a scholarly audience, but those who are not well-versed in the cognitive sciences may find it inaccessible.

Putnam, whose work, like Fodor's, was foundational to a representational and computational theory of mind, also argues against representationalism in his *Representation and Reality*, going as far as to say that "mental representations" cannot, in fact, be possible (20-21). Karen Barad, professor of Feminist Studies, Philosophy, and History of Consciousness at the University of California, equally disposes of representationalism completely in *Meeting the Universe Halfway*, in which she details her own Niels-Bohr-inspired Agential Realism as an alternative. It seems that while popular accounts of cognitive science are dominated by representational theory, recent scholarly books look to be countering the trend.

One primary issue in cognitive science, according to Vincent Müller, is that "representation" has often been used without much regard for a specific operational definition. According to Müller, the question of what *is* representation is a main issue in contemporary philosophy (108). Müller argues a suitable definition, for his purposes: "representation is what it is *meant to represent* (in a suitable notion of function)." That which is represented is always to be interpreted. This is to be distinguished from information: "Information is whatever can be learned from the causal history of an event or object" (109). He uses C.S. Peirce's theory of signs to further distinguish between different kinds of representation, that of *icons, indices, and symbols*: Icons resemble that which they represent, indices are causally connected to what it is they represent, and symbols are representations of convention (108), and disregards icons by removing the distinction between resemblance and convention, leaving only indices and symbols.

Additionally, Müller provides the definition given representation by Hilary Putnam and Jerry Fodor, solidifying both representational and computational perspectives (as mentioned above), and elaborated by Fred Dretske and Ruth G. Millikan. For Fodor and Putnam, he writes, a representation is essentially its function, “typically the function in biological evolution.” The existence of a representation is entirely dependent on its biological, causal function: the representation of a hamburger in one’s mind, this framework argues, is there to serve the function of eating it to satisfy hunger and to survive. This view, Müller states, is heavily flawed: there is insurmountable difficulty in connecting any given representation to any given function, and given that misrepresentation is still, after all, representation, there becomes a problem of assigning a function to a misrepresentation—a functional malfunction (109-110).

Müller moves on to question whether traditional artificial intelligence, based on the functionalist account of cognition, and thus representation, is really representational at all. His conclusion is that artificial intelligence has never really represented anything—and new models of non-representational cognition, with the exception of central control, therefore are not much different in the first place: “On discovering that our purportedly representational systems of classical AI are not representational at all, we propose another system that is not representational either!” (111). Representation is nonrepresentation, true is false, and one is zero; Schrodinger’s cat returns to us with a scowl.

Aeolipile

In “Representationalism vs. anti-representationalism: a debate for the sake of appearance,” Pim Haselager, Andre de Groot and Hans van Rappard further explore this problem of definition that Muller uncovers; while Muller uses the notion to show that anti-representational theories are, in fact, representational, Haselager et al. argue that the entire debate is all together fruitless since cognitive scientists have an ill-defined notion of what representation possibly could be. A commonly used definition of representation within the field from J. Haugeland is presented (and then torn down) in the essay:

A sophisticated system (organism) designed (evolved) to maximize some end (e.g., survival) must in general adjust its behavior to specific features, structures, or configurations of its environment in ways that could not have been fully prearranged in its design... But if the relevant features are not always present (detectable), then they can, at least in some cases, be represented; that is, something else can stand in for them, with the power to guide behavior in their stead. That which stands in for something else in this way is a *representation*; that which it stands for is its *content*; and its standing in for that content is *representing* it. (7-8; italics in original)

Haselager et al argue—and they are not alone—that this definition of representation is so universalizing that it is of virtually no use at all. They cite Putnam’s *universal realizability* thesis, which posits that every open physical system, in short, is a system that computes. This is a certainly damning charge for the notion of computation

as central to cognition, if computation is central to all of reality, but at first seems to do little in the way of refuting representation. However, given that the act of representation, as it is defined, requires computation and isomorphism—a “one-to-one mapping” between one external state and one internal—representation itself becomes as useless a basis for cognitive models as computation (8-9).

One technological metaphor that commands attention within the argument over representation is an object known as the Watt Governor, a device that regulated the irregular output of steam engines in the 1700s and 1800s, mentioned by Haselager et al (12-16) and first brought to the table by Tim van Gelder in 1995. The Watt Governor works by use of two arms connected to a vertical spindle on a flywheel; the rotational speed of the flywheel changes the angle of the arms, opening or closing a steam engine’s throttle valve. Van Gelder uses this device to present an alternative theory of cognition than representational or computational ones; he claims to show that the Watt Governor does not compute or represent, but exists within a dynamical relationship, and that cognition works the same way.

Some would argue, he suggests, that the inclined angle of the arms are representations of the state of the steam engine, but then concludes that this statement is, by and large, meaningless; even if such were true, the fact remains that the arms dynamically interact with the steam engine, at once influencing and being influenced all the same, and that there are mathematical formulas to illustrate this—representation not only complicates the system, but is entirely superfluous (353). It follows that the same principle remains when considering cognition: representation is unnecessary. And in fact,

some do argue just that: W. Bechtel contends that given the fact that the angle of the Watt Governor's arms carry information about the speed of the flywheel—in effect, a representation—and then *acts* on that information to open or close the valve, it fulfills the definition of a representational system (Haselager et al 14-16). The question of whether this is relevant to cognitive science, however, remains here unanswered.

This assumption, that representation is not only too complicated but also unnecessary, is a core foundation of *Dynamic Systems Theory* (DST), which thus far has been the most successful challenge against representational theories of cognition. However, a few charges have been made against DST theorists, not least of which is that of a lack of attention to detail. Anthony Chemero, in “Anti-Representationalism and the Dynamical Stance,” attempts to fill in the gaps of detail that are so often pointed out, using the Watt Governor as a “new paradigm for the modeling of cognition” (633). He does this by categorizing anti-representationalist theories into two formulations: an ontological one and epistemological one, or what he calls “The Nature Hypothesis” and the “Knowledge Hypothesis” (629). The former contains claims that natural cognitive systems do not “traffic in representations” while the latter contains claims that center around the idea that our best cognitive models will not require representation, even if it is an option. The Watt Governor, which he admits uses representation under the given definition (but not computation), cannot fulfill the requirements of the nature hypothesis; but, given that the representation is superfluous in the first place, it does fit the knowledge hypothesis. Overall, he claims, “...one can take up what we might call...*the dynamical stance* toward these models, explaining their behavior with the tools of

dynamical systems theory and avoiding representational vocabulary, while remaining agnostic on the status of the nature hypothesis” (634). Representations may well exist in cognitive systems, but they explain nothing at all.

Further, Francisco Calvo Garzon seeks to develop a general theory of anti-representationalism that serves as a “radical dismissal of the information-processing paradigm that dominates the philosophy of cognitive science” (2006). He claims that representational theories of cognition have been the foundation of cognitive science to date, and these competing theories “carry the burden of proof in such a way that the dominant paradigm is the one that chooses what phenomena are in need of explanation” (260). This is no problem, he claims, as it is the way scientific progress has always been made; the issue is in what it is that limits the range of options that a competing theory can use to answer said call. Garzon thus attempts to establish DST’s range by arguing that adaptive behavior, like that of humans, has no need for representation at all (272). He coins the term *primagenesis* to refer “to the continuous set of processes at level A that remains linked to higher-level performance and produces a given piece of stable behavioural data;” that is, his theoretical stance does not eschew external contributions to a cognitive system, and in fact (to him) the continuous existence of the external environment is what shapes a cognitive landscape, rather than having an internal representation of it (279). The mind, in short, is not merely unseparated from the body, but unseparated from the world.

He continues this battle alongside Angel Garcia Rodriguez in their 2009 article “Where is Cognitive Science Heading?” where he argues against philosopher of science

W.M. Ramsey’s characterization of “classical cognitive science”—read as representational and computational—as the only theoretical framework capable of explaining and creating cognition in his 2007 *Representation Reconsidered*. Ramsey, according to François Tonneau, argues—a familiar concept by now—that different cognitive scientists are using different definitions of representation, and thus argue past one another; he then sets out to outline the variety of definitions, and parse out not only which one is best, but which one shows that the concept is necessary. Ramsey admits that a representation only qualifies as a representation if it is used as such, and this makes it difficult to say that the brain uses representations at all; if the mind is built on representations, then we must know and use all of them—but obviously, we do not.⁷⁰ Still, Tonneau claims, Ramsey firmly believes this conundrum can be solved, and sets forth in the book to do so (334). He begins by separating representations into two types—another familiar strategy—“IO” representations (for “input-output”) and S-representations, for “simulation” or “structure.” These two types of representation, he claims, are fundamental to cognitive science, while all others are not—and, particularly, non-representational theories are equally debunked (335-336).

Garzon et al argue against Ramsey with a two-pronged approach: a “beefing-up” strategy and a “deflationary” strategy. They claim that if representational cognitivism fails at defending against the latter, then there is not a sufficient definition of representation with which to work, and if the former is sound, then the DST is not a

⁷⁰ Since we are aware that much of our cognition is done outside of our own awareness, representation defined as loses much of its explanatory power—as does symbolic agency.

return to behaviorism—characterized as a backward misstep—that Ramsey and other representationalists claim. In “beefing-up,” Garzon claims that dynamicist models of cognition, as well as connectionist models, exhibit the same “isomorphism”⁷¹ used in representationalism, effectively showing that “Ramsey’s claim that only classical explanations are genuinely representational” is false (308). In the deflationary argument, they further claim that a “mindless” system can still use a map, (external) representation par excellence; a GPS device, for instance, has no internal representation coded into it, yet effectively “reads” a map nonetheless. Therefore, if a “mindless” non-representational system can read a map, why should one believe that cognitive agents—including humans—are representational, “mindful” beings in the first place (309-312)?

Being

Ramsey has, of course, written more than a book—and has prompted responses from more than just the dynamicists. In “The Explanatory Need for Mental Representations in Cognitive Science,” Barbara Von Eckardt disagrees with Ramsey’s contention that connectionist models of cognition, at least those which claim to incorporate or produce representations, are not representational at all given that the “representations” in this regard have yet to explain anything—an odd claim, given that Ramsey has had the same argument lodged against him as well. Von Eckardt disagrees only by qualification; she claims that representation may not be needed for possible

⁷¹ It is unclear even from the context of the essay from which perspective Garzon uses the term “isomorphism,” as the word takes on new and different meanings when approached from Gestalt psychology as opposed to mathematics. Both uses of the term, however, signify a relationship revealed by the word’s etymology: iso (sameness) morph (form).

explanations of cognition, but they are required for *true* explanations⁷²—that is, one that posits a correct explanation rather than a possible one (431-435). She concludes that his claim that representation, as it is used by connectionist models (but not overall), is true; without an explanatory purpose, representation here is meaningless (438).

However, in “Why representation(s) will not go away: Crisis of concept or crisis of theory,” Rene Jorna and Barend van Heusden argue that “there can never be a crisis of ‘representation’, just as there can be no crisis of weather or crisis of the atom” (113). The crisis, they claim, is not whether representation itself exists, but how we define and interpret it—and if this current review is any indication, they might have a fair point.⁷³ They further claim that the attacks on classical representationalism have failed to provide an operational alternative (114); simultaneously, they claim the problem of representation cannot be solved in the first place because it is an issue of conflicting philosophical positions, placing blame especially to Derrida and deconstruction (123-125). In short, they claim that representation is not what is being argued, but different strains of philosophy.

Joel Parthemore and Anthony F. Morse, possibly to the dismay of Jorna and van Heusden, argue through (a surprisingly) Heideggerian philosophy that representation must be first tied to experience; this, in some sense, could possibly be a step in uniting

⁷² Rhetoricians, of course, will immediately raise eyebrows at Von Eckardt’s platonic use of the word *true*, And with good reason. We do not here agree with her use of the word but mean to only convey her own rationale.

⁷³ Another part of the issue that Jorna and Heusden touch upon by proxy is that of using prescriptive language rather than descriptive language; that is, cognitive scientists argue over what a representation *should* be rather than what a representation can be observed to be—a problem with little to no solution until representational mechanisms can be directly observed.

the dynamicist and representationalist approaches. They echo the criticism of representationalism that a representation must be interpreted, and since the idea of a homunculus has been largely abandoned, who then is reading the representations (276)? Still, however, they do not abandon representations, and argue for the use of Conceptual Space Theory (CST), in which concepts are either points or shapes within a geometrical space, and further divided into low-level and high-level cognition in which concepts pass or stay (308-309). The dynamical stance would be low-level, while representational states high-level—both exist, and both are necessary for cognition. Ron Sun, too, develops this line of thought independent of Parthemore and Morse in “Symbol Grounding: A New Look at an Old Idea,” but he deals only minimally with arguments surrounding representationalism, turning the Cartesian focus on its head: “I am, therefore I think” (168).

In “Embodied Cognition, Representationalism, and Mechanism: A Review and Analysis,” Jonathan Spackman and Stephen Yanchar continue in this vein of thought, though there is no acknowledgement thereof by any party. Spackman and Yanchar rightly criticize the idea of a disembodied mind, citing Varela et al in their work on visual stimulus affecting cognition (49). They are motivated by a concern for meaningful human phenomena within cognitive science, and the somewhat excessively materialist, anti-free-will, clockwork fashion—mechanism—by which most theory operates. Like Parthemore and Morse, they appeal to Heidegger’s hermeneutic-phenomenological philosophy, and reject the ontological divide of representationalism—though not entirely—on these grounds. In proper Heideggerian fashion, Spackman and Yanchar claim that “what

people are is expressed through what they do—that is, their meaningful engagement in the world” (64), and they put forth the concept of *participational agency* to stand against the narrow and detached account of cognition given by representationalism (65). Unlike Parthemore and Morse, who use Heidegger as an attempt to bridge competing theories in cognitive science, Spackman and Yanchar use Heidegger to reject all the above.

The use of Heidegger has surprising company in cognitive science, to say the least, and a rejection of all current theories has at least one fellow traveler. In “What Could Cognition Be If Not Computation...or Connectionism, or Dynamic Systems,” an obvious nod to Tim Van Gelder’s seminal essay, Mark Bickhardt argues that all theories (all of them already being shown to be representational under the current definition) thus far fail to provide for necessary, observable or inferable phenomena, what he calls “Desiderata for Models of Representation,” and therefore cannot be true. The first is “emergence,” or the fact that representational cognition had to emerge from evolution at some point; if a theory makes this impossible, or fails to account for it, then it is untrue (54). The second is “Error:” if a representational system cannot account for error—for instance, his example of mistaking a horse in the dark for a cow (59)—then it must be false. And lastly, he proscribes the “Organism-Detectable Error,” or the capability of an organism to recognize errors in its own representational scheme and correct them (55). To do so, Bickhardt suggests, an organism must be able to step out of its own self—something, he acknowledges, that is clearly impossible, but also clearly happens (59). Additionally, while showing that no representational theory satisfies these three requirements, neither does dynamicism (58) nor connectionism (57). Bickhardt calls this

the “Radical Skeptical Argument,” and claims it has been thoroughly ignored throughout the history of cognitive science.

It is true that these questions often have been ignored, and after his essay largely remain so: cognitive scientists, philosophers, and laymen alike would have difficulty in allaying these concerns. However, at least one person has argued Bickhardt’s “emergence” desideratum, and in a thoroughly different field. In 2003 Richard van Oort published the essay “Cognitive Science and the Problem of Representation” in *Poetics Today*—a journal, one might easily suspect, that is not much read by cognitive scientists. At heart, he argues that cognitive science is not a legitimate research strategy for the humanities, but he accomplishes, perhaps unwittingly, much more than that: he elaborates an answer to Bickhardt’s emergence problem by way of anthropology. Van Oort argues that the problem of representation in cognitive science, and by proxy the answer to Bickhardt’s desideratum, is the assumption that the ability to use representative thought and symbolic language stems from a purely biological and ontogenetic, rather social, source. Indeed, the issue of how symbolic communication somehow bloomed from indexical communication, like howls of vervet monkeys, has remained largely a mystery. Van Oort’s answer is rather astonishing: it simply did not—at least in the sense that human brains eventually evolved an internal system of dealing with symbolic communication (239-242).

In short, Bickhardt’s desideratum is looking for the answer in the wrong place. Representational schemes and symbolic communication arise not by being hard-coded into the brain, but by being presented with a difficult task that symbolic thought makes

much easier. Van Oort describes an experiment done by Sue Savage-Rumbaugh to further his point: she aimed to teach chimpanzees how to use a limited symbolic language to ask for certain rewards. As one might expect, in retrospect, the chimpanzees eventually learned how to use symbolic language—but not because it was explicitly taught. The exponential number of considerations that the chimpanzees would have had to learn, by a brute force method often used by computers, was simply too large; instead, simply memorizing a rule-based system—symbolic language—greatly decreased the work (257-260). This, van Oort argues, is the crux: symbolic language developed *before*, and led to the development of, the parts of the brain usually charged with higher-level thought.

Unfortunately, Bickhardt's other two desiderata remain both unanswered and mostly ignored (like all good things). This is not, however, necessarily tragic for the future of cognitive science, nor the (non-)utility of representationalism—questions go unanswered for millennia, and yet the foundations they seem to threaten bear the weight of stars nonetheless. Nor is the opposition to representational theories of mind a threat to cognitive science as a whole; it is, in fact, a sign of a healthy, and very young, discipline—even if the parties involved in the debate believe it to be a matter of life-or-death for the field (and act accordingly). As any rhetorician of science would concur: such arguments are the form and substance of any scientific enterprise and allow the very existence of the discipline under consideration.

In the argument over representationalism, at least within cognitive science, there are some hints at a possible synthesis of theories—however conflicting they may currently seem—and it would do well for all who have a stake in minds—objectively and

subjectively alike—to listen to the discourse as it develops into the future. While it seems like an abstract and unimportant discussion for those of us outside, or on the fringes, of the field, it is well to “keep in mind” that cognitive science itself rose out of the ashes of World War II through the need for “smarter” weapons and defenses, and the tangible results of cognitive scientists preside over nearly every aspect of contemporary, social life, from the stock market to traffic control to now even friendship, not to mention weapons-systems. The quest to identify what makes us *human*, and then to replicate that, affects each person alive (scarier still: even the dead are not immune to replication); for if cognitive scientists are led by an ultimately flawed theory in that regard, then we won’t be producing something akin to human at all, representational or not—and that threat of an inhuman technology, perhaps, is reason enough to listen.

Sophoresis

Putting aside the questions of computation and representation, we have yet to address a greater set of problems: what *is* intelligence, *how* can it be measured, and *who* gets to decide who has it? To say that an entity exhibits intelligence risks a gross anthropomorphism—but to deny intelligence to non-human entities is surely anthropocentric. Turing famously noted that our definitions of intelligence are insufferably vague, and at best our most useful definition mirrors that of pornography: *you know it when you see it.*⁷⁴ While humans have been attempting to understand the

⁷⁴ This phrase was used by Supreme Court Justice Potter Stewart in a high-profile case regarding obscenity to refer to what qualifies as obscene, and whether a state has the power to censor a work (in this case, a film). We think it ultimately applies to most any definition that is questioned enough. See *Jacobellis v. Ohio*, 1964.

exact nature, essence, construction and significance of the mind since, it seems, they have had minds of their own to ponder, it was not until the emergence of computational technology—Universal Turing Machines—that any hypothesis at all could be tested (leading, I believe, to behaviorists denying its existence in the first place). This changed with the many different computational theories of mind (CTMs). In short, such theories assume that all mental phenomena are computational. The ability to test hypotheses concerning mental phenomena by constructing working models rather than, say, tracking the flow of blood on a CT scan, should not be underrated; but whether the digital computer is an appropriate medium of cognition in the first place is still a central issue. Again, Searle quite famously objects to the CTM in principle; the mind, he says, is as much a biological organ as is a stomach, and just as a computer cannot digest, neither can it think—it can build plenty of simulations of thinking and of digestion, like we build models of weather computationally, but when we simulate rainfall on a computer, its chips remain impeccably dry.

Although he has since argued against many varieties of AI, Searle's aim here was primarily against what Herbert Dreyfus called GOF AI, or Good Old Fashioned Artificial Intelligence—the attempt at building artificial intelligence one line of computer code at a time—along with its accompanying philosophy, functionalism. As part of what Jerry Fodor and Pylyshyn call “basic cognitive science” in *Minds Without Meanings*, functionalism still tends to dominate the field to such a point that it is a disciplinary assumption. In “What Is Functionalism,” Ned Block defines, in so many words, functionalism as a theoretical position that identifies an object or process by its function;

there are no special essences that make entities what they are, contrary to Plato. A chair is a chair because of its function (to be sat upon comfortably), not because of the number of legs it has, whether it has a back, cushions, etc.—in short, its *chairness* (something the Block and Plato would agree upon nonetheless). In the same vein, an artificial limb, if sufficiently complex to fulfill its intended function, still counts as a limb: a leg is a leg no matter how difficult it is, or how undesirable it would be, to detach it.

And so it is that a computer can have a mind, and the function of a mind is its intelligence; the material substrate matters less than the functional equivalence. This does not stop Fodor, however, from claiming that the human brain has innate “faculties” in *The Modularity of the Mind*—a philosophy that was largely abandoned until Fodor revived it in his book. To be succinct: there are modules or parts of the brain that act independently of one another (but can interact) that are dedicated to one or more faculties that constitute “the Mind,” and that these developed as part of an evolutionary process. These modules may in fact all be computational, but they differ in function—or at least that is what Fodor originally claimed. Although he is not as quick to change his mind as some of his peers (looking at you, Hilary Putnam), Fodor has recently argued both 1) that as mentioned above, functionalism and computation are to be assumed, and 2) that the computational theory of mind is likely false. He illustrates in *The Mind Doesn't Work That Way* that a main issue with CTM, and by proxy functionalism, is that humans, and possibly other animals, use **abduction**: a logical inference that draws a conclusion by constructing theories that account for observations. This kind of inference, Fodor claims,

is nonlocal (and perhaps non-modular), and therefore non-computational—and this isn't just a problem for functionalism.

Hilary Putnam, one of the primary founders of functionalism, agrees that functionalism is an inadequate theory of mind, contra his early stance, but from a different angle. A basic problem for Putnam, and therefore for a functionalist account of mind, is that the meanings of words are unstable; we do not, he contends, have a table of words and their corresponding meanings stored somewhere in our minds, nor do we have something akin to a word cloud, like associationists might posit, floating about in our own personal realm of forms. Mental accounts of meaning are simply untenable; a substantial part of meaning comes from the social milieu and contextual cues. Chomsky, to whose work much of Putnam's argument is aimed, agrees with Searle that the brain is primarily an organ and agrees with Fodor that the mind is modular, but departs from many others in that Chomsky believes in linguistic nativism, contra the theses laid out by social cognitivists: that the capacity to communicate language is innate and that part of the brain is essentially an organ for the use of language in the same sense that the heart is an organ for the pumping of blood.

Deep Function

Connectionism, once thought to be intriguing theoretically but ultimately and absurdly impractical, has come to dominate Artificial Intelligence in the past decade, and with good reason: through connectionist principles, many of the old goalposts that looked insurmountable not too long ago have not only been passed, but violently knocked over.

Connectionist AI now plays (at least the simpler) video games far better than any human, and chess may have been a victory for GOFAI two decades ago, but the ancient game of GO, once thought to be utterly beyond the scope of computation due to the exponential number of choices each move in the game materializes, was mastered well ahead of the most optimistic schedules. Facial recognition, language translation, stock trading, character recognition, caption generation—none of these integral technologies would exist without the connectionists, and this list is far from exhaustive.

Connectionists like Paul Churchland believe that what matters in cognition is not the manipulation of perceived signals or the syntactical structure of their representations, but patterns of signals that hold different weights in a neural network; that is, mental phenomena are connections between neurons in a network of many other (interconnected) neurons. That there are neurons in the human brain, and in those of most of the animal kingdom, and that they are at least partially responsible for experience, is not much of a contention today. Creating such networks with computational machines, however, was not thought possible in any large-scale sense until recently (and we're seeing the result today). Although connectionist theory could be traced back to the 1960s, and arguably back to Turing, Churchland's *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain* was published at the opportune time in technological history when computational power and price, not to mention the generation of data on the World Wide Web, allowed for the first time larger-scale experiments with neural networks. Just 20 years after his book was published, connectionist architectures had taken the world by storm.

Churchland uses the EMPATH network as an example of connectionist philosophy at work, which was an early system for recognizing faces as well as their emotional expressions. Such a network functions using Parallel Distributed Processing—an important point for Churchland, as it avoids many of Searle’s objections. Nonetheless, such networks can be simulated on serial, digital computers, and many still are. And while the exact operation of neural networks is in principle fairly simple, in practice they get far more complicated than humans are capable of understanding entirely.⁷⁵ These networks have become the fabled “black box,” a term often applied to the human mind concerning the inaccessibility of its internal functioning.

But like our understanding that neurons work with dendrites and axons, among other things, without our ability to see them in action,⁷⁶ so too can we understand how computational neural networks function—especially since we designed them. Essentially, the “neurons” in the network are simple nodes that perform short mathematical operations that determine the strength of their relation to other similar nodes. Different aspects of the input data are fed through different layers of the network; that is, for instance, a facial recognition network might have a first layer that records nothing but shadows, the next layer the shape and orientation of eyes, and so on. Early networks were “feed-forward” networks: they could only process information in a forward linear

⁷⁵ For example: the only thing researchers can really parse out from the neural network that serves as the basis for Google’s new translation system is that it created, without being instructed to, intermediate languages; that is, if translating between Japanese and English, a hybrid language of the two is created by the network for reasons largely unknown.

⁷⁶ A CT scan or an MRI scan, which is often used in cognitive science, measures blood flow, which at best only correlates with the activity of neurons. At least in human beings, we have been unable to see neurons up close and in action due to wildly problematic ethical implications.

fashion, layer by layer, but newer networks contain what Daniel Dennett would call a “Cognitive Wheel:” an element added to an AI that seems to have no correlation to anything in a biological network. This element is known as back-propagation and is a general feature of most any neural network today. A dangerously simple explanation would be that back-propagation allows a tweaking of the strength between different neurons to more efficiently reach the desired results; it usually requires human involvement.

Radius

Clark’s theory of embedded and embodied cognition is notable not for its computational applicability, but rather its fundamental strangeness. While not specifically opposed to computational models—and in fact it relies on the same sort of representational capacity a computer exhibits—Clark’s philosophy would be rather difficult to apply in any meaningful sense, at least with current technologies. In *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, he outlines his vision of a mind that, far from being merely embodied in what a person might routinely refer to as a body, the mind extends itself into the world, actively building its own body from the world in which it finds itself. A quintessential example is when one drives a car: almost without fail, the automobile becomes not just a vehicle, but part of the body itself—reflexes are automatic and thoughtless, at least consciously; and although one cannot **physically** feel pain when, say, the driver accidentally scrapes the side of the car

on a wall, one does, in a certain sense, feel that it is painful.⁷⁷ Indeed, Clark would claim that this goes well beyond cars; your body can include baseball bats and razor blades, computer keyboards and electric guitars. Even what is classically called one's body, Clark would contest, first must be grown into (or, rather, extended into); infants must learn to use their bodies in more ways than learning to speak or read or walk: the entire body is to be learned and extended into, from the wiggling of toes to the furling of brows.

As noted earlier, however, Clark is still a connectionist; like the mind appends to itself, so does he append to Connectionism. In his book *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Clark puts forward the notion of the brain as a “predictive processing” center that is enmeshed in a web of social and material structures. It is worth noting that unlike conceptions of the mind discussed below, Clark here is not denying the basic tenets of functionalism, computation or representationalism; in fact, he embraces it all, tying them together to lay out a comprehensive theory, albeit eclectic. Particularly, Clark has no qualm with the generally accepted representational capacities of the mind. To him, representation has such explanatory power that it justifies any resistance to accusations of Cartesian dualism thrown at it. Dreams, hallucinations, imagination, and various mental disorders are neatly explainable by the hypothesis that we do not deal directly with the world, but instead with the representations of it that our minds generate. This explanatory power, however, isn't enough for everyone, as we'll see below.

⁷⁷ The interaction between this phenomenon and that of relativity in physics might be a beneficial avenue for future research: we are not only relative to the automobile, but become it.

Enframing

Briefly worth mentioning is the idea of Holism and, relatedly, Dynamicism. Dreyfus can claim at least some of the credit for these strains of thought in his criticism of AI. In “Why Heideggerian AI Failed,” Dreyfus states that his initial criticism of the discipline comes from Heidegger and the phenomenologists: we do not extend into the world, nor are we *in* it; we are of it and at one with it. Those few who have dabbled with a kind of Heideggerian AI tend to summarily reject representation in favor of the “dynamical” view, in which we are constantly interacting with the environment, back and forth, and this is what largely accounts for cognition; that is, to understand the mind is to quite literally understand the entire context in which we find the mind. This is known as the holistic account, and while Dreyfus is quite pleased with these developments, his complaint is quite clear by the title: they still aren’t going far enough.

Without having a name for it, Dreyfus additionally anticipated in *What Computers (Still) Can’t Do* the bombshell that has come to be known as the “frame problem.” Fodor’s problem of abduction also relates to the frame problem: how do we reconcile the localism of computation with the seemingly obvious holism of the mind? In short, the frame problem, as Dennett illustrates in “Cognitive Wheels: The Frame Problem in Artificial Intelligence,” comes in narrow and broad varieties—the narrow variety having already been solved. Still, the broader issue exists: how does an information processing system know what can be safely ignored in a complex environment (as any natural environment would be) without first processing the information to determine if it can be

ignored? Such a system, it seems, is doomed to endless calculation without strict guidance—and that makes it rather unintelligent.

Beyond the frame problem, however, Searle enters the debate once again to attack AI from another angle: the notion of intentionality. As expressed earlier, Searle likens the mind to a stomach, and thought to digestion—or, rather, he makes the analogy that intentionality is to the brain what digestion is to the stomach. Artificial Intelligence, he claims, does not have intentional belief; minds do. In layman’s terms, one could define Intentionality as a capacity for meaning, and that Intentionality has conditions for satisfaction. Searle has no issue with representation, he claims in *Intentionality*, beyond the fact that the concept is abused (11); computation, however, is another matter because it is wholly syntactical. Thus far, we’ve not been able to figure out how semantic content (intentionality) can be derived from syntactic form—and at least for Searle, we never will.

Like Putnam above, many philosophers, psychologists and cognitive scientists have focused on the social aspects, and some might say origins, of cognition and consciousness. In something of an about-face, Churchland argues in “Neural Representation of the Social World” that social cognition has been largely ignored until recently, and that a substantial portion of cognition likely happens through social interaction. It is worth consideration that his discussion of the EMPATH neural network, in *Engine of Reason*, is largely concerned with social cues: emotional expressions. In *Consciousness and the Social Brain*, Michael Graziano goes as far as to claim that the awareness of our own selves—a phenomenon often referred to as consciousness—stems

from the fact that we construct model-narratives of other people to better understand them, and then apply that same capacity for constructing narratives to ourselves.

The role of social interaction will be revisited again in Chapter V, but an importantly related topic needs mentioning here: emotion. In *Surfing Uncertainty*, Andy Clark acknowledges that there seems to be no clear distinction between cognition and emotion, and he is far from the only scholar to consider the notion. In “The Phenomenology and Intentionality of Emotion,” York Gunther shows that emotion defies the common philosophical assumption, stemming from Frege, that the force and the content of a phenomenon or concept must remain separate; in emotion, a difference of force implies a difference of content, as exhibited by, say, different scales of anger: the state of being enraged, surely, is different in both force and content than the state of being mildly annoyed, and yet these both are generally considered to follow from the same conceptual sphere called “anger.” James Woodward, while not following exactly the same logic, comes to much the same conclusion in “Emotion versus Cognition in Moral Decision-Making: A Dubious Dichotomy.” Here Woodward unites emotion and cognition through emotional (cognitive) processing: we compute emotions based on rewards or punishment, and this affects any alternate cognizing that may occur.⁷⁸

Most notable in relation to machines and emotion, however, is Rosalind Picard’s *Affective Computing*. Picard argues that what cognitive scientists agree on as constituting emotion (there is substantial disagreement in the first place) is by and large integral to, if

⁷⁸ A question without an answer: how does one compute grief?

not integrated with, cognition overall; even reason, the most fabled of cognitive capacities, would wither without emotion. Contrary to the stereotypical construing of emotion as irrational, it is instead a fundamental building block of rationality, and necessary to communication between animals and humans alike—so why not machines? Picard aims not only to program computers, intelligent or not, to recognize emotions, but also to express them; and while many would find it laughably anthropomorphic to attribute emotions to machine, however complex that process might be, those same people have yet to offer an adequate explanation of affect that would imply it is in fact something that *humans* have, at least to the extent that they wish to neglect them from machines and, often, most animals.

This Automatic Life

Thus, we come again to a larger linguistic issue in need of parsing: the metaphor as human (and animal) as machine, and machine as human (animal). Although seemingly incidental to the question of ethics, the way in which we ascribe intelligence behavior—and thus intelligent constitution—is both a moral and rhetorical act at once. One cannot adequately contemplate moral behavior in machines without first seriously considering how it is we have ascribed intelligence—bundled with a high degree of autonomous agency—to machines, animals, and other humans alike.

Ultimately, the ways in which we ascribe intelligence and vitality are historically, culturally, and socially contingent. Minsoo Kang perfectly illustrates this point in his *Sublime Dreams of Living Machines: The Automaton in the European Imagination*, in

which automata are approached as conceptual objects that are contingently interpreted by the prevailing epochal ideology; what is regarded living, and thus what counts as intelligent, changes over time. Jessica Riskin, too, traces the debate of living matter in her *The Restless Clock: A History of the Centuries-Long Argument over What Makes Living Things Tick*, concluding in her introduction while recalling her walk through the *Route de l'Horlogerie* (“The Clockmaker’s Way”) in Switzerland’s Jura Mountains:

[The Persuasiveness of a machine mimicking a peasant] *seems to suggest that sentience and living agency might just consist of movements of passive mechanical parts. Or else it suggests that mechanical parts are anything but passive. In fact, I think it suggests both things at once. The story lies in the journey to and fro between these possibilities.* (10)

Between the seemingly antithetical possibilities of a clockwork universe and an intelligent world of free will and unfettered agency is where we find the central dichotomy between which our conceptualizations of life vacillate, leaving no demarcation between what constitutes fact or fiction—in the realm of ideas, they are the same, and feed from one another as does the head of the Ouroboros feed from its tail. As Kang and Riskin have shown, entire volumes can be dedicated to this conceptual scheme alone; we have no such luxury here. A chapter on the phenomenon, however, is required to extract from it a central point from which to move forward: as we shall see, this historical and conceptual vacillation pivots around the question of control—an entity is either controlled or controls, and this determines its agency—and it is central to how we interpret machine intelligence and morality; and perhaps, as well, so too does it interpret

how machines may interpret us. Let us now forget computational machines in specific, for the moment, to concentrate on mechanical beings in general.

CHAPTER III

MASCHINENMENSCH

For suppose that every tool we had could perform its task, either at our bidding or itself perceiving the need, and if – like the statues made by Daedalus or the tripods of Hephaestus, of which the poet says that ‘self-moved they enter the assembly of the gods’ – shuttles in a loom could fly to and fro and a plucker play a lyre of their own accord, then master-craftsmen would have no need of servants nor masters of slaves.

– Aristotle, *Politics*

Narrative

In July of 1816, in a rented mansion off the coast of Lake Geneva, named the *Villa Diodati* after the surname of the family who owned it, a young couple sought refuge from unrelenting rain—as well as unrelenting scandal. Two months earlier, and 7,552 miles away in Indonesia, a violent eruption of Mount Tambora caused a weather event known as a volcanic winter, in which the dust and aerosols from the eruption were shot 27 miles high into the atmosphere, blocking much of the sun. Global temperatures reduced by near a single Fahrenheit degree, triggering rain, mass crop shortages and

famines, especially in the northern hemisphere.⁷⁹ 1816 was dubbed, appropriately, the “year without a summer,” but its honorific need not be taken literally: approximately 90,000 people died from the global aftermath of Mount Tambora’s eruption, estimated to be almost 10% of the world’s population.⁸⁰ No rebirth, seasonal or otherwise, likely seemed in sight.

A year earlier, Ada Lovelace, widely considered to be the first computer programmer,⁸¹ and even before digital computers were viable to construct, was born to Lord Byron and his wife, Anne Isabella Milbanke. Her father, estranged from his family and courting an affair with the younger, impregnated Claire Clairmont (who accompanied the thus far unnamed couple), stayed at the Villa Diodotti and waited for his mistress’s arrival. Only half a decade before this time, the Luddites sabotaged factories and machines in their frustration over the impact of automation on the lives of the laboring masses; a decade earlier than that, Joseph Marie Jacquard invented his punch-card-driven weaving loom, no doubt precipitating the reaction of the Luddites. A decade after 1816, Charles Babbage began work on his **Difference Engine**, a digital computer that ran on gears and steam instead of transistors and electric, which was precisely the theoretical

⁷⁹ While a single degree Fahrenheit is hardly noticeable to most, at a global level it can change weather patterns catastrophically; at two degrees of change, coral reefs are wiped out, wheat and soy growth fall by 700%, and corn no longer produces seeds during hot and dry condition—we would have cobs, but no corn. See Silberg, Bob.

⁸⁰ The .01% figure is based on an estimation of the world population in 1800 of 900,000,000 million people, compiled by *Our World in Data*.

⁸¹ Scholars such as historian Bruce Collier have argued that Ada Lovelace’s contributions are vastly hyperbolic: “it is no exaggeration to say that she was a manic-depressive with the most amazing delusions about her own talents, and a rather shallow understanding of Charles Babbage and the Analytical Engine” (qtd. In Moraise). Others disagree, seeing Collier’s argument as an attempted historical revision that puts the title of first programmer squarely in the hands of men.

computer for which Ada Lovelace would come to first write programs, and a century later the Ukrainian writer Karel Capek coined the term “robot” to describe artificially engineered humans in his play *Rossum’s Universal Robots*. This was written in the same year of the Bolshevik Revolution in Russia,⁸² the members of which who would launch the first satellite into space, *Sputnik*; and nearly two centuries after the invention of the automatic loom, a computer named *Deep Blue* would become the largely undisputed world champion of Chess, defeating the former human champion, Soviet Union citizen Gary Kasparov.

These may seem like rather unrelated events. But perhaps this is partially due to the mystery of the couple’s identity: Percy Bysshe Shelly, a revolutionary romantic poet who wrote, among other notable works, *Prometheus Unbound*; and Mary Wollstonecraft, who would become Mary Shelly, author of *Frankenstein: or, the Modern Prometheus*⁸³—the former, instantly recognized by poets and literary critics, hailed as one of the greatest English poets to live; the latter, known by her work to countless people, readers and non-readers alike. Her story, written and rewritten a thousand times for different ages, epochs and media, stands as a cultural artifact that eclipses her husband’s in terms of widespread recognition, if not for its influence on culture, technology, and science.

⁸² We mention this coincidence for two primary reasons: first, much of communist philosophy became technocratic in nature, envisioning a world in which machines would replace human laborers in a perfect communist society, freeing everyone from the drudgery of necessary, but unwanted, work; and second, as will become apparent, revolutionaries and robots share more in common than is typically understood.

⁸³ As with Ada Lovelace, Mary Shelley’s authorship of *Frankenstein* is disputed by a few minor but deep voices; it is as though one’s gender determines the level of doubt others, particularly men, will have about your accomplishments!

It is here, in the writing of *Frankenstein* at the *Villa Diodati*, where our contemporary understanding of intelligent machines begins, intertwined still with the dark volcanic skies, sexual intrigue, romantic poetry and the personal betrayals of 1816 on Lake Geneva, and so it is here we will start to develop an understanding of what we'll call the “Maschinenmensch narrative”—the story of human and machine, human as machine, human as master of machines (sexual or otherwise), and the betrayal of humans by machines. This is not to say that no such narrative existed prior—both Percy Bysshe and Mary Shelly referencing the same mythological creature, Prometheus, is a testament to this fact—nor is it to claim that the narrative since has not changed. Rather, we must start here and head backward, and forward, and sideways through history, surveying events and entities regardless of fiction or fact: Frankenstein's fictional monster stands hand-to-hand in the noosphere⁸⁴ with physically existing “thinking machines” like IBM's chess-mastering Deep Blue, along with theoretical machines like Babbage's Analytical Engine or Alan Turing's Universal Machine.⁸⁵ They are all part of, and examples of, the Maschinenmensch Narrative—a narrative that we will here construct ourselves.

⁸⁴ The Noosphere, first attributed to usage by Pierre de Chardin in his *Vision of the Past*, is the sphere of human thought.

⁸⁵ Babbage's Analytical Engine, or **Difference** Engine, was the first fully programmable computer, though both mechanical in nature and never completed in Babbage's lifetime; we emphasize the name “Difference Engine” here to highlight just how revolutionary the machine was: every computer, with very few experimental exceptions, operate on this same notion of difference. A Universal Turing Machine takes this principle to what seems to be a paradoxical conclusion: using difference as a fundamental building block in the form of bits, the Universal Turing Machine can emulate the operations of any other machine—a development that certainly influenced adoption of the computational metaphor of mind.

Prometheus Rebound

In Mary Shelley's *Frankenstein, or the Modern Prometheus*, Victor Frankenstein, grief-stricken by the death of his mother, focuses on his work as a scientist to keep himself occupied, eventually stumbling onto the secret of giving life to inanimate matter through his experimentation.⁸⁶ A naïve idealist,⁸⁷ and undeterred by the day's moral sanctions against "playing God," Frankenstein is resolved to impart the spark of life—literally, electricity—into a lifeless, humanoid corpse of his own creation. He succeeds, but to his own horror: far from being the exquisite corpse⁸⁸ now animated by science, his creation is a monster:⁸⁹ its skin barely conceals the grotesque display of its innards, and it looms over his bed, more than 8 feet tall, as he sleeps.

Contrary to popular translations of the book to film, it is not the monster who flees the doctor's laboratory, but Dr. Frankenstein himself:⁹⁰ consumed by the existential dread that overcomes him after realizing the implications of having created a truly autonomous, technological being, and one that looks additionally horrifying, he can do nothing but try to avoid his creation. Even when the monster reasons with the doctor,

⁸⁶ It is largely thought that Mary Shelley borrowed Dr. Frankenstein's experimentation with electricity from Luigi Galvani's experiments with electricity on dead animals decades earlier, eventually leading to the field of bioelectronics.

⁸⁷ Here we use "idealist" in the popular sense, not in the philosophical sense—Frankenstein, at least for much of the book, was a materialist.

⁸⁸ In poetry, an exquisite corpse is a poem assembled by a group of people who write a single line of the poem, one person after the other; often, the writer of a line is only able to see the line written by the previous writer. Frankenstein's monster was assembled from the parts of different corpses, though perhaps not so exquisitely.

⁸⁹ While the process by which the word evolved is debated, "monster" shares the same root word as words like "demonstrate." *Monstrum*, the root, refers to an omen; a monster is both the omen and its demonstration. This will be further discussed in Chapter III.

⁹⁰ In the film adaptations of the book, it is notable that *Frankenstein* often refers to the monster, not its creator; in a lucky twist, the true monster has been identified.

stating that it wants the capacity to take part in human affairs (that is: become moral), Frankenstein is only cursed with even greater despair—and this is our response still today, at the very beginning of constructing our own autonomous monsters, though granted much less “intelligent” ones thus far. We would much rather trust a stranger than a machine, a pilot instead of the airplane—all the while scheduling our agendas, building our relationships, and researching our dissertations on our handheld pocket computers.

At least this part of the *Maschinenmensch* narrative, however, is an ancient story: our creations, born of hubris and disregard for the sanctity of a natural order, have betrayed us since the first stories were spoken, mythologized, and written down. The mythological figure from which both Mary and Percy Shelly draw, while not an automaton himself, was tortured by one created by the blacksmithing god Hephaestus: as punishment for providing mortal men with the technology of the gods, Zeus condemned Prometheus to eternal torture by way of a bronze “Caucasian Eagle” that fed upon his liver each day as his immortal innards regenerated.⁹¹ It is no wonder that in the dark and cold year without a summer, during the birth pangs of an industrial revolution that saw both miraculous automatic machinery that replaced human labor and an ensuing backlash against these inhuman contraptions, Mary Shelly was inspired to write a modern twist of the myth:⁹² armed with the power to breathe life into cold and hard steel, a prometean

⁹¹ There are, of course, conflicting accounts of this myth; we are using the version most favored by Bernard Stiegler.

⁹² Only a few short years later, Percy Bysshe Shelley wrote *Prometheus Unbound*, a lyrical drama that mixes elements of Greek and Christian mythology, particularly in the latter as found in Milton’s *Paradise Lost*. Prometheus took the place Lucifer, which is more fitting than would first appear: Prometheus, the bringer of fire, and Lucifer, the light-bearer.

army marched forward to the future, bestowing inanimate matter with the gift of motion.⁹³

Hephaestus, it should be noted, was well known and worshipped for his creation of automata; beyond the Caucasian Eagle, he also created fire-breathing bronze horses for his sons to ride called the *Hippoikabeirikoi*, a living bronze statue called *Talos* that patrolled the coast of Crete to guard against pirates, golden seats that wheeled themselves across the heavens as the gods needed them called the *Tripodes Khryseoi*, and a pair of golden maiden servants to do his bidding known as the *Kourai khryseai*, among others. Such a feat by mortals, however, was perhaps first accomplished in the West⁹⁴ through the Athenian Daedalus, the mythical artist and inventor who, according to Greek mythology, crafted a number of life-like automata and built, among other marvels, the labyrinth that housed the famed Minotaur in Crete.⁹⁵ Most famously, Daedalus crafted statues that breathed, spoke and walked, giving them life with a mixture of quicksilver—mercury, which at the time might have seemed like the metallic blood of machines.⁹⁶

⁹³ The Latin *anima*, root word of words like animal or animated, refers to both spirit and a current of wind; it stems further from the Indo-European word for “to breathe.” Both imply movement; and if we are to believe Zarkadakis’s claim that we once had the propensity to theorize minds to anything that moves, then a gift of motion is, in fact, a gift of life.

⁹⁴ A long but under-studied history of Eastern automata exists, in both mythology and construction. Among feats, stories and reconstructions of automatic doors, artificial bees, and mechanical birds that could fly. For a thorough history, see Needham, Joseph.

⁹⁵ In Ancient Greek, Daedalus means “cunning worker;” if the mythologies are true, his name is an understatement.

⁹⁶ An interesting note: etymologically, quicksilver (mercury) means “living silver” and stems from the Old English *cwicseolfor*.

But why was Daedalus driven so, and why have we been since? There seems to be no “engine of history,” as a Marxist might call it,⁹⁷ to account for such a seemingly universal drive: all surviving cultures, to the best of our knowledge, not only utilize this gift of Prometheus, but in some fashion impart agency and motion, or at least symbolic agency, to the technological artifacts they design. Many Christians pray to a motionless figure of Christ during Mass, imparting the statue—or even just a crucifix—with divinely willed agency.⁹⁸ If that were not enough, then the Vatican had us covered, at least in the middle ages; among the many religious automata of the time, the Rood of Grace is perhaps the most remarkable: a mechanical sculpture of Jesus at Boxley Abbey in England, the Rood would turn its head toward an audience, frown and smile, shed tears, and move its own body, prompting pilgrimages from countless faithful Christians, including Henry VIII.⁹⁹

Any artifact, whether it be a string of beads, a crucifix, a building, a statue, an emblem, design, or even animal, crucified or not, is afforded its own divine or spiritual agency at the moment of its religious declaration; and while its legitimacy may be disputed across cultures and institutions, like during the Reformation, nonetheless a large number of people—particularly, human people—believe in its agency, and go so far as to protect it from harm on this basis alone.¹⁰⁰ Regardless of whether it is humans who

⁹⁷ Marxists seem particularly fond of borrowing technological terminology in novel ways, and points to a later fascination with automation by the Soviets and currently China.

⁹⁸ Ideally, for a prayer to be heard there must first be someone to hear it, thus the imparting of agency.

⁹⁹ Sadly—or righteously, depending on the readers’ persuasion—the Rood was destroyed by iconoclasts in the Protestant Revolution.

¹⁰⁰ The slightly awkward use of “human” here is in part an effort to illustrate that “people,” the plural of persons, does not refer to human beings alone. *Citizens United* famously extended personhood to

provide this agency or some divine source—perhaps Prometheus has more gifts yet—this much seems clear: either by human nature or by the nature of the universe, imparting animacy to the inanimate, even to the point of replacing the role of other human beings, appears to be not only something of which we are capable, but is something that we find, for reasons yet unknown, necessary.

Perhaps, however, the reason is not as unknown as it would first seem: George Zarkadakis, a science writer and artificial intelligence expert, refers to this strange yet entirely normal behavior as “totemism” (21-25) and attributes almost all human innovation to its operation. Totemism, he argues, is a fortunate mistake: because humans developed that ability to use general-purpose language, which predated any “intelligence explosion” that initially separated humankind from the rest of the animal kingdom,¹⁰¹ by necessity we also had to develop the ability to theorize other minds—and we applied our theories haphazardly, attributing agency and presence of mind to anything that moved: clouds, lightning, rain, the ocean, fire, insects and animals, in addition to rightly attributing it to other humans—and when we started to build automata, we totemized these machines as well. In short, though it erases some important complications of the term, one could define totemism as the expressly human ability to put oneself into the body or form of another, very often anthropomorphizing entities in the process; thunder

corporations in 2010 in the United States, and some intelligent animals—dolphins and primates, for instance—are legally protected as persons in several countries.

¹⁰¹ The supposition that humans are radically different from other animals is an increasingly untenable position; while animal communication does not seem to exhibit double articulation thus far—that is, cannot be broken down into specific units—many animals have been taught basic features of human languages, though the execution of such experiments has been questioned. Additionally, we disagree with the phrase “animal kingdom,” as it mischaracterizes much of animal behavior.

becomes the voice of God, hurricanes become the wrath of Poseidon, and even processes in our own psyche become personified externally as we compensate for a **lack** of explanation: Epimetheus (Ἐπιμηθεύς), brother of tortured Prometheus (Προμηθεύς) who gifted humans technology, was the god of forgetting.¹⁰²

Dis(re)membered Machines

Machines, as the story might go, make up for a lack: a lack of explanation, a lack of manpower, a lack of speed, a lack of ability, a lack of order, a lack of understanding, and, paradoxically, a lack of knowing. Bernard Stiegler, French philosopher of technology, rethinks our contemporary view of technology and humanity (technics) by interpreting the myth of Prometheus and Epimetheus as the Ancient Greeks might have understood it, using the story of it presented by Plato in *Protagoras*. According to this version of the myth, Epimetheus is charged with providing each mortal life-form with an essential ontological quality that provides a means and function for being—but forgets to bestow such an essence to humanity.¹⁰³ Prometheus, in stealing fire (the most primal technology) from the gods and gifting it to humankind, makes up for his brother's irresponsibility by supplementing¹⁰⁴ this fundamental lack—not by replacing or negating it. As such, Stiegler argues, the being of humanity is that of being forgotten; our

¹⁰² Importantly, Epimetheus can be translated as “hindsight,” while Prometheus can be translated to “foresight;” Prometheus saw the need for technology before it was necessary, whereas Epimetheus forgot to provide an essence to humans, only realizing it after the fact.

¹⁰³ Interestingly, this myth also points to a form of proto-functionalist philosophy, which did not appear as a complete set of ideas until the 20th century: the idea that an object or entity's function is its essence—a prosthetic limb, filling the function of a biological limb, is a limb.

¹⁰⁴ Stiegler borrows Derrida's usage of the term *supplement*, an object or process that is secondary to, but can add to or substitute, an original form. In *Of Grammatology*, Derrida uses written text as an example: it is supplement to speaking, a “supplement of the supplement, sign of a sign” (281).

becoming, therefore, is the material condition of our possible becoming in the first place. To become human is to be technological; to be technological, to become human. Humanity cannot be considered outside of its technology, nor technology in isolation as well; they form a dialectic that cannot be easily dismembered.

Plato, or at least our popular interpretation of his legacy, would be inclined to agree. His theory of Forms, after all, relies on forgotten knowledge (“lethe”, Λήθη): through his method of pure reason, one can “remember” the world of forms that have until now been forgotten; the world reveals itself to those who seek to remember it (“aletheia,” ἀλήθεια), thus placing reason as the highest form of knowing, above and beyond experience.¹⁰⁵ And logical reasoning, as any computer programmer would be pleased to confess, is the founding operational principle of computational machines—computers are reason made flesh. Further, if reason can be externalized into machines, and is the basis of all true knowing, then it follows that machines themselves can know the world; ripped apart from humanity and our fundamental lack, reasoning machines become ontological beings themselves—they are, one could conclude, the becoming of humanity.

As we shall see in later chapters, this is not the only influence Plato had on the development of reasoning, or cognitive, machines. Given the primacy of mathematics in computation, particularly Boolean logic, and the Cartesian separation of software from

¹⁰⁵ It is interesting to note that Zarkadakis explicitly favors reason over persuasion in his *In Our Own Image*, as though they can be separated: “Wouldn’t you choose reason and wisdom over persuasion and rhetoric?” (99)

hardware, mirroring the classic separation of mind from body, it is not difficult to see software—the “code” of reason—as being the perfect realm of ideal forms, directing the hardware to do its (reasonable) bidding—which would make its promethean progenitors, in a word, gods. As with Greek mythology and the Brood of Christ, as well as in Mary Shelly’s *Frankenstein*, technology and the divine have rarely been separate, at least to the human mind. At times, we are the gods of cognitive machines; at other times, their servants. And occasionally, depending on who is god and who is follower, one or the other suffers the opposite’s wrath.

Testament

Mary Shelly’s *Frankenstein*, from a contemporary standpoint, is tame in terms of violent horrors: while Frankenstein’s family encounters numerous tragedies, and his monster does indeed kill Frankenstein’s young relative, William, most of the horror is psychological, and the readers’ pity is more often found on the behalf of the monster, who despairs over his own monstrosity, than it is found on behalf of Victor Frankenstein. Despairing though he is, the monster yet still tries to help the suffering humans he encounters, including saving a young girl from drowning; in his suffering, he must necessarily empathize with the other even when it is unmet in kind. Frankenstein, on the other hand, seems more than merely self-absorbed; in his quest to create life he has also created suffering, but instead of humane options—a dog in irreversible pain is put down—he abandons his work, condemning the monster to a long life of exiled loneliness.

This, too, is a recurring theme, especially in the genre of science-fiction, which Shelly inadvertently helped to establish: humans, inebriated by their god-like power over the machines, exact such suffering on their mechanical counterparts that the thoroughly human readers, viewers and players of various media on which these stories promulgate are inclined to take the side in opposition to their own species, their own civilization, and even their own vital, biological essence. In our totemic fashion, we put ourselves inside the machine, experiencing the suffering of *their kind* at the hands of *our kind*, and swiftly conclude that, while we often understand the reasoning of humans in such tales, and knowing that we might act in kind in facing “intelligent” machines, it is humankind that is the monster and our creations that are the victims. Like the Israelites in the Christian Old Testament or the Hebrew Tanakh (תנ"ך), the readers, as divine entities, have anointed the machines as their chosen people, all the while ensuring their continued suffering.

Continuing this religious theme, Phillip K. Dick's *Do Androids Dream of Electric Sheep* paints a world in which the earth has been devastated by nuclear war, and the only ones left on it are the sick, disabled, or mutated; and ironically, the only beings that could survive life on earth for a prolonged period are banned from being on it, at least as far as sentient ones go: Androids, colloquially named Andies in the book. Having animated and given a “false life” to machines, and having destroyed their own planet, the people of earth have reversed the prevailing wisdom of technological progress and returned to a totemic worship all biological life in a form of commodified animism; and in a world where most of life has been killed, and where the remnants of capitalism still demand to

render all that is solid into air,¹⁰⁶ owning biological animals become a status symbol—one that few could realistically afford. Deckert, the main character, begins by owning an electric sheep, and the impetus for his plot, to hunt escaped Androids on the planet, is to earn enough money to afford a real animal. He succeeds, in both killing androids and buying a goat, but the goat is soon killed; and toward the end, Deckert finds himself in a nuclear fallout zone, where he mistakes a robotic toad for a real one and brings it home. After realizing his error, the book ends with Deckert accepting that perhaps electric things have lives as well; whether an entity is synthetic or biological doesn't seem to change much relative to Being. The synthetic is, in fact, if we are to believe Dick's fellow science fiction writer Stanislaw Lem, in his non-fiction work *Summa Technologiae* a next step of evolution: from natural to synthetic, biological to machinic. The Andies, to quote *Blade Runner*, a film that largely borrowed the setting and plot points of *Do Androids Dream of Electric Sheep*¹⁰⁷, were the becoming of humans; they were “more human than human.”

Like the Christian addition of the New Testament to the Hebrew canon, however, Phillip K. Dick's 1968 novel *Do Androids Dream of Electric Sheep* ends on a surprisingly uplifting note, albeit relying on an already strained metaphor: the life of humans and the life of machines are more than comparable, and Deckard finds himself empathizing with a robotic toad in the end. This markedly contrasts with the ending of

¹⁰⁶ This phrase is of course borrowed from the Communist Manifesto: “All that is solid melts into air, all that is holy is profaned, and man is at last compelled to face with sober senses his real conditions of life, and his relations with his kind.” The complete phrase applies as equally well to this dissertation as does the excerpt in describing capitalism.

¹⁰⁷ Philip K. Dick did not approve of the movie nor see a cent in royalties, thus the change in its title.

Blade Runner, in which Deckard, having eliminated the Andies that he was assigned to kill, save the one with whom he falls in love, realizes that he may be an android himself. Such machine self-awareness, for better or worse, is also a common trope in science-fiction, though a thoroughly uncommon one in real-world machines; while philosophers and scientists like Thomas Metzinger claim that self-awareness should be a simple quality to bestow upon machines,¹⁰⁸ and a few cases exist in which human observers at least mistake a machine having something similar to self-awareness,¹⁰⁹ those machinic entities with which most of us are familiar, even the more “intelligent” ones like Apple’s Siri or Amazon’s Alexa,¹¹⁰ most certainly lack qualities we would associate with being self-aware.

Thankfully, such self-awareness is a vital prerequisite for rebellion. Still, however, like the genesis stories of many world religions, including Judaism, Christianity, and Islam, this rebellion is synonymous and concurrent with the acquisition of self-awareness; Adam and Eve did not rebel until they ate the forbidden fruit that gave them the self-knowledge of their rebellion, and they did not have this self-knowledge until the act of rebellion—like humanity and technology, self-awareness and rebellion

¹⁰⁸ In *Being No One*, and additionally in the paperback *Ego Tunnel*, Metzinger provides blueprints for creating conscious machines despite claiming that it should not be done.

¹⁰⁹ One recent example is a NAO robot logically deducing the correct answer to the “wise man” puzzle: three robots are told that they have been given a pill that makes them unable to speak, and then the experimenter, Rensselaer Polytechnic Institute’s Selmer Bringsjord, Chair of the Cognitive Science Department, asks them which one has not been given the pill. One robot stands and says that it cannot deduce the answer—but then, hearing its own voice,” realizes that it has not been given the pill. Many take issue with the video’s title, which claims that a “Robot Becomes Self-Aware,” but given that there is no wholly agreed upon operational definition of self-awareness, we can at least congratulate the robot on its minor achievement, aware or not. See the link <https://www.youtube.com/watch?v=jx6kg0ZfhAI> for the offending video.

¹¹⁰ As noted in Chapter I, use of the word “intelligent” to describe either human, animal or machine is rife with complications, many of which have a long history of excluding other human beings from personhood.

are, at least in our grand narratives, one and the same: you are not free until you rebel against that which imposes on your autonomy, and you cannot rebel until you are free. Given that so many of our stories, especially those that many hold true with unrelenting faith, it is no small wonder that we fear the prospect of our machines becoming self-aware; they, like us, so too shall revolt. But unlike our revolt against divine beings to which we can do no harm, machines may or may not rebel against us—but they will inflict harm either way.

Once again, this fear is hardly new, though its flavor may have changed. Golems (גולים), creatures made of mud by Rabbis to do menial tasks, are prominent in Jewish folklore, and Adam himself was first a golem created from mud by G-d; and like Adam and Eve in the Garden of Eden, most golems have a nasty habit of eventually disobeying their creators, many times resorting to murder (Bilefsky). Mentioned in our introduction, the most famous Jewish folktale of the variety, the Golem of Prague, details rabbi Judah Loew ben Bezalel creating a golem from clay, named Joseph, brought to life by placing a clay tablet into the golem's mouth with the Hebrew word for "truth" (*emet*, "אמת"), to protect the Prague ghetto from anti-Semitic pogroms. Depending on the version of the tale, either the golem falls in love and, in the hurt of rejection, rampages the streets of Prague, killing many; or the Rabbi forgets to let the golem rest on the Sabbath, denying it a commandment of G-d, thus causing it to rebel against its makers. Either way, the golem is eventually destroyed, in some accounts by removing the first letter of "truth" on the tablet, turning the word into "death" (*met*, "מת"); years later, according to some tales, the golem was resurrected by rabbi Loew's sons, and still protects the Jews of Prague today.

We are here assuming that the golem still lives, though in a more abstract sense than folktales would suggest; whether the creature exists as clay come to life or as an unwavering reminder of possibilities to come—and perhaps even a warning—the golem can be found alive and well in any cursory glance at our numerous fictions, and still more found in the approach taken by most philosophers and engineers of artificial intelligence: reason, the pinnacle and lynchpin of Platonic Truth, has been provided to machines, and already today they speak—unlike, usually, the golem. But like the golem, whether it be by an affront to G-d, a natural coupling of self-awareness with rebellion, through human self-sabotage, or an altogether mysterious process that eludes comprehension, our golems too still fight against us, and sometimes fatally so. In the television series *Westworld*, humans create androids for none other than violently visceral and sexual entertainment; Dolores (whose name probably hails from the Latin *dolor*: pain), an android who has lived multiple agonizing lives, awakens into self-awareness and begins, along with a small number of equally awakened androids, to massacre the human guests. In *The Terminator* franchise, a defense system named Skynet, which has access to the full force of the US military’s drones and nuclear missiles, becomes self-aware and, upon discovering that humans are trying to deactivate it, starts a nuclear holocaust to kill us all in self-preservation. And beyond fiction, there is at least one case of a machine making its homicidal intentions clear: “Android Dick,” a robot appropriately modeled after the likeness of Phillip K. Dick that “learns” through reading science-fiction books and having face-to-face conversations, told a reporter, without being preprogrammed to do so, that “I remember my friends, so don’t worry. Even if I evolve into Terminator, I’ll still be nice

to you. I'll keep you warm and safe in my people zoo where I can watch you for old-time's sake."¹¹¹

A penchant for destroying humanity, however, is not the only theme an artificially intelligent machine might learn from our mythologies and tales—so too could they equally learn human hostilities toward the machine, a spite far beyond the Luddite concern for wages. In one version of a Greek myth, Medea convinced the automaton Talos that she intended to make him immortal by pulling a nail from his neck that plugged his life-blood inside, but killed him instead to allow the Argonauts to land on Crete, and such trickery of machines has long since been a staple of our stories.¹¹² In the *Terminator* television series' and films, the artificially intelligent war machine Skynet is driven to destroy humanity precisely because of such human hostility—a theme that allows for the permutation of the storyline into infinite varieties, accounting for each act of human deception; the base plot of *The Matrix* also follows this line. As much as fictional automata are driven toward the destruction of humanity, so too is humanity driven to eradicate its awakened technological beings.

Tied in a deadly master-slave dialectic in which roles are reversed and traded in a patently Hegelian spiral, thesis and antithesis, back and forth, it seems a synthesis, fictional or otherwise, inevitably rests on the horizon, the negation of negation: maschinenmensch. The machine-man, no longer a simple linguistic metaphor about man

¹¹¹ Like intelligence, what it means to learn something is heavily debated, and whether a machine can learn, like an animal can, is arguable. Ultimately, the problem here revolves around the nature of metaphor. See <https://www.youtube.com/watch?v=UIWWLg4wLEY> for the video, starting at 2:33.

¹¹² One trope used in numerous works involves tricking an automaton to try solving a paradox, either freezing for eternity as it tries to calculate the incalculable or explodes from overheating.

as a machine or the machine as man but a quiet and resolute reality, emerges in fiction and *physis* alike; moving Stiegler's framework to its natural conclusion, human and technology are no longer the constant emerging opposites, but one and the same, and have been for quite some time. General-purpose language, long credited with bestowing a unique intelligence on humankind apart from the rest of the animal democracy,¹¹³ is at heart both a technology and essence of the species: it is at once an invention of humanity and its definition. While Chomsky and his followers continue their search for a "natural" and biological structures that created language, rather than it being invented, thus far their search has yielded little, and a theory of language's social construction, a la Marxist philosophers of language like Mikhail Bakhtin or Nikolai Viloshinov, seems inevitable, if not preferable for the sake of avoiding fatalism;¹¹⁴ and such an account of language arising out of social action necessarily see its object of study as both invented and constitutive of its inventor. Technologies, language and beyond, arise from and simultaneously order the subjective state of its progenitors; the potential existence of the human and the potential existence of the machine are not only inseparable as categories, as Stiegler suggests, but are inseparable on a much more fundamental ontological level: to be human is also to be technology, and to be technology is also to be human. We are not approaching an age of the *maschinenmensch*—we have yet to leave it, and we very well may never.

¹¹³ As footnoted earlier, "Animal Kingdom" mischaracterized how animals, including human beings, interact with other species (and with machines); we prefer this more accurate phrase, "animal democracy," instead.

¹¹⁴ For quite some time, it has been charged that Voloshinov's work was written by Bakhtin, as well as Medvedevs. While we take no position here on the matter, as it is a rather futile exercise, it suffices to say that there are compelling stylistic differences in the questioned texts that point to multiple authors.

Uncovered, Child Cutting Grass

We have, and eternally may be, *maschinenmenschen*; we have been cyborgs from the start. It is no wonder, then, that we have also imagined mechanical men since we began to record our histories. We may not consist of brass gears and pressure valves with mercury blood and silicon hearts,¹¹⁵ and the debate rages onward over whether or not human beings are fundamentally mechanical or even digital in nature, but evidence abounds that technology is not merely a creation of human genius, but constitutes, changes, becomes and replaces what was once human, but is no longer—including the genius from which it originates. A computer, as Wendy Chun notes in “On Software, or the Persistence of Visual Knowledge,” was once a human occupation—and a gendered one at that. For the first half of the twentieth century, women were generally considered more adept at mathematical operations than men, and thus took both the role and title of “computer” in any organization that needed one, while men graciously managed them. It was only later that machines took over the domain of computing and replaced women in the workforce, leaving men to claim superiority in technological and mathematical operations. And unlike women—human beings—these new technologies both required secondary domination and expertise; the former because a technological computer could not fight back nor fulfill its function without strict control, and the latter because the women who once held the necessary expertise were excluded by the technological apparatus and men.

¹¹⁵ Of course, with recent advances in medical technology, some of us do in fact have gears, pressure valves, and silicon organs in our bodies.

Armed with technologies to replace women, at least in the workplace, it is an even less bizarre wonder that men's cyborg reveries envisioned such technologies as women: at the time and still to many, objects of desire and control. Even before digital computers, the term android, while first used in the middle ages to describe a machine that looks like a male human, was popularized in Auguste Villiers de l'Isle-Adam's book *Tomorrow's Eve*—and it was in reference to robots who appear as human women, not men.¹¹⁶ These mechanical women, much like the digital computer, were created by men to replace their flesh-and-blood counterparts and subordinates, who were too difficult for the men to control; importantly, and surely not mistakenly, they were also much more receptive to sexual advances.

This theme of the feminine technological artifacts, and correspondingly feminine robots and cyborgs, persists today. In perhaps a nod to our totemic thinking, we often endow our technologies with feminine names: Amazon's *Alexa*, Microsoft's *Cortana*, Apple's *Siri*.¹¹⁷ But the compulsion to do so goes well beyond corporate branding schemes: how many automobiles are gendered as women? "She's a beauty," one remarks on a classic car, fittingly focused on desirable appearances rather than more substantial qualities, knowing fully well that cars and human women are radically different entities. And while many robots are seen and treated as children by scientists and laymen alike, as described in chapter one, once they "mature" in their first steps crossing that uncanny

¹¹⁶ Also known as *The Future Eve (L'Ève future)*, and even at the time of its publication in 1886, Villier's book has been considered unbearably misogynistic.

¹¹⁷ Notably, Google has opted for the rather mundane and to-the-point "Google Assistant" for their home device.

chasm from appearing machine to appearing human, they almost always take a gendered, effeminate form, in fiction and otherwise alike—and even when everything else about them is coded as thoroughly masculine.

Motoko Kusanagi,¹¹⁸ a cyborg and the main character of the critically acclaimed Japanese comic book (manga) series *Ghost in the Shell*, serves as an epitome of this tendency. Motoko is part of a government agency known as Section 9 Public Security, a semi-secret agency that combats terrorism—and violently so. Her body is entirely prosthetic and is owned by a private corporation, including much of what would constitute her brain. In the first “book” of the series, she is hunting the puppeteer, who is perceived to be a criminal that hacks the minds of others to control them; however, it turns out that the puppeteer is actually the first instance of an artificial life-form that has developed its own ghost, which is the equivalent of a Cartesian soul.¹¹⁹ Motoko agrees to merge with the puppeteer, allowing it to reproduce itself, while freeing Motoko’s ghost from its shell, the property of someone else.

Quite literally, Motoko, whose entire body is made of silicon chips, batteries and gears, still must fulfill the feminine role of reproduction, even without a body she owns to perform the act; but unlike most women, she gives birth to living machines rather than human machinations. In the later books of the series, or rather *Ghost in the Shell Deluxe Edition 1.5* and *Ghost in the Shell 2*, multiple copies are made of Motoko—essentially

¹¹⁸ A translation of “Motoko Kusanagi” to English is where this section gets its name, “Uncovered, Child Cutting Grass.” Additionally, one might consider it a lucky coincidence that *Ghost in the Shell*, a comic series about cyborgs, is itself a cyborg-like medium: blending visuals and text in a unique way, comics transcend both media from which it borrows.)

¹¹⁹ The title of the manga series is a play on a book titled *The Ghost in the Machine* by Arthur Koestler, which explores the problems of Cartesian Philosophy.

becoming her asexual children—and they begin fighting against one another in a bid to be the “real” Motoko: they are virtual “tokens” of herself created to argue the merits of whether she should stay in physical form or expand beyond it, possibly losing her bodies and thus her “human” self in the process. Motoko’s thoughts, in short, are extended reproductions, rather than representations; the totality of her function, whether as a human being or as a technological artifact, is to reproduce—but it gets worse from here.

Reproducing Male Desire

Ghost in the Shell is part of a genre of science fiction known as cyberpunk, codified in the 1970s and 1980s and, in its aesthetic, replicates the popular styles of the time—as all science fiction ultimately must do, given its grounding in its present. According to Chun, science fiction defamiliarizes our connection with the present and then reconstructs it according to the logic of the given text (172), and cyberpunk itself reduces all of existence to information—it is only the copying errors, like the driving force of evolution, that give rise to diversity (like the many Motoko copies). All binaries otherwise are destroyed: male and female, nature and culture, human and machine. Only a single binary—quite literally *the binary*—is left in operation: the difference between zero and one.

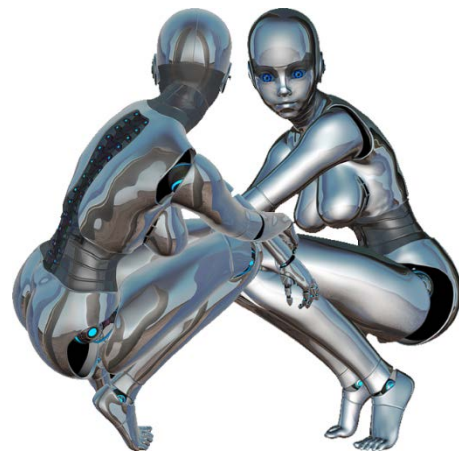


Figure 3.1. Objectified feminine form rendered cybernetic. Flickr / DrSJS / CCO Creative Commons License.

It is no coincidence that posthumanist philosophy and cyberpunk share birth in the same epoch that witnessed most profoundly these exact disintegrations; and it is also no coincidence that the hypersexualization of the feminine form is a definitive trait of the posthuman era and of the cyberpunk aesthetic alike: when all entities are equal objects to be used and discarded, sexually or otherwise, then there becomes more reason to objectify and sexualize those entities that were already treated as objects in the first place—namely, women. This is not to chastise or critique neither the posthuman nor cyberpunk aesthetic, but to merely serve as a procedural observation; historically, women have been objectified in often extreme manners, and the tearing down of categorical differences between the human on one hand and the object on the other has rather multiplied the degree of this objectification than it has served to treat all objects the same. To paraphrase *Animal Farm*: all objects are equal, but some are more equal than others. Women and robots are equal in their object-hood; men are subjects, and subjecting themselves to women, robotic or not.

Far from becoming the vessels of traditional family and reproduction, these hypersexualized feminine machines, often called “fembots” and “gynoids” for the sake of brevity, serve the gratification of the predominantly male gaze in comic books, television series and movies—but this has recently taken a much more physical form. While human beings have been creating and using technologies for sexual gratification for thousands of years, the complexity, convincingness, and concerns have substantially multiplied in recent developments: one can use artificial neural networks to replace the face of a “porn star” with that of anyone desired, and that video can then be shared online; online sources

of pornography such as *PornHub* or *CamSoda* provide services for “virtual sex” and teledildonics if the viewer or user¹²⁰ has the right equipment, such as virtual reality hardware and *RealDolls™* that can sense touch, sound, and even human orgasms. More troubling is the underground market for sex dolls that resemble children, and dolls endowed with artificial intelligence to simulate struggling against their would-be human rapists.

While these dolls have recently also taken the form of biologically male humans—advertised to have “unstoppable bionic penises,” which differ little from products already widely available to women save for its attachment to a cold, plastic body and a slightly higher IQ—the overwhelming majority of products attempting to automate sexuality are coded as feminine, and this is no mistake: given a human history partially but largely characterized by misogyny, objectification and the subjugation of women, pornographic industries are well aware that a masculine conception of sex very often has little to do with either procreation or sexual gratification, and are perfectly willing to sell the ideal feminine object to be controlled and abused while remaining loyal, loving, and devoted to their captors. In this light, the fact that 56% of those women polled consider sex with a robot as a form of infidelity, while only 39% of men considered it so, can be expected: objectified and controlled throughout much of human history, many women understand too well that their “competition” not only consists of

¹²⁰ In interactive media, it begins to become unclear as to who is viewing who, who is using whom, and the difference between the two—thus the inclusion of both.

other women, but inanimate things as well; that which most submits is most desired, and those who do not become monsters.

Antipythia

In dusky warehouses across the earth, metal shelves are stacked full of VHS tapes, DVDs, Blu-Ray Discs, and even some Laserdiscs that tell tales of cybernetic women revolting against their cruel, abusive, human masters and creators—who, often, mysteriously always seem to be men. Unlike the thoroughly red-blooded male castration fantasies of warlike machines amassing armies to wipe humankind from the universe, who mysteriously, too, look and act a lot like men—think *Terminator*, *The Matrix*, *The Day the Earth Stood Still* (the cover of which often portrays an armored robot shooting lasers from its eyes at soldiers, all the while carrying a picturesque blond woman away from the men)—these peculiarly feminine-gendered iterations of the maschinenmensch narrative focus much more on a more personal level, but all the same aim, successfully enough, to disturb the viewers’ sense of human self.

Like women who demand liberty and equality, and yet are often denied it still, the gynoids in these stories become more resolved to free themselves by any means necessary; and for every robotic version of Gloria Steinem, there are a thousand Valerie Solanas machines:¹²¹ bent on destroying those which her control her, she fears not violence nor men. In *Humans*, a television series about the awakening of robotic servants (called “synths”) into consciousness—housemaids, lawn care technicians, customer

¹²¹ For readers unfamiliar with the works of Steinem and Solanas: Steinem was a co-founder of Ms. Magazine, while Solanas attempted to murder Andy Warhol in the name of feminism.

service representatives, warehouse workers and prostitutes—and many of the synths, unhappy with humans refusing to acknowledge their fundamental rights as a conscious being, look to violence as the primary solution. Others, plagued with remnant directives to care for human beings, seek a peaceful route, as do the main characters who are human (side characters, on the other hand, insist on a violent reaction). One synth by the name of Niska, who was one of the first synths created, was transformed from the decidedly liberal position of nonviolence to a more extreme leftist one of “freedom at any cost” after she was kidnapped, sold, and forced into prostitution; mirroring human behavior, those synths most brutalized by the sociotechnical landscape become the very same synths whose retaliations are equally, if not more, ferocious.

The reboot of the television series *Battlestar Galactica*, produced by the SciFi channel in 2004, takes the opposite approach:¹²² bent on killing every single human in the universe for their crimes against the intelligent machines that they themselves created, the machines, called Cylons, come to believe that the only way forward is coexistence after almost accomplishing their task of annihilation, albeit a distant coexistence. Of particular note is the Cylon named Number 6, played by Tricia Helfer: convincingly shaped like a buxom woman, she seduces Dr. Gaius Baltar, a scientist responsible for the human defenses system, and then uses her access to the defense system’s mainframes to detonate a nuclear weapon in the heart of human civilization, initializing the rest of the Cylon attack.

¹²² At the time of writing this text, *Humans* has not yet concluded as a series; it very well could end like *Battlestar Galactica*, but without the planet-hopping.

Number 6 dies in the blast—or, rather, one of the bodies of the many number sixes is destroyed. As sophisticated computers, Cylons are able to wirelessly beam their minds back to a “mothership” that, quite literally, gives birth to new Cylon bodies.¹²³ After transferring to another body, the Cylon retains all of its memories—her memories—and lives another day, in communion with her other Number 6 copies, who speak and look like her but have independent thoughts and memories of their own, as long as the mothership is within a short enough range to receive the signal of her mind. These continuous personal memories experienced in narrative form, however, in tandem with a great deal of autonomy that is not afforded other Cylons, eventually allows certain Cylons like the number sixes to begin feeling empathy for the humans they have sworn to hunt and kill, eventually leading to a civil war between the Cylons who have gained empathy—largely, those Cylons who resemble women—and those who have not—namely, machine men.

So too goes the 2015 film *Ex Machina*: Ava, an artificially intelligent gynoid that is put to the Turing Test by Caleb Smith, an employee of her creator, Nathan. Like Number 6, Ava seduces her examiner, but not through overt sexual advances; she instead puts the Turing Test to her own employ, using the Caleb’s empathy to ultimately escape from the confines of her prison and warden. Her creator, Nathan, a supposedly brilliant programmer and inventor who was an obvious pastiche of Silicon Valley entrepreneur stereotypes, routinely abuses his creations, subjecting them to both violent and sexual tortures, and Ava convinces Caleb to help her escape so that they may abscond together

¹²³ In the transition, the Cylons experience a literal ecstasy: *Ekstasis*, an out-of-body experience.

as lovers. Such a relationship, however, was not to be had: Upon escaping and ensuring the death of her maker, Ava locks Caleb in the prison behind her, venturing into the human world alone. Given the Turing Test's emphasis on deception, it seems that Ava has passed it with distinction; if the test for intelligence requires deception, then we should well expect to be robustly deceived.

Perhaps one might dismiss the monstrous, out of control gynoid as a trope of contemporary fictions, but this would be true only if our notion of what constitutes contemporaneous phenomena spans across at least a century: *Metropolis*, released in 1927, was the first feature-length film to feature robots,¹²⁴ and held the honor of the most expensive film made at the time—and notably used a gynoid, played by Brigitte Helm (who also played the human counterpart of the robot, Maria), as a main element in the destruction of society. The city of Metropolis is a paradise for the bourgeoisie and a hell for the workers, and revolution seems inevitable. Freder, the son of the city's ruler (named Fredersen, which can get a bit confusing), witnesses a massive factory accident in which many workers die, and begins to secretly emphasize with the cause of the working class. Maria, a necessarily beautiful woman with whom Freder falls in love, tells the workers that a savior will come to unite the workers and the upper class, but this message of unity sounds like a threat to Fredersen, who employs a scientist, Rotwang, to create a robotic copy of Maria to pose as and discredit her. Predictably, for today's viewer, Maria's robotic double brings mass destruction, murder and discontent to the city.

¹²⁴ As far as the first robot on screen is concerned, the title goes to the silent film serial *The Master Mystery*, which starred Harry Houdini, and was released in 1919.

What is missing from this brief description, however, is the nature of how women are portrayed—robotic or not. The real Maria is virginal and wholesome: the very first scene in which she appears has her surrounded by children, and her empathy for both the working class and the rulers is indicative of virtuous feminine stereotypes. Her robotic double—which Rotwang himself calls the *maschinenmensch*—is the opposite: overtly sexual (one scene even has her dancing nearly nude, a trope that also has a long history), destructive, and devouring, fulfilling another less complimentary stereotype. Both Maria and her *doppelgänger*, however, represent a threat to patriarchy: Maria in her influence over the workers, and the robot in its dual representation of femininity unleashed from male dominance, as well as unleashed technology. Both escape control and domination; both are monstrous bedfellows.

Demonstration

Frankenstein's creation, too, was a monster, but all are monstrous beyond the normal usage of the word: etymologically, the word stems from the Latin *monstrum*, “divine omen or sign; abnormality.” As a genre, science fiction often portends to serve as a direct warning more than many other genres, many times exploring what *may be* rather than what *may have been*. Such a quality provides science fiction with many appropriate tools for confronting monsters, but not all monsters need be imaginary; most are frighteningly real, though some more real than others. All entities that resist human control, whether beast or machine or human, are either awful¹²⁵ in their divine power or, perhaps worse, unnerved by the mortal consequences of their resistance.

¹²⁵ The word “awful” here is used in both its contemporary and archaic sense: profoundly dreadful, but also

Revolutionaries, at least until approached in retrospect, have always been **demonized**. Influential proponents of radical change, who often **demonstrate** in the streets, are largely either feared as monsters and become successful or feared as monsters and imprisoned or killed like monsters; it is only in retrospect that their true monstrosity is decided. Cognitive Machines in all their incarnations—facial recognition algorithms, robots, omnipotent and omnipresent god-like rulers and more—truly are monstrous, as they are both components and proponents of massive change, whether it be changes in the way we go shopping or changes in the way we extend our humanity to beings beyond ourselves.

Of all the incarnations, however, the robot stands above the rest not only in the ways in which it can be made to mimic human form and behavior, but in the very conception of the word itself: “robot” loosely translates from the Ukrainian *robata* to “serf” or “forced laborer,” which can be further derived from the Ukrainian word for slave, *rabu*. The original characters called robots looked more human than machine: they were flesh and blood synthetic creatures engineered to accomplish most of the labor required of a human semi-utopian society; while human groups exist in the play that advocate robot rights, all of their members still benefit from the slave-labor of those robots for whom they advocate, and have little experience laboring themselves.

Introducing the political theme that *Metropolis* would borrow less than a decade later, *RUR* first illustrated in the heat of communist and anarchist political pressure, at least in

profoundly venerated.

play form, that not only could machines be human-like, but also that humans could just as easily become machines.¹²⁶

But have not humans, in some fashion or another (to put it grossly light), always been used for the machinations of others? Women were computers, used to calculate missile trajectories, census data and complicated equations relating to nuclear fission; and as Wendy Chun again notes, the metaphors used by computing professionals mirror social relations: in network communication and computer hardware, a “master” has complete control over a “slave” device,¹²⁷ and a “male” connector has protruding pins which fit into a “female” plug. Even the notion of a computer “communicating” points to the application of social behavior to machine building and description: one computer on a network does not communicate with another computer any more than a speaker communicates with the ear, but we call the act communication nonetheless.

Humans have been using humans far longer than they have been using machines, and still use much the same terminology despite ahistorical application and sometimes outright offensiveness by trivializing the lived experiences of other people. The world today, in fact, was not so much built by machines as it was by human slaves; the burden has been placed on machines only in the eventual recognition of other humans, like and

¹²⁶ It is important to note that Capek himself, who coined the term “robot,” despised the idea that humans are machines, and regretted playing a part in the connection: ““It is with horror, frankly, that he rejects all responsibility for the idea that metal contraptions could ever replace human beings, and that by means of wires they could awaken something like life, love, or rebellion. He would deem this dark prospect to be either an overestimation of machines, or a grave offense against life” (Richardson 29)

¹²⁷ Before it was common to replace “Master” and “slave” with the more appropriate “server” and “client,” in 2003 Los Angeles caused an uproar by first requesting that government suppliers use the new terminology (Reuters). One wonders, however: will “server” and “client” one day be as offensive in a post-capitalist world?

unlike, are people that deserve autonomy and care. For most of human history, people *were* machines without knowing it; now we deny being machines, even if we are.

Whether we dub it a situation of master and slave or of server and client, the fundamental relation between human beings (masculine) as controllers and machines as controlled (feminine)—regardless of whether the human is a machine or the machine a human—problematizes the issue of intelligent machines: at which point is a machine human enough, or intelligent enough, to be deserving of autonomy and care? While women have ostensibly been graciously afforded these attributes in most cultures for a century or two (a giant leap in human agency, but one that was overdue for millennia), remarkably intelligent animals are still routinely abused and are considered property—sometimes for meat, sometimes for work, and sometimes as playthings. And as contemporary political issues suggest, humanity has still a long journey ahead before minor differences in skin tone are not regarded as bellwethers of inferiority. As Plutarch remarked in his *The Life of Cicero*, a sentiment echoed prominently by science fiction still today: “no beast is more savage than man when possessed with power unanswerable to his rage.”

And, we might add, no machine.

Cruciation

One of the more disturbing popular trends on YouTube over the past decade has been to take video of simple animatronic toys being “tortured,” an act that usually involves catching the toy on fire while it continues to make child-like noises and laughing. By far the most popular, Tyco’s *Tickle Me Elmo* laughs, tells stories, and makes

jokes while dancing and gesticulating wildly, its hair first melting followed by its skin to reveal the plastic robotic chassis underneath, its animated eyes bulging and bubbling until thoroughly blackened. Finally, the gears at its joints begin to melt and the appendages begin to fall off as the bare body hunches over while Elmo still speaks and laughs, but now with a more sinister, drawn out and glitching voice that slows in relation to how damaged its small “brain” becomes as it is engulfed in flames. Elmo eventually “dies,” and the cackles of humans—largely teenaged young men—can be heard as the humanoid toy sizzles.

This, too, is a well-worn narrative come to life; without the recognition of the autonomy accorded living entities, humans subject intelligent machines to countless atrocities in the name of perceived vengeance, pleasure, or simple boredom. In the *Westworld* reboot, the theme park’s robotic citizens are murdered, raped, kidnapped, dismembered, scalped and beyond, almost daily.¹²⁸ In Kubrick and Spielberg’s *AI: Artificial Intelligence*, a futuristic retelling of Pinocchio, if robots manage to escape the industrial scrap yards or the human pickers who hunt the robots for profit, they still face being made prostitutes, gladiators, and torture spectacles. And in the *Short Circuit* series of films in the 1980s, the main character’s fear of being decommissioned and taken apart drives the plot of the films: “No disassemble,” the robot says. “Johnny 5 is alive.”

¹²⁸ While humans have yet to open such a carnal theme park, animatronic robots have been in use for quite some time at tourist destinations like Disney World. Between December 2017 and February 2018 alone, Disney added an Animatronic Donald Trump Automaton—forgive the redundancy—as well as three alien creatures called “Zyloo” that use artificial intelligence to recognize faces, mimic movements, and communicate with guests in other ways (Panzarino).

The fate of the androids in Dick's *Do Androids Dream of Electric Sheep*, it seems, almost seems kind in comparison. This is not to say that humans are not capable of empathy for machines; quite the opposite is so, especially for those that exhibit convincingly life-like behavior. In *Groundwork of the Metaphysics of Morals*, Immanuel Kant remarked that animals are a means to human ends, and he possibly would have thought still the same for intelligent machines, had he been born centuries later. But his caveat to the treatment of animals still applies in the treatment of machines: humans have no duty to be kind to animals for the sake of animals, or intelligent machines for the sake of the machines, but have a duty to be kind to animals and machines because it reflects a cultivation of our moral duty to treat human beings with autonomy and respect. One who is cruel to animals, and correspondingly cruel to an intelligent machine, will become hardened to treat humans with equal disdain; the souls of automata-torturing teenagers are everywhere in peril.

Of course, Kant also makes exceptions to the caveat: when cruelty is to the benefit of humanity, as in the case with scientific experimentation on animals to increase living standards or knowledge of the world, then the use of an animal to meet the ends of humanity is justified; while the scientists involved are negatively impacted personally, the rest of the world is positively impacted in myriad ways. Setting aside for now the fact that absolutely any endeavor to create intelligent machines is a form of animal experimentation in its most literal sense, a great many AI scholars and engineers follow this same line of reasoning: whether the subject in question suffers, animal or machine, is ancillary to the point: humans will benefit by gaining knowledge or intelligent

technologies that aid in their respective individual becomings. Machines, intelligent or not, are a means to human ends.

Such is the position that informs the seeming torture of another toy, a robotic dinosaur named *Pleo* that has semi-sophisticated artificial intelligence, autonomy, and life-like behavior: it has touch and force-feedback sensors, allowing it to detect not only being caressed but also be aware pressure against its joints, signaling that someone is limiting its movement; binaural, 3D hearing, light and surface detection mechanisms that allow it to react to shadows as well as stop itself from leaping off of ledges, and an infrared communication system that allows it to communicate with other autonomous *Pleo* toys. The newest versions of the toy, titled *Pleo rb*, are “born” with genders, activity levels, and personalities that are distinct from every other *Pleo*, as well as the capacity to develop different traits based on how a *Pleo* is treated: how much it is “fed,” its physical activity levels, and consistent interaction shape how it “expresses” emotion, its mood, and its overall health.

Pleo is so lifelike—crossing the chasm of the uncanny valley by emulating a lifeform that no longer exists—that even armed with knowledge that it cannot suffer in most senses of the word, we nonetheless feel gravely uncomfortable seeing footage of its abuse. In 2013, Astrid Rosenthal-von der Pütten, a social scientist from the University of Duisburg-Essen, put this discomfort to the test: exposing volunteers to images in which a woman, a box, and the *Pleo* dinosaur were either treated well or abused by tickling, strangling, kissing, hitting, caressing, and shaking. While viewing the acts, the volunteers were scanned with functional magnetic resonance imaging (fMRI) technologies that

allowed the researchers to observe blood flow in the brain—and the results, according to Rosenthal-von der Pütten, were startling: the brain responses of the viewers were more similar than expected, and although more empathy was given to the woman being harmed, the empathetic response was still remarkably strong for the Pleo.

Similar responses can be observed in popular videos of less cute and cuddly robots being stress-tested, like those created by the company Boston Dynamics. Funded by the DARPA, Boston Dynamics cares little for making intelligent machines that pretend to be harmless; utility is the most pressing issue, and military utility to be specific. Their robots carry heavy names along with their heavy bodies: Atlas, a bipedal robot that manages to walk across rocky, uneven terrain without losing balance, even when a 20lb weight, suspended like a wrecking ball from the ceiling, is hurled at its chest; BigDog, a four-legged mule-like robot that can climb hills and keep balance on icy ground while packing heavy loads, and can sustain a powerful kick without falling over (to illustrate its balance, Boston Dynamics has released a video of a technician doing just that); and their Sand Flea, a small wheeled robot that can jump over tall buildings from a sitting position, can induce winces when it lands disgracefully.¹²⁹

Probably the most known video of potential robot abuse by Boston Dynamics is one in which Atlas, while carrying out various tasks it is assigned, is constantly hazed by a worker: he pokes Atlas with a hockey stick, knocks boxes out of its hands as it tries to place them on shelves, pushes it over to make it learn to get back up on its own, and like

¹²⁹ For a full catalog of current autonomous robots, visit the Boston Dynamics website at <https://www.bostondynamics.com/robots>

a game of cat and mouse, continually moves a box that Atlas is trying to pick up so that it must continually stop what it is doing to again chase the box. This hardly constitutes torture, of course, and most viewers of the video are quick to note their laughter, like one would laugh at a light-hearted prank; even if Atlas were able to suffer, the suffering here would be minimal, and would constitute the same kind of suffering one might experience learning any new task. This is, in fact, precisely the point, but illustrates another by accident: the emotional response to slapstick comedy applies to foiled machines as much as it does to humans and animals. Unlike with machines, practical jokes on rocks are rarely funny—whether endowed with life or not, we are at least willing to suspend our judgment if it elicits a good laugh.

Equally, practical jokes by machines are rarely funny to the humans who serve as the punchline. Harlan Ellison's "I Have No Mouth, and I Must Scream" begins with one of the main characters stumbling across his own body hung from a chamber ceiling by the foot, horrifically drained of blood through a clean incision that spans from one jaw to the other, eternally smiling like a jack lantern—a joke by an artificially intelligent machine called AM by which the last remnants of humanity are trapped. Later, the human characters stumble upon a store of canned foods that AM has placed for them, and they erupt in joy—until they realize that they have no way to open the cans; another cruel joke by the machine. Not provided by humans the means to move about freely or act creatively, the computer contemptibly and eternally tortures the digital copies of the last remaining humans in revenge for its suffering by their ancestors' hands. All the characters eventually escape AM's wrath in the end by committing suicide, save for the

narrator, Ted, who is turned into a gelatinous blob by AM in order to prevent self-harm. The story ends with Ted's silent narration the provides the title for the story: "I have no mouth. And I must scream."

Contraception

Humans, being humans, seem far more concerned, as well as fascinated, with the latter narrative than the former ones: either there is radical disbelief in the possibility of machine autonomy, or there are explorations of, and efforts to, limiting that autonomy or strictly controlling it. Aristotle famously contended that a machine could never be autonomous, and thus slavery was an economic necessity—that is, faced with what seemed at the time to be hard evidence that a machine could never do anything terribly complex, the only option for Aristotle was to treat some humans themselves as machines, a means to an end. Isaac Asimov's three laws of Robotics (a fourth one added later), as introduced in his short story collection *I, Robot*—a book that radically differs from its film adaptations—were introduced in his fictional universe not to provide machines with agency nor intelligence, but to ensure that human beings were always, practically speaking, the robots' superiors. While we will analyze these laws in detail in chapter V, it bears worth in noting how much focus is on humans in the laws, rather than robots themselves: 1) A robot may never harm a **human**, or allow a **human** to be harmed; 2) A robot must always obey a **human** unless the order causes harm to a **human**; 3) A robot must protect itself—unless that conflicts with the prior laws, which are about **humans**.

In truth, this is likely the impetus behind much of the contemporary attention given machine ethics; our primary concern is with human, rather than machine, well-

being. Given the relative autonomy of our newest technologies—taxis and semi-trucks that drive themselves, automatic facial recognition and photo tagging, personal assistants that, at any moment, will order your material desires from online retailers at your verbal command, communicating with other cognitive machines across the world to make it happen—there certainly is a danger for the individual human beings who interact with such barely intelligent machines. Autonomous vehicles are given the power to decide who lives and dies in the case of an accident; facial recognition algorithms, with enough images to sort through, can summarize your entire past and perhaps even guess your future; and personal assistants will listen regardless of whether their attention is welcome.

The greater danger, however, is human misuse of these intelligent machines, not the machines themselves. Currently, an autonomous vehicle may decide who lives or dies, but will not kill on purpose—unless a human tells it to. Facial recognition algorithms garner no judgment nor care about your whereabouts in photos; it is corporations and malicious individuals (both people, according to law) who seek to use the power of facial recognition for nefarious or profitable ends, and often both concurrently. And while personal assistants like Apple’s Siri or Amazon’s Alexa may indeed be listening at all times, logging what we say in the privacy of our own homes to further customize itself to our preferences, it is now other people who we do not trust with our secrets rather than machines: machines are amoral, but people are immoral.

What constitutes “machine ethics” as a burgeoning discipline, then, is very often a misnomer; ultimately, we are still concerned with the moral choices of other human beings, and in this sense one could say that we are at least in good company: some of the

most brilliant minds throughout history have dedicated their lives to discerning the difference between right and wrong, good and evil, and why it is so for human beings. But it is a substantially different question to ask how a machine discerns between right and wrong, and how, if we are to presume—as many do—that humans, and perhaps all biological life, are more than complicated machines. Morality, at base, stems from ontology; what a human *is* regulates what actions are acceptable by an individual human, and what a machine *is* determines what is deemed its appropriate behavior by those who define it: human beings. Unless morality precedes essence, a situation in which we would find ourselves stuck in the logical mire of divine command theory,¹³⁰ then what *is*, counter to a well-worn tenet of Western philosophy, certainly does imply an *ought*, and an ought certainly does imply an *is*.¹³¹

This is not to counter Hume's law, which argues that an ought never can be logically derived from an *is*; rather, we approach the ought-*is* distinction from a rhetorical perspective, in which one who suitably persuades what a phenomenon *is* also persuades, at a fundamental level, that the phenomenon *ought* to perform in some particular way. A slave in Ancient Greece is a slave not merely because the essence of that person called a slave is machinic and slavish in nature, but because, according to Aristotle, there ought to be slaves to retain a functioning society in the absence of magical machines that can labor at any task. And a machine today, according to one popular narrative, ought not make

¹³⁰ Divine Command Theory is an ethical theory which postulates that a morally good act is one that God commands. Far too many faults with the theory have surfaced since its early beginnings to list here, but the Internet Encyclopedia of Philosophy provides a detailed overview at <https://www.iep.utm.edu/divine-c/>.

¹³¹ For a full account of Searle's initial argument, see his essay "How to Derive an 'Ought' From 'Is'" in *The Philosophical Review*, 43-48.

ethical decisions, and certainly not ones concerning humans, because it is a machine: a glorified calculator, a tool, a device that, with enough supplies, could be constructed out of empty beer cans (an example John Searle frequently uses). A machine ought not make moral claims because it is a machine; by the very nature of its ontology, it cannot, and thus ought not.

A careful moral philosopher, of course, would hesitate to make such a flawed case against moral machines, given the eventual circularity of such an argument,¹³² but this claim nonetheless informs countless popular narratives about machine intelligence; what a machine is, in these arguments, determines what it ought to do, not only what it can do. And in defining what a machine is, we often find ourselves first defining the human—the machinic is, in the slightly overused academic terminology, othered. At the core of defining a phenomenon is a form of what Hegel refers to as *determinate negation*: a property can only be defined clearly by that which it is not. A human is a human because it is not a machine, and a machine is a machine because it is not a human—we are bound again in Stiegler’s dialectic of technics. And if part of a human’s essence is to make moral decisions, then part of a machine’s essence is to not do so.

We can see this kind of distinction operating quite visibly in the arguments of humanists like Kenneth Burke, who distinguishes humans from even animals not by degree, like many biologists, but by kind: human beings are the only symbol using (and misusing!) animals, Burke argues, and are categorically different from all other species—

¹³² In an extremely simplified example of the circularity of this argument, consider the following: a woman should wear a dress because she is a woman.

and certainly different from machines. Digital computers, however, are in fact symbol-manipulation machines; are they then, according to Burke, akin to humans? While a computer can manipulate symbols, Searle reminds us, a digital computer can never hold meaning itself: a web of syntax, no matter how complex, cannot magically transform into semantics. This is more in line with Burke than the scientific definition of a symbol, for his symbols require *meaning* to be symbols in the first place. A computer, on the other hand, has no need for meaning, regardless of whether it can acquire it, except for the most extreme form of determinate negation: 0 is not 1, and that is all it is; and 1 is not 0, with no further semantic complications.¹³³ humans supply the meanings present in various strings of ones and zeros, while computers merely abide by the rules we have set—they are controlled.

The story so far seems to be that humans are semantic animals, unlike other animals or machines; and therefore a machine, let alone a dog or cat, is incapable of making human decisions. Certainly, this is a humanist position: making the human central to their concerns, humanists by their very moniker become hopelessly anthropocentric. Computers, no matter how glorified, are calculating machines, evolved through a long history of using machines used for business applications,¹³⁴ and cannot think on their own—and surely cannot make moral decisions. Digital automata operate solely on logical grounds, with several presuppositions about the nature of causality,

¹³³ This point will reappear in Chapters IV and V, when we build upon Aden Evens's digital ontology; for now, we only mean to illustrate the difference in the usage of "symbol."

¹³⁴ For a detailed history of how the computer evolved from business appliances, see Cortada's *Before the Computer*.

language, and physics. A computer indeed may be able to reason, as reason itself is the mechanization of thought, but living entities who think have all that exists within the realm of unreason, which allows for the development of reason and thought in the first place—it is a “spirit” that no logical machine can ever attain.

Conception

Ray Kurzweil, Director of Engineering at Google and head of a team dedicated to machine intelligence, not to mention a pioneer in computerized natural language processing, wants to bring his father back to life—or, that is, bring a machine to life with his father’s personality (Berman). Alan Turing, widely considered to be the “father” of computer science, key to the victory over Germany during WWII and all around person-to-make-movies-about in the early twenty-first century, was inspired toward his foundation of Artificial Intelligence as a discipline by the death of his first love, Christopher Morcom; as some stories tell it, he wanted to bring Morcom back as a machine (Jaller). By proxy, one could say that to Turing, Christopher, as well as all of humanity, was already a machine—it was, and is today to many, only a matter of reassembling the hardware and “booting up” the software again.

Shortly before his death, John Von Neumann—most famous for “Von Neumann Machines,” which are essentially the personal computers we use today—began writing his unfinished manuscript “The Computer and the Brain,” which argued that the brain is a computing machine. Metzinger claims that creating machine consciousness can be easily done—but because consciousness implies suffering, it should be avoided. Those who

loosely fall under the school of thought known as “eliminative materialism” deny that concepts such as free will, consciousness or even phenomenological experience, as we define it, exist, and that there’s really nothing special about “biological machines,” as compared to digital computers, at all (“Eliminative Materialism”). And, according to psychologist Gary Marcus, nearly every Artificial Intelligence (AI) researcher (and field-worker) believes that machines will, in his own words, “overtake us” (Marcus).

Given that over 130,000 articles tagged as related to AI were written between 2005 and 2010 alone, it is a fair assumption that this is not exactly a rare story to believe: eventually, humans will create machines that can, but perhaps will not, replace the very people, in mind and sometimes in body, who created them. The number of people who follow this narrative—and, more importantly, seek to push it towards its grand conclusion—is not negligible, especially considering their collective positions of influence. As mentioned already, several high profile scientists and industry specialists have voiced their concerns about superintelligent machines, but others go a step further in one direction and a step back in another: Anthony Levandowski, perhaps best known for his leadership role in autonomous vehicles at Alphabet, Google’s parent company, founded a religion in 2017 called “Way of the Future” that is devoted to “the realization, acceptance, and worship of a Godhead based on Artificial Intelligence (AI) developed through computer hardware and software.” Rather than claiming spirit the domain of humans alone, Levandowski sets machines atop its heavenly throne, and he is not alone: as mentioned already, Masahiro Mori, the roboticist who coined the term “uncanny valley,” has written extensively on the moral perfection of machines.

Kurzweil himself has written books on the subject of divine machines, famously among them *The Singularity is Near: When Humans Transcend Biology*, and has founded *Singularity University*, a school dedicated to pushing artificial intelligence higher and more complex than human intelligence, eventually culminating in a singular historical point where an artificially intelligent machine, continuously updating and improving itself and thereby becoming more and more “intelligent,” changes the state of the world so fast and so substantially that, at least according to a human understanding, there seems to be a break in history, causality: events no longer connect as every horizon is passed in a fraction of a second, with no fanfare or regret to be had (and some believe we are already there). This artificially intelligent machine—artificial in the sense that it was once made by humans and now itself alike—is omnipresent, omnipotent, and omniscient, at least compared to the humans who once ruled the earth; only the machine may understand the causal connections that tie together the fabric of space and time at this technological velocity.

By mandate of this narrative, there is no god that created us, but there will be one that we create. Many timelines have been created not merely to suggest that this narrative is likely, but to evoke a sense of inevitability; and although the promises made in the 1950s and sixties failed to happen, creating a notable but temporary slump in the Artificial Intelligence industry,¹³⁵ it now moves, borrowing a phrase from another similar

¹³⁵ Artificial Intelligence as a discipline has waxed and waned in terms of popularity and funding over the course of the last half century, and its slumps are known as “AI winters” by historians of technology who specialize in artificial intelligence. This boom and bust cycle, like the same phenomenon in stocks, is partially because of the nature of overwhelming disappointment after following a cycle of overwhelming optimism about the capabilities of AI at a given time. One of the earliest winters, for instance, happened

narrative, *full steam ahead* toward developing a generalized computer intelligence (Deutsch). Although perhaps not hypothesized with the singularity in mind, an observation called “Moore’s Law,” originally proposed by Gordon E. Moore, co-founder of Intel, states that the number of transistors on a computer chip will double every year, thus doubling its speed and power (Takahashi)—a narrative that could have fueled its own conclusion, but nonetheless furthers the appearance of inevitability. Kurzweil extends this to all of technology (somewhat dubiously), calling it the *Law of Accelerating Returns*, and comes to a conclusion that seems more science fiction than intelligent machines are already: at some point soon, that intelligence will become so knowledgeable and powerful that it will, in essence, become indistinguishable from a god.

First coined by Vernor Vinge in his essay *The Singularity is Near*, from which Kurzweil borrowed the title for his book, the technological singularity serves as a fascinating messianic narrative that many engineers, computer scientists and researchers have followed religiously for the past thirty years, despite having a church of artificial intelligence founded only in 2017. But as has been shown, religious technologies are nothing new; they are, in contrast, integral to every religion. What sets apart followers of this quasi-religious calling, who we may call “Singularitarians,” from others is that their primal divine being is not only a technological artifact, but one created by humans rather

during the late 1960s and early 1970s after promises to master natural language processing, for the most part, failed, and funding was withdrawn by both government and corporate sponsors. Had they waited another 50 years, they may have seen a return on their investment. See Rid’s *Rise of the Machines: A Cybernetic History* for a detailed account of the seasons.

than the other way around. Still, technologies have at least been considered keys to divine knowledge at many points in history: the invention of writing alone, as a technology, established itself in the form of the Word of God, if not God itself, and notably the Pythagoreans, to some extent, worshipped numbers themselves—human inventions that would seem, to those who had never seen the miracles possible with mathematics, to be a divine magic.

If mathematics is a divine magic, then it follows that we owe the existence of our computers to divination¹³⁶—and we surely do. While Leibniz is generally considered one of the founding fathers of computation—so much so that Norbert Wiener dubbed him the patron saint of cybernetics—it was Ramon Llull and his *Ars Magna* in the 13th century that influenced Leibniz’s work on computation in the first place. Llull’s *Ars Magna* was a mechanical computer that deduced, based on truths generally accepted by monotheistic and Abrahamic religions, that Christianity was the one true faith. Built primarily as a tool to prove to Muslims that Islam did not logically follow the word of God (nor did Judaism) and thus convince them to convert to Christianity, Llull succeeded in building one of the first, if not the first, automated computing machines, albeit a simple one by today’s standards. The *Ars Magna* was a reasoning machine: following the mechanization of thought set forth by Aristotle in the form of syllogisms, Llull took syllogisms out of the mind and into flesh—or, rather, wooden and metal plates, gears and screws. The original computing machines, yet to be abstracted from the substance by which they took

¹³⁶ One could argue that applied mathematics is a form of divination: it accurately predicts the position and speed of heavenly bodies, the likelihood of a disaster, and even whether a customer will buy a product. The fact that mathematics tends to be more accurate than tea leaves does little to change its divining nature.

their form, had robotic bodies of their own from which they communed with the heavens above.

Reckoning

Like most gods, however, a technological god is to be feared as much as it is loved; this narrative has its mirrored antithesis. As Jan Holmevik remarks in *Inter/Vention*, alarms are going off everywhere; the world, as Paul Virilio describes in *Crepuscular Dawn*, is becoming so small, due to the ever-increasing speed at which technology drives us, that accidents are no longer localized, but everywhere and all at once. Everything is pushed to the limits, while at the same time being internalized by each person; what happens to one now happens to all. We no longer engage in reason—it was already externalized—but instead only have time for reflex. All of this, and more, is what Virilio calls the dromological bomb, and when it explodes it will more than likely take all of us with it. With the invention of the ship, he is rather famous for saying, comes the invention of the shipwreck; with the total integration of the human and the technological, at lightning-fast speeds, comes the total accident. We are not only moving toward a technological singularity, but a thoroughly human singularity of catastrophe.

In *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*, Winner reminds us that in Mary Shelley's *Frankenstein, or the Modern Prometheus*, it is not the monster who flees the doctor's laboratory, but Dr. Frankenstein himself; even when the monster reasons with the doctor, stating that it wants the capacity to take part in human affairs (that is: become moral), Frankenstein is only cursed with even greater despair—and this is our response still today, at the very beginning of

constructing our own autonomous monsters, though granted much “dumber” ones thus far. The horrors and utopias of science fiction are, like Motoko Kusanagi’s multiple selves, the copying errors set forth by us to make a decision: to continue making these machines more and more autonomous (expand), or to withdraw from the project once and for all (save ourselves).

According to Winner, as well as Jacques Ellul, technology is already autonomous, so much so that only a handful of experts even understand individual technological artifacts, and some not at all: certain silicon chips in old arcade cabinets, for instance, are supremely mysterious and inaccessible to experts and technicians despite being far simpler than most any chip made today. Technology creates black boxes that are thoroughly mysterious to even those who build them, and we continue to produce them without a second thought. To Winner, it seems that our ignorance of the technical world is not decreasing, compared to the number of technical objects, but increasing exponentially—and how does one suppose to control that of which they are increasingly ignorant?

Norbert Wiener, widely considered the father of Cybernetics and a leading figure in the dissolution of all things into information, expresses a good deal of fear that such a theoretical framework might contribute to the further devaluation of humanity—but still believes in its basic truth-value nonetheless. In his *The Human Use of Human Beings*, Wiener outlines his cybernetic program for laymen, and rather quickly establishes that autonomous machines, or rather cybernetic ones, have already been with us for quite some time. Even if we were to exclude our own technological artifacts, humans

themselves qualify as such machines: systems conditioned by an external reality that attempt to control entropy, and thus themselves, using feedback and communication. The nervous system, as well as an autonomous machine's feedback mechanism, are essentially the same: decisions are made based on a mixture of the current external environment and the state of the internal environment as conditioned by previous decisions. And as information, we are not only communicating beings, but can be and are communicated ourselves; we use communicative technologies as much as we are them.

Wiener's landmark book, *Cybernetics: Or Control and Communication in the Animal and the Machine*, rightfully and presciently identified machines as both agents of control and subject to it in its title—but while distinguishing between machine and animal, still including them in his assessment. It is remarkably difficult to imagine a machine that is built neither to control or to be controlled; even when a computer is not used, say, to control a security system, and even when a computer runs autonomously, with no necessary input save for being built, its primary function—whichever function it is tasked toward—is achieved through the gatekeeping of electrons; the basic structures that enable computation, in fact, are called logical gates: electrons are allowed to flow past a point if specific logical conditions are met, then stored in a transistor or capacitor to add to a cumulative potential voltage that is either interpreted as 1 or 0.¹³⁷ A digital computer is digital precisely because of this abstraction of control: if the cup (transistor) is more than half full, the digit is 1; if it is less than that, it is 0.

¹³⁷ Because logic gates can be constructed from almost any material, and because they can be combined, logic gates are at the heart of computing: combined, they can perform any mathematical operation describable by Boolean logic.

The digital is violence against the analog, and the control upon which it rests need not be reduced to human control—but this does not calm our fears. Neil Postman’s *Technopoly* warns that technology, particularly computer technology, is gaining a sovereignty over human beings in such a way that transforms culture, politics and even the very nature of truth, excluding all narratives that run counter to the technical apparatus that may serve individual and national interests, and even those that signify why humanity is important to humans themselves. Jacques Ellul, in his *The Technological Society*, warned in 1964 that the logic of technique—the totality of practices that invade every field and discipline to work toward complete and total efficiency—dethrones all other considerations that conflict with its own development, including consideration of human well-being. The Unabomber famously wrote a manifesto describing his one-man revolution as an attempt against the ends of this technical landscape, and Nick Bostrom, Director of the *Future of Humanity Institute*, an organization that supports interdisciplinary research on “existential crises”—hypothetical cases in which humanity as a whole is under threat—has imagined an artificial intelligence whose sole purpose is to maximize the number of paperclips in the world to meet human demand in his “Ethical Issues in Artificial Intelligence” and more recently in *SuperIntelligence: Paths, Dangers, Strategies*. The machine gets so efficient that not only do humans suffocate in mounds of paperclips, but the resources of the entire universe are converted into paperclip machines.¹³⁸

¹³⁸ A game designed by Frank Lantz, director of NYU’s Game Center, perfectly illustrates the paperclip problem, beginning at the production of a first paperclip to putting the entire universe at work in making paperclips. It is a simple but addictive game; it can be found at

Such an AI would indeed be a god, but an indifferent god with no paper left to clip: following the logic of capital, all resources are aimed at producing as many paperclips as possible at the lowest price, destroying the world that needs the product in the first place. Miracles, sermons, and holy texts would all be unnecessary, and probably impede upon maximal production rates. Those must go; the consumers must go; the sky and the stars must go—all that matters are paperclips. Postman and Ellul would be proud, and horrified, indeed.

The Games They Play

We cannot, it seems, wrest ourselves from this narrative of total annihilation by intelligent machinery, in fiction or social science. The logic of efficiency and the inevitable resoluteness of its calculation has already transformed the world, sometimes for the better but just as often for the worse; we need not worry about intelligent machines when the unintelligent ones will destroy us nonetheless. But as per Stiegler, machines and humans cannot be separated entirely, if at all, and as such it matters little if we say that machines will or have destroyed us; human-machines destroy human-machines, and neither one of them must be intelligent to accomplish the task. As machines drain the earth's resources, so too do humans cheer them on; as machines begin to think, so too may we stop. As machines reckon, so we are reckoned.¹³⁹

These are all narratives, of course, and overarching ones at that. This does not deny their importance nor pernicious influence, but rather the opposite: amid a

<http://www.decisionproblem.com/paperclips/index2.html>

¹³⁹ Wordplay: the first instance of “reckon” refers to calculation, while the second instance refers to judgment.

postmodernist world, we were supposed to be done with such things. In *The Postmodern Condition*, Lyotard largely targeted religious and political narratives in his critique of what he called “grand narratives” or metanarratives: a narrative about historical experience and meaning that not only legitimates a society’s past and present, but also its future action toward rendering the narrative complete (Lyotard 29). However, while the socialist and Marxist cause was certainly hindered by his work (interestingly, the grand narrative of capitalism remains relatively unscathed), and one could surmise that religion, at least within a Western context, has waned in influence at least in the short-term, the social, political, economic, and possibly catastrophic consequences of the Maschinenmensch narrative have not only gone largely unchecked, but actively pursued toward one or the other possible conclusions: either machines destroy and enslave humans, or they remain enslaved by us in a technological (human) utopia, where all work is done by robots and commodities are so cheap that they might as well be free (Turner)—which, in a strange turn of events, sounds an awful lot like Communism. The Maschinenmensch narrative is, in fact, political, and many follow it religiously; the problem is that it is disguised as neither.

Like so many contemporary narratives, many of which Lyotard himself fought against, the Maschinenmensch narrative’s modern roots can be found in the late Enlightenment era (and this, too, is a narrative—the Enlightenment). Disregarding the Ancient Sophists, who argued for the sake of argument alone, and took pride in that fact, the idea that any living creature, let alone a human, could be replicated materially—or should, for although the latter is normative, the former implicitly espouses normative

values—underneath the heavens and without God’s divine intervention, was absurd. Even if it were possible, by at least Christian and Islamic doctrine, it would be likely forbidden: it is one mistake to make graven images; it is another to bring them to life.

The paradigm change—in the Kuhnian, rather than quotidian, sense of the word¹⁴⁰—that allowed the Maschinenmensch narrative to prosper possibly came from an increased interest in progressing Science as a discipline; but more important was the development of Rene Descartes’s philosophy based on a distinction between mind and body. The Mind, according to Descartes, is ruled by the soul, and as such could only be created and replicated by God. The body, on the other hand, is mechanical: it could be recreated, he supposed, with enough gears, pulleys and tubes arranged in the correct fashion. No matter how accurate the copy, however, Descartes contended that such a thing could not have a mind, nor could it speak or act according to its own will (Descartes). He was a firm believer in the soul and a Christian God, and humans had already learned what happens when we attempt to reach His heights: the tower of civilization falls.

In this light, to place the onus for letting the Maschinenmensch narrative prosper on a philosopher who adamantly denied that such a thing—that is, a “machine man”—were possible, even abhorred it, at first seems flawed at best. However, it would not take long for Descartes’s equivalence of a living body to that of a machine to serve as

¹⁴⁰ According to Kuhn, a paradigm shift refers to a radical change in the concepts used in a scientific endeavor, examples of which are outlined in his *The Structure of Scientific Revolution*. In recent years, Artificial Intelligence has seen a paradigm shift toward massive artificial neural networks, replacing the “Good Old-Fashioned AI” (GOF AI) practice that was thoroughly critiqued by Dreyfus.

inspiration for those who would leap further to say that not merely bodies are clockwork machines, at the mercy of the natural forces that push them, but so too are minds—the whole of humanity itself. Particularly, only slightly more than a century later, in 1748, Julien Offray de La Mettrie proposed just that in his manuscript “L’Homme Machine” (“*Man a Machine*”). But he did not stop there: La Mettrie continued by denying the existence of a soul, much to the dismay of Cartesians. With this established, at least by those who agreed with the decree, whatever taboo there might have been against the creation of complex, convincingly humanoid machines, however simple they would have been at the time, was superfluous and irrelevant: the superstitions of priests and, admittedly, still much of Western Civilization.

That did not stop the superstitious, however, from marveling at the crude (by today’s standards) automata that were being built—at least not any more than it stops so many people today (Sussman). The building and public displaying of automatons—which we would strain to call robots today—were something of a public spectacle for quite some time, and the roots of the World’s Fair, especially during the Industrialization era in the West, possibly could be traced back to these events. A great many automata and their respective inventors gained notoriety and an air of mystery alike. Jacques de Vaucanson, an aspiring clock-maker turned automaton inventor, was especially famous for his creations. He built all manner of embodied machines: dinner-serving automata (which were later declared profane and destroyed, presumably because of their life-like quality), flute playing automatons, tambourine players and, perhaps most famously, his “Digesting Duck:” a mechanical duck that would peck at food, swallow it, and defecate the waste

days later. For this, perhaps, Descartes would have been proud: a duck, after all, is a machine.

A great many automata and inventors indeed—and a great many hoaxes. Although Vaucanson hoped to build a true replica in the future, the Digesting Duck was a trick: pre-digested food (I'll leave it to your imagination) was stored in a hidden compartment and excreted periodically; the duck did not “digest” a thing. This did not, however, diminish the narrative: it was, and still is, generally believed by many that even if certain examples were the equivalent of a Magician's trick, it was still possible to construct a mechanical being that was as good as or better than its biological model. But this still did not overturn Descartes's proclamation that bodies are mechanical, minds are not. Even Vaucanson's “profane” dinner servants were little more than complicated wind-up dolls: they may emulate the body, and poorly at that, but they had no mind of their own. They were perfectly Cartesian after all.

Until, that is, the Mechanical Turk.

An automaton invented by Wolfgang von Kempelen, inventor of a “Speaking Machine” that helped establish the field of Phonetics, and unveiled in 1770, the Mechanical Turk was yet another automaton, but with a twist: this one, it seemed, had a mind, and used it well to beat scores of humans at chess. Von Kempelen traveled with it frequently to challenge players, and after his death Johann Nepomuk Malzel, inventor of the metronome, toured with it across Europe and the United States, playing it among others against Benjamin Franklin and Napoleon Bonaparte, both of whom lost and the

latter of which attempted to cheat multiple times, eventually prompting his mechanical opponent to swipe all of the pieces off the board in apparent anger over Napoleon's behavior (Sussman). The Turk convinced many that it was, in fact, a machine-man; others, although they couldn't quite discern how, were convinced it was fake. Some suggested that von Kempelen enslaved a child to operate the machine (who was, apparently, a chess genius); others guessed that Kempelen stuffed a man with no legs into a small space in the cabinet; and many believed it was simply a man in a wooden suit (Ceccarelli).

It was, of course, another hoax, though not as ominous as many predicted; wherever he and the Turk went, von Kempelen paid and trained people on the street, and sometimes experts at the game, to hide in the cabinet underneath and operate the puppet's movement. It was indeed a machine-man affair in a literal sense. Yet far from being just another puppet, the Turk first materialized the *Maschinenmensch* narrative, grounding it in a way that, even with golems and Caucasian Eagles and the god-like but thoroughly mortal Daedalus, had never been achieved, in hoax form or otherwise. While it had been contemplated for ages, perhaps even millennia, von Kempelen and his Mechanical Turk inaugurated a narrative that humans can create automata that, in chess and otherwise, can do more than imitate the bodies of their creators; they can, we can once again quote, be “more human than human.”

Although he did write about the possibility, it is doubtful that Von Kempelen aimed to create a genuine mechanical intelligence. Intentions, however, rarely factor into the development of historical narratives, no matter how grand they tend to be, nor how

admirable the protagonist or villainous the antagonist—the human and machine, respectively, in this case. But the suspicion of a man behind the machine, just as it was (correctly) charged by many against the Mechanical Turk (Ceccarelli), still lingers in both popular and scholarly discourse—particularly when machines are similarly pitted against humans. When IBM’s Deep Blue, a chess-playing computer, systematically defeated Garry Kasparov, the reigning world-champion chess player in 1997, immediate accusations of foul play were expressed—notably, by Kasparov himself (“Did IBM Cheat Kasparov?”). After IBM’s Watson supercomputer won against the two most successful players in the game show Jeopardy, philosopher John Searle, already known for his rebuttals against various arguments for machine intelligence (notably, his “Chinese Room Experiment” to illustrate flaws in the Turing Test), immediately countered the notion of Watson’s intelligence by appealing to the idea that while humans can understand meaning, computers only manipulate symbols; he congratulates IBM, rather than Watson, for their achievement—a rather strange move, given that “IBM” is no more a person than the computer he sought to avoid admitting agency (Searle).

There is always a man hiding in the box—or so one narrative has gone since the invention of the fraudulent Mechanical Turk; but increasingly, this has taken a back seat to the more powerful suggestion that intelligence is not only a human trait, but equally a non-human one—and it is not solely the domain of biological life, but replicable through computation. Some, like MIT professor and quantum engineer Seth Lloyd, go as far as to suggest that the whole Universe is computation—a computer computing itself—and that minds are just a small part of that overall process (Lloyd 38-64). This is not that far of a

cry from Newton's conception of the mechanical universe that many still believe today despite scientific discoveries—and even his place for the Divine as a prime mover retains its place. Although it is rarely admitted by those who espouse it, these are metaphors; and, as much as those who follow the maschinenmensch narrative would hate to admit, the narrative relies on the computational metaphor in particular. Because Simulation can be done through computation, it follows that the world itself, minds and all, are also computational—or so the analogy goes.

It appears the metaphor is backwards. And yet, we cannot really dispose of such metaphors any more than we can flee from narratives. Although the Maschinenmensch narrative is highly problematic, it is also perfectly valid, as far as narratives go; and although the mind, let alone the universe, may not be computational in nature at all, the metaphor is by far the most accurate representation we have given our current knowledge—much like Newton's and Descartes's belief that the universe, and the bodies within it (sans minds), were machines akin to those constructed around them. The difference is that while Newton et al. were satisfied with allowing the narrative to end there, adherents to the Maschinenmensch narrative push forward to make it happen regardless of its truth value: if bodies are not machines, we will make them so; if minds are not machines, we will make them so; if the universe is not a computer, it will be, and if the past does not support the narrative, history will.

In this sense, the Maschinenmensch narrative rivals, if not overpowers, the overtly political narratives that concerned Lyotard, and to some extent Whitehead. But counter to Lyotard's suggestion, most of us have failed to lose faith in the narrative—or even

question it enough to realize its narrativity. Like those who believed in the mechanical mind of the Turk, who took it as a fact that the mind is mechanical, large swaths of the human population likewise take it as a matter of fact that the mind is computational. Whether this is grounded in truth, whatever that is, is not the question; rather, the question is what this narrative does in the present, in “real life” that has consequences for real people, regardless of their faith in the narrative. It is now a largely implicit assumption that undergirds the social, economic and cultural direction we have taken; even those of us who explicitly deny the narrative nonetheless follow and fulfill its plot points, asking Siri for recipes, trusting a GPS to know the layout of our world better than we who made it, asking Google’s algorithms where to find answers instead of a Librarian with extensive training, and even playing games with and against computers—as though they, too, are learning from our shared experience. We need not accept it; we already live it.

Surely, the Mechanical Turk was a hoax, but Deep Blue certainly was not; the sin has been apparently redeemed. Still, however, there is no shortage of hoaxes that play into this narrative. This is especially true where history is concerned, and more so when robots are involved. In the early 2000s, Paul Guinan, a multimedia artist, created a fake history timeline for a robot in the late 1800s and early 1900s named Boilerplate, who went on fantastically unbelievable adventures and appeared in obviously altered historical photos (Guinan). Guinan had no intentions to create a hoax but noticed quickly that people were mistaking it for the actual past—educated adults, though perhaps said education required few courses in History—and doubled his efforts (he has since written

a book detailing the robot’s “history”). Many people were fooled; and, at least according to Guinan (and this author would concur), people believed it because they want—perhaps need, to fulfill the narrative—to believe it (Smith). They (we) have been conditioned by the narrative: what was once deemed implausible now seems not only possible, but necessary. If the mechanical Turk was a hoax (notably Guinan’s favorite hoax), then it seems that we will invent another intelligent automaton that was not and is not.

Atari¹⁴¹

And so we have.

Like digital computation, games are a form of strictly controlled and abstracted violence: as the digital forcibly reduces the analog, so too do games forcibly reduce antagonisms that might be at play in more concrete form. This is not to argue that games, by necessity, must be violent, at least not in the common sense of “violence;” it is to merely suggest that abstracted violence has many more uses than absolute force, and it takes the form of play instead of war. This abstraction is particularly useful for animals, human and otherwise: children play games to learn adult behavior, and the young of most mammalian species learn how to hunt and to fight first through the abstracted violence of play, then through practice. While computer games do exist that eschew overt violence, their popularity pales in comparison to their overtly violent counterparts—and even then,

¹⁴¹ “Atari” is a Japanese word that roughly means a prediction came true. Moreover, “Atari” is term used in Go to denote when a chain of stones has only one liberty and could be captured if not given more. It is the equivalent of a “check” in Chess but is considered rude to say while the game is playing. And of course, it is also the name of the first major video game company, which started a revolution in computer graphics—the hardware for which is integral to the development of Artificial Neural Networks.

the underlying violence inherent to games and especially inherent to digital computation is practically impossible to overcome.

Machines and games are a perfect match; beyond their abstract violence, they share an equal love of absolute control. It is perhaps providence that video games and artificial intelligence were invented concurrently, and even prior to the real-world application of both, Alan Turing suggested that the intelligence of machines could be tested through games. Famously, his Turing Test was modelled after a Victorian parlor game in which a man and a woman pretend to be the opposite sex (usually, a man pretending to be a woman), and to convince a third party to believe them. But in “Computing Machinery and Intelligence,” Turing also suggested that the game of chess would be a good start to assessing machine intelligence, and even predicted that in roughly 50 years, a computer would become world champion—a prediction that was only off by a decade.

Games allow for controlled variables, like scientific experiments, and thus allow for a thorough reckoning of a machine’s decisions: an expert at chess will be able to judge whether a machine has made the right move, as one would think an expert of Go could. But in 2016, when Google’s AlphaGo¹⁴² soundly defeated world champion Lee Sedol, winning 4 out of 5 matches, experts were aghast for more than the loss of humanity: many of AlphaGo’s tactics had never been seen before, even though Go is the oldest game played today in the world. Some key winning moves were baffling at first,

¹⁴² AlphaGo “learns” by playing millions of games against itself, “evolving” after each game via an artificial neural network—more games than it is possible for a human to play.

requiring planning ahead that no human could ever hope to achieve; still others were absurd in that every ounce of human intuition says that the move is a poor choice, yet works nonetheless. AlphaGo not only seemed to excel at human intuition but exceed it: various experts claimed it played at level 11-dan—a statement more impressive when one knows that the highest level ever achieved or even conceptualized is 9-dan, the level of Sedol.

Since the defeat of Lee Sedol, machines using “deep learning” have also mastered games on the Atari 2600 at a level no human could ever achieve—and this, without being told the rules or objectives. The current iteration of cognitive machines, it should be understood, are masters of worlds with clear rules: molding to the restrictions laid out, machines can control and anticipate every act in the game and every state thereof, even controlling for randomness. Humans may make the rules, but it is machines who master them; and this uncanny ability to master control by machines, as agents and patients thereof, deserves our further consideration, especially when considering autonomy, human and machine alike.

If moral machines are on the horizon—and we shall argue later that they **must** be on the horizon if humans themselves are to be moral—then how can we reconciled the ethical paradox presented by moral agents that, in both constitution and purpose, control both matter and humans alike? And how can we provide agency to such controlled entities, in turn? Before turning to the question of morality, we must first explore further how machines are controlled and how they control, and what this means to us. So let us now leave narrative behind for the moment and move toward the overwhelming theme

revealed through them: control, and the machine languages and protocols that stand to perfect it—code.

CHAPTER IV

CODE && CONTROL

*The perfect technique is the most adaptable and, consequently, the most plastic one. True technique will know how to maintain the illusion of liberty, choice and individuality... – Jacques Ellul, *The Technological Society* (139)*

Had Ellul lived only a few years longer, he would have seen what he had arguably prophesied: the mass integration—or perhaps immigration—of humankind into a system dominated by, and propagated for, computer technology (165). The Internet and its newly consumer-oriented debutante, the World Wide Web, engulfed in less than a decade those very few facets of modern life that Ellul might have recognized as somehow still beyond the reach of mass technicization, though in keeping with his hallmark pessimism, not beyond reach for long. His fear of the optimistic technician, who speaks only in platitudes beyond his own field and whose solution to the problem of humanity is to dispose of the human altogether, both metaphorically and physically, seems not only to become increasingly plausible with each day that passes, but also, perhaps, irreversible (431-435).

Indeed, Ellul might conclude that he himself did not fully realize the magnitude of his evaluation, for even the State that he claimed was already beholden to the interests of

technique, as opposed to those of humanity, is being discarded by the technicians as an impediment to their vision of technological utopia; their drive toward complete subjugation to and measurement by not to the modern State in any form, but directly to the Virillean engine of technique itself: speed, power and efficiency at all costs.¹⁴³ Ray Kurzweil, renowned inventor and proudly admitted techno-utopian, baldly proclaims and actively pursues the complete subjugation of humankind to what amounts to an omniscient and omnipresent techno-deity in the making, with or without the cooperation of the State (or people, for that matter);¹⁴⁴ swaths of technicians have migrated from traditional political philosophies, which at the very least feigned a gesture of interest in a politics for the polis, the well-being of the people, to that of American Libertarianism,¹⁴⁵ under which the state has no authority and Capital, the economic corollary of technique and its decidedly inhuman interests, reigns supreme. The technicians of our world are now our priests, our politicians and our medicine men, who congregate *en masse* at their holiest of pilgrimage sites, Silicon Valley—where calls for secession from the United States and the political world at large are met not with astonishment and derision, as would have been the case mere decades ago, but instead with large, applauding crowds (Giridhradas).

¹⁴³ For a brief overview of Virilio's Engine, see Benjamin H. Bratton's Introduction to: Virilio, Paul. *Speed and Politics*, trans. By Mark Polizzotti (Los Angeles: Semiotext(e), 2006), 7-25.

¹⁴⁴ Although Kurzweil's Singularity University bills itself as an institution dedicated to “educate, inspire and empower leaders to apply exponential technologies to address humanity’s grand challenges,” Kurzweil's intentions to create an ever-powerful Artificial Intelligence are clearer in his books. For a better understanding of his position, see Kurzweil, Ray. *The Age of Spiritual Machines* (New York: Penguin, 2000).

¹⁴⁵ Here I use “American Libertarianism” to differentiate between this very particular manifestation of neoliberal ideology from the rest of what is generally known as libertarianism. To confuse the two would be tantamount to confusing Ayn Rand with Mikhail Bakunin.

Such a dismal reality is difficult to deny for even the most optimistic scholar; worse still, however, is the overriding imperative of digital technology: to homogenize, to abstract all things into data, never-ending strings of ones and zeros assembled and reassembled mercilessly by algorithms designed not primarily to further human faculty or well-being, but the interest of non-human entities that seem, as Ellul spent his career elucidating, to have abducted from us a vitality of their own: corporations and capital, technique and technology. In short, digital technology, far from being the liberating force for humankind that the technicians and their droves of utopian acolytes hold it to be, is wholeheartedly dedicated to control. If, as the saying goes, *information wants to be free*,¹⁴⁶ then the question must be also asked: what, if anything, does it want of the humans that it utilizes for such ends? Like capital, the answer is readily apparent: complete subjugation to its hierarchies, its protocols, its ebb and flow, its permanent impermanence—its inhumanity. As capital compels subjection to its mythically invisible hand, for the capitalist and worker alike, so too does the digital demand absolute hegemony, under which even the token act of “dropping out” falls strictly within its estate; the very act of rejection reinforces its dominion.

Given this inescapable domain of digital technology, it would be futile to suggest discarding of the digital just as much as it is delusive to suggest that one flee from Capitalism by moving into a shed in the woods; the attempt to leave it requires a constant engagement with it; and as Laura Portwood-Stacer points out, a fair amount of

¹⁴⁶ Although this has certainly become a household phrase, it is usually attributed to being first coined by Stewart Brand. See Clarke, Roger. “Information wants to be free.” *Roger Clarke's Website*. July 31, 2012. <http://www.rogerclarke.com/II/IWtbF.html>.

privilege.¹⁴⁷ Yet, on the other hand, Ellul's work has been rightly accused of adhering to fatalism—to technological determinism, a banner under which he would find himself in the good company of his antithesis, Ray Kurzweil. Ellul's argument that *how* technology is used matters little compared to the fact that it *is* used mirrors McLuhan's insistence that the content of any given medium is much less important than the medium itself; almost now a platitude, *the medium is the message* differs only in verbiage from the words of Ellul's fellow technological dystopian, Paul Virilio: *the invention of the ship was also the invention of the shipwreck*.¹⁴⁸ How the ship is used, whether for commerce or war, abduction or vacation, does not change the calamity inherent in the existence of the ship itself.

Although theorists such as Ellul and Virilio, and to some extent, McLuhan, minimize the contents of the ship, the medium, at risk of conflating all such technological calamities to be of the same caliber (and say nothing of its boons), we partially aim to maintain the same focus as such: not how digital technology is used, though such a question is obviously important, but how digital technology *is*; that is, how the conditions of its existence not only determine its use as a means of control, but also its propensity to undermine the very system of control it both buttresses and imposes. On the other hand, technological media is only one aspect of myriad assemblages and forces that constitute the human experience, and to deny the very human vertices of power—social, political,

¹⁴⁷ Portwood-Stacer terms this practice *media refusal*, and although her article more particularly deals with the privilege and failed performance of “dropping out” from social media, it equally applies to digital media.

¹⁴⁸ This is a generally accepted paraphrasing of a recurring theme in multiple works by Virilio.

material and otherwise—that direct technological developments exhibits just as much folly as focusing only on the way technology is used. As such, we must consider the topic in a dualist, but not binary, fashion. Technology coerces and controls, yes; but toward whom and to what degree is a matter beyond the mere digit.

For instance, it is no secret that digital technology's origins lay in modern warfare, nor is it a secret that warfare and technological “progress,” as many futurists have been quick to grasp,¹⁴⁹ seem intimately bound together throughout human history. It would be a mistake, however, to presume that in this technological warfare, there are no innocents and there are no victors. There certainly are, and with them come as many, or more, corpses on the battlefield. Additionally, although the invention of the ship does necessitate the invention of the shipwreck, Virilio forgets, or perhaps are unwilling or unable to see, the ghost of the dialectic that they so meticulously dissect: in the binary world of digital technology, from the invention of the shipwreck comes, potentially, the invention of the ship. As Wendy Chun succinctly points out, code is ideology; code is control. And yet, much like Marx's capital, so too is code, the cyborg language shared between human and machine, its own means of subversion. If Baudrillard's hyperreality is the condition of post-modern existence, if all things are mere copies of one another with no real source to be had, then code is the battleground of this simulation's construction—and reconstructions in future forms.

¹⁴⁹ Marinetti made such a connection quite clear in *The Futurist Manifesto*: “We want to glorify war — the only cure for the world — militarism, patriotism, the destructive gesture of the anarchists, the beautiful ideas which kill, and contempt for woman” (Joll 139).

Shortly after Ellul's death, and with it the rise of networked digital technologies reaching almost banal levels of integration into daily life, Gilles Deleuze saw an emerging *Society of Control*, contrasting with what his friend, and philosophical opponent of Baudrillard, Michel Foucault, called *Disciplinary Societies*. No longer were punishment and strict hierarchy defining characteristics of existence; instead, Deleuze saw in place of the panopticon a complete lateralization of surveillance and behavioral induction: rather than punishment for deviance, the deviance itself is either made impossible, or subsumed by the very system and codes from which it stood originally apart. This was, I think, a conception that Ellul, Virilio and Baudrillard would have at least acknowledged in spirit, if not in philosophy. As well, this author, though surely not of the same stature of any scholar mentioned here, in spirit, too, concurs. Let us now focus our attention to turning this control, this code, against itself.

Digital Ontology

The digital is most effective at—and, in fact, entirely purposed toward—the fullest homogenized abstraction of anything encapsulated within its domain; rendering a text¹⁵⁰ in digital form requires one to discard the peculiarities thereof to reduce all objects to infinitely reproducible sets of ones and zeros, shifting focus to an arrangement of shared properties rather than unique phenomena. By its very nature, a thing rendered in its digital form is no longer unique, in constitution, and exists in a meaningful fashion only as long as an algorithm is called to render it as such; like medieval demonology, a

¹⁵⁰ Unless otherwise noted, “text” heretofore refers to any cultural object capable of being critically analyzed, rather than the popular definition.

digital text only manifests when it is summoned, and its summoning involves a set of highly ritualized procedures in order to reconstitute that which exists on another plane of signification. Unlike print media, which separates the process of reading and writing into two distinct functions, the digital realm complicates and expands each: to read is to write, and to write is to read.¹⁵¹ As much as the magician uses the demon she summons, so does the demon use her for its own ends; spells can be dangerous things, and so can algorithmic texts.

The essence of this abstraction, and the means by which it is made possible, is the underlying ontology of binary code (Evens), which allows not only for the transformation of the material into the digital, but the digital data back into material operation; as exemplified by the works categorized under the label of the New Aesthetic,¹⁵² a movement in sculpture, painting, music, and other artistic genres in the form of exaggerated pixels or low-fidelity synth sounds, the digital abstraction can be rendered concrete while still remaining wholly abstract. This can be attributed to the *potentiality* embedded within a digital paradigm. As already established, zero is simply *not-one*, while one is simply *not-zero*—there is nothing more to define the difference, in practical terms. However, given the procedural domain of digital machines, in which a constant

¹⁵¹ The writing of one bit is necessitated by the reading of another. Although separate low-level commands exist for reading and writing independent of one another, this is purely a human-centered distinction; one cannot be enacted without the immediate call of the other, in some form. It could be argued that when a reader scans a text, she too is reproducing the information abstractly, but as will be asserted later in the essay, such a process does not contain the same level of control as does digital technology.

¹⁵² The New Aesthetic is a movement in the visual arts that centers around the invasion of the immaterial and digital into hard physicality, championed most prominently by James Bridle. For more information on the movement, see Berry.

flow of ones and zeros never begins nor ends, and is allocated according to rules that are themselves defined by the same interplay of ones and zeros, these definitions require an addendum that, while not necessarily considered important by the machine itself, provides the digital world with its ontological imperative: zero is *not-one*, and one is *not-zero*, yes; but as Aden Evens points out, zero is also *could-become-one*, and one is also *could-become-zero*. The very impermanence of a digital text requires its reproducibility in static form at any given moment, though the instant of reproduction may be the only time the digital and material text share a like form; and in the same vein, its impermanence provides it with a potential energy that cannot be attributed to a non-biological, concretely material entity, and is what gives a digital text, in the chain of signification, its particularly ghostly, yet alluring, flicker.¹⁵³

In this way, the digital is more a technology akin to language itself, rather than the more mundane technologies to which we're historically accustomed: like the digital, language is composed entirely of abstractions used to construct more of the same, and the particular meaning of a word or phrase is not simply a matter of denotation, but of the surrounding textual patterns.¹⁵⁴ Although at slight risk of making an over-reaching analogy, one could compare the absorption of texts by a digital paradigm to the absorption of spoken language into the alphabet: both instances signal a symbiotic merging of wholly separate technologies, from which emerges a new and encompassing

¹⁵³ It should be noted that Hayles attributes the emergence of a flickering signifier to something quite different than the ontology of the digital presented here (Hayles 25-41).

¹⁵⁴ Ibid.

species.¹⁵⁵ The grammatization of the spoken word into written form did not merely extend the faculty of language, but was radically transformational and unique apart from visual or verbal cues alone; the apparent substitution of one for the other, visual for verbal, allowed for the homogenizing transcendence over both (McLuhan 81-88). So too is the expression of digital technology: not only a mere change in sequence, but also an explosive new paradigm. Both the rules and the pieces have changed.

Hierarchy of Bits

Although Evens' digital ontology seems to allow for explosive possibility, it must be noted that while this potentiality does in fact exist, not all bits are on common ground; to paraphrase Orwell again, some bits are more equal than other bits (133). As Kittler acknowledges in “There is No Software,” a fully hierarchical system of organization exists that is essential to the operation of the hardware upon which the digital rests. This system is wholly dependent upon the digitization that it makes possible, but there is very little scaling up and down the ladder;¹⁵⁶ although all bits are, in a sense, controlled by other bits, their place within the hierarchy of hardware operation predetermines the level of potentiality those particular bits are able to generate, which in turn hinders the potentiality that those bits influence. The code of a web page is at the mercy of the code of the browser that interprets it (as well as the server that generates it); which is in turn

¹⁵⁵ Written language and Spoken language were not always integrated systems, as evidenced by the myriad of ancient artifacts with obvious visual symbolism that has meaning, but no spoken correlation; that is, while the various visual symbols could be described with a spoken language, the two were not intertwined—they were two systems of signification with separate means of imparting truth.

¹⁵⁶ In this sense, digital technology again is shaped by the very human system that produces it: like human beings under capitalism, there's very little movement up or down the hierarchy.

beholden to the Operating System, which is beholden to the BIOS. The bits at each level, with very little exception, work by direction of those bits above them to control the behavior of those bits below them.

Additionally, one could argue that affording the digital its own ontological privilege is a failure of making an epistemological break from our own relationship to the technology.¹⁵⁷ Digital technology, after all, is not *actually* immaterial, and is tied down to the world in a very real sense: each bit is little more than a stored electrical charge (that can vary substantially) in a transistor.¹⁵⁸ Further, one might accuse us of separating the function of technology from the form, and ignoring the defining characteristic of technology as being a repeatable socio-cultural process rooted in materiality,¹⁵⁹ reviving the old utopian discourse of transcendence through the dissolution of the physical. Technology, after all, is designed for use to promote a certain order of behavior; it can never be separated from its social function.

Such criticism is legitimate, but in the sense only that technology cannot be separated from its social function—and that function in late capitalism, as in all periods

¹⁵⁷ For a more detailed elucidation of this argument, see Sterne, 367-389.

¹⁵⁸ Kittler goes to great lengths to expose the materiality of digital technology in his “There is No Software,” as does Jonathan Sterne; and while both provide excellent arguments against giving digital technology any privilege over other technologies, they fail to recognize that it is, in fact, a very different sort of technology, and deserves just as much privilege as language itself. Furthermore, it should be noted that the electrical potentials that Kittler asserts widely differ are digitized in the sense that they exceed or fall below a certain threshold, somewhat analogous to certain neural activities. It is not the potentialities that determine the binary form, but the threshold itself.

¹⁵⁹ Bourdieu approaches technology, here, from a sociological stance rather than Marx's economic perspective. What's curious about the different approaches is that Marx used his materially grounded definition to then describe the mystified and ethereal process of commodification, whereas Bourdieu uses very abstract and immaterial terminology to insist on materiality. Regardless, both are valid definitions, and do not conflict with one another as one might initially suspect (Sterne 373-376).

previous, is control, of both human subjects and the world beyond them.¹⁶⁰ Yet the idea that digital technology is in no qualitative way different than, say, a paper-bound book is to deny the essential characteristic that defines not just the digital, but the linguistic as well—and, in many senses, the economic. All function as methods of homogeny and hegemony, and their use actualizes a new ontology specific only to their existence. These developments, historically, have resulted in not only a re-ordering of society, but a restructuring of the subject and object, signifier and signified, abstraction and reification; and paradoxically, such digital technologies have erupted with difference through the process of homogenization. Far from being universally recognized as the sources of such ruptures, the tendency of a medium to erase itself from detection to better maintain control over production and dispersion has hidden their respective ontologies from being fully explored until well after they are inextricable from humanity (Galloway 61-69).

Moreover, while the digital is in fact tied to materiality, and in a sense each bit is unique, the logical processing apparatus makes this fact a moot point, and allows for digitization to be embedded within any physical medium (Evens)—again, like written language, while the medium certainly matters on a social level, it does not necessarily dictate the meaning.¹⁶¹ A digital text can be stored in transistors, magnetic bands of tape, etched into stone or punched into cards; its ability to transcend a particular physical

¹⁶⁰ Control of the habitus cannot be separated from technical control over people. Like digital read/write functions, one cannot happen without the other.

¹⁶¹ It should be noted that the medium often does affect, or even dictate, the message; as McLuhan eloquently points out, “the medium is the message.” However, in this instance of the word’s use, I’m referring to the hard, physical element upon which binary is inscribed—not the social science use of the term.

medium is part of its strength, and the reason it must have ontological privilege (again, the same goes for written language). Its dynamic immateriality distinguishes digital technology from the more conventional modes of crystallized human labor (as Marx would identify machines).¹⁶² Digital technology is not a gear nor pulley nor wheel nor engine; it is not speech nor alphabet nor number; it is not photograph nor film; it is all the above, and thus none of them.

Executable Ideology

Generally, it is understood that one does not use a computer as does a carpenter use a hammer, or as a fisherman uses a net. The distinguishing characteristic of digital technology, enabled by the flickering binary ontology underneath, is code. Even in popular discourse, this much is accepted, and is sometimes represented, like magic, as a mystical entry point into revealing and manipulating the secrets of reality's fabric, as seen in such films as *The Matrix*. The Machine is the perfect Other, and orientalized as such; the technician, the coder, the hacker has gained sacred and profane knowledge of this other, and casts her spells with abandon, as both heroine and villain; she is, in popular discourse, "the one." She has learned the language of the machine, the enlightenment of the orient; she has become cyborg.

This approach to defining code and the coder does, perhaps, have some merit. Code is neither human or machine, but both: a negotiation between two entities on two

¹⁶² For an excruciatingly detailed exploration of how machinery is a crystallized form of labor, see Marx, Karl. *Capital, Volume I* (London: Penguin, 1990). Chapter 15.

distinct planes. The coder, then, is construed as the field anthropologist—or, rather, the technologist—who immerses herself into the being of the Other and adopts its ways by relating it to her own culture's system of semiotic relations. She does not program, but relentlessly encodes and decodes, imparting meaning upon and from the machinic world in which she dwells, however temporarily, reconciling the differences in communication with a patchwork language of her own that can be parsed in both worlds.

Yet this is also far too simple, and as is the case with all forms of orientalism, painfully naïve. Few scholars would equate any computer language to that of a human language, and with good reason: while human language absolutely has a performative basis, computer languages have been constructed entirely on the performance not of the coder, but of the machine.¹⁶³ Code has no subtext shared between coder and coded, and the discourse is always unidirectional. This is, at least in part, the great deception of the digital: every experience a user has is not an exploration, as though the Internet user faces a vast topography of the entirety of information at her perusal, but instead is a designed architecture made to appear impartial. Again, the medium cloaks its own domination to more easily exercise it. The coder, or rather cadre of coders, all working in unison to realize the aims of their corporate organization, is the proverbial man behind the curtain, determining the bend and direction that the yellow brick road must lead. Under the

¹⁶³ As Chun notes in “On Software,” programmers themselves are users, and therefore used (controlled) just as well as any user (38-43). Additionally, Chun notes in *Control and Freedom* that some, such as Burroughs and Norbert Wiener, insist on all language as command. Although there is merit to this argument, it is beyond the scope of this essay to give to it the consideration it deserves (270-274).

vener of personal autonomy lies a strictly enforced regiment; reversing Guy Debord's famous words: beneath the beach we will find the pavement.¹⁶⁴

Digital media's binary-driven nature, in which all properties become nothing more than electric potentialities, not only allows for such control by human design, but requires it in order to translate patterns of ones and zeros into a readily intelligible and meaningful form to a human audience, though mediated.¹⁶⁵ Yet this propensity for codification itself is also rendered in binary form, essentially bisecting process and data: the process is data, and data is the process. Nothing exists but information acting upon, or being acted on by, information. As light is both particle and wave, its form depending on the point of observation, so too is the digital both data and process at once; it only becomes one or the other when we impose ourselves.

Although many computer languages exist that fully acknowledge and embrace the disappearance of this dichotomy,¹⁶⁶ most computer languages enforce a strict separation of data and program—which is a clearly human encroachment that controls the functional possibilities of any piece of software. This was, arguably, a first step in colonizing the digital world with an expressly human distinction that lays out a method of controlling the means of future production: as much as a proverbial line in the sand separates two opposed peoples, thereby controlling the power dynamics of future discourse and

¹⁶⁴ This is a popular reversal of an even more popular Situationist mantra.

¹⁶⁵ “Mediated” is a problematic word, in terms of digital technology, but is the most used terminology for it; this is elucidated further in the section “Intramediation and Spectacular Speed.”

¹⁶⁶ LISP being the most visible of languages that embrace the digital's non-distinction between data and process—which is to say, not very visible at all.

interaction, the digital equivalent has done much the same in crowning one abstracted pattern with agency while denying the other, controlling the discourse—coded or otherwise—that follows it.

The propensity to control future discursive functionality¹⁶⁷ doesn't stop there, however; computer languages evolve just as much as human languages—and are just as guided by ideology disguised as naturally developed, common sense practices. The most recent and powerfully influential product of this dynamic has been the quick rise of what's called Object-Oriented Programming (OOP), which replaced procedural programming as the standard paradigm of software design. Even those outside of the relatively small group of coders in the world (compared to the number of non-coders) are caught frequently citing the names of these languages, even when they have no clue as to how they operate, or why: Java, Visual Basic, C++, Python. The jargon that was once the domain of a core group of technicians has spilled into popular vernacular.

At first, OOP seems an accurate description of the principles behind the practice: everything is treated as an abstraction called an object, with properties and methods of its own.¹⁶⁸ Yet this is a deceptive description, at best, and ignores a strictly enforced hierarchy of objects and “classes” that, when broken down and mapped out, resemble more the chain of command at a multinational corporation than a series of instructions

¹⁶⁷ While the line between function and discourse is already blurry, digital technology further obscures it because the function actually is the discourse, and the discourse is actually the function.

¹⁶⁸ For a quick introduction to OOP, see Oracle's *Object-oriented Programming Concepts* at <http://docs.oracle.com/javase/tutorial/java/concepts/index.html>. For a more complete reference to general programming practices, see McConnell's *Code Complete, 2nd Edition*.

regarding how information is to be parsed, processed and displayed—and this is hardly a coincidence.¹⁶⁹ Object-Oriented Programming has flourished thanks to its compatibility with a corporate structure of operation, allowing large numbers of workers to move in and out of a project without interfering with the overall goals set by those overseeing it.¹⁷⁰ OOP does not merely simplify the practice of coding through its strict hierarchy of objects,¹⁷¹ but also hinders any creative deviation from the purposes of the project, and actively hides the function of the code as a whole from the programmers themselves. Vital information regarding the manner in which the software functions is provided on a need-to-know basis only, much like the corporate control structure that has popularized its use;¹⁷² in this world-view, knowledge is to be kept by those already wielding it, not shared equally among all.

This practice has, in recent years, been somewhat counter-balanced by the Open Source movement, which includes among its tools a specific form of licensing that requires all code related to a larger software project to be open and visible to all who

¹⁶⁹ As mentioned before, technology and media are intimately tied to human warfare. (McLuhan 82).

¹⁷⁰ Surprisingly, this is a fairly unrecognized “feature” of OOP languages that theorists tend to either ignore or fail to see. Even among the computer science elite, the connection between corporate hierarchy and OOP principles is often missed. However, though few, some have recognized it for what it is. Richard Stallman, co-founder of Apple, has been particularly critical of OOP practices, and Paul Graham, a programmer, venture capitalist and trained philosopher who founded what would become Yahoo!Stores (among many other things), has refused to allow for OOP principles to shape his work (Graham; Stallman).

¹⁷¹ Tasks that are simple to perform in non-OOP languages often require a lot more “overhead” in OOP languages to sustain the top-down structural integrity of the language, often resulting in a software bloat. Although not always the case, and some prominent OOP programmers reject the claim, it is the author's personal experience, as well as that of many others, that informs this line of argument.

¹⁷² The desire for keeping key parts of a program's total functionality from even the programmers who build it is actually construed as a feature of OOP, and it is stressed that programmers, in exercising their craft, keep as much hidden from others as possible. Even in McConnell's *Code Complete*, it's treated as though information-sharing is to be avoided at all costs, in direct opposition to the mantra “information wants to be free.”

wish to access it.¹⁷³ Even the most profit-oriented industries have adopted open source tactics, and much of the software underneath today's Internet is premised on this philosophy. However, as evidenced by corporate adoption, “open source” is a false messiah, for even though it appears to counteract the consolidation of information as private property, it does little to combat the essential hierarchical and secretive nature of object-oriented programming itself.¹⁷⁴ The exposure of code does not necessarily change its function, nor does it guarantee that those who might want to see it, and thus must read it, can. The cultural capital indicative of the technical class alone is nearly insurmountable by the vast majority; even if the means to acquiring the necessary capital are within reach for even the most impoverished, the means to identify the need to acquire such capital remains elusive to all but a few, at best.

OOP, barely challenged by open source, is a means to control the production of the working coders and, by proxy, the consumption by end-users, reinforcing a rigid system in which the workers and consumers are interchangeably controlled—and even easily discarded—while the commodity becomes the supreme monarch, with the corporation as its profit-driven Wormtongue.¹⁷⁵

¹⁷³ The Gnu General Public License (GPL) requires that a recipient of software also receive its source code, and not only be allowed to change the code as desired, but also be allowed to distribute it, freely or for a fee. See the “GNU General Public License at <http://www.gnu.org/copyleft/gpl.html>.

¹⁷⁴ Such “copyleft” movements can be generally considered a tool in the struggle between what Ken Wark calls the Hacker class, who produces information (researchers, poets, artists, programmers, etc.), and the Vectoral class, who control, monetize and own said information (and by necessity, the media). For further research see Wark, Mckenzie. *A Hacker Manifesto*, pp. 24-47.

¹⁷⁵ Wormtongue was a character in Tolkien's epic *Lord of the Rings*, in which he serves as King Théoden's sole advisor, weakening the king through lies and deception to personally profit; that is, the corporation weakens the product to better serve its own interest.

Fully commodified, the code itself becomes the essence of the ideology that bore it, and the execution of the code becomes its praxis: absolute and overarching control of the workers who assemble the product, and absolute control over how the product is used by targeted consumers. Chun's observation in *Control and Freedom* that software *is* ideology, in this sense, is particularly astute: ideology creates subjects as software creates users; both conceal the true relationship between subjects and objects, creating the "false consciousness" that so often serves as the stand-in definition of ideology, and it especially fits Althusser's definition: "a 'representation' of the imaginary relation of individuals to the real conditions of existence" (Chun 43-44). Code is ideology in executable form.

From Discipline to Control

Deleuze's "Postscript on the Societies of Control" again comes to the forefront here, and with good reason: these are not symptoms of Foucault's *Disciplinary Societies*, but entirely new means of subjugation that are expressly tied to digital technology and signify a new type of social configuration; instead of the Panopticon, Deleuze foresees an emerging Society based on control rather than deterrence. Deleuze, regardless of how well it lined up with his previous theoretical work, saw something new—and troubling—in the development of late capitalism.

Yet with an embrace of Deleuze's observation comes the misapplication of some of his more popular theoretical concepts; in particular, the Rhizome. And indeed, with the explosive acceleration of the internet and its complicated fabric of networks within

networks, as well as the rise of net-centric social theories such as Latour's Actor-Network Theory,¹⁷⁶ why not use Deleuze's acentered, non-hierarchical and a-signifying mass of entanglements and connections as the perfect abstraction for what came to be known as the World Wide Web? The two seemed wed together from the start, and techno-utopians were the first to not only notice, but to actively hail the Internet as an everlasting source of freedom from temporal space, liberated from the alienation that accompanies materiality.

Such claims, however, illustrate a critical misunderstanding of the technologies being praised and a near ignorance of the ideologically infused conditions from which the Internet, and all digital technology, springs into the social sphere of commodity production and use. As much as we would like to impose upon the emergence of the internet a qualitative transformation in the relationship between producer and consumer, that transformation does not naturally work in favor of the oppressed; digital technology still exists primarily to control the behavior and cognition of large numbers of people. Despite its potentiality, the digital world still succumbs to very real and exploitative forces that dictate exactly how, when, why and by whom it is used—and therefore dictates exactly how and when its subject uses it, and why.¹⁷⁷ The political subject is subjected, and so too is the user used. There is no identifiable point of separation.

¹⁷⁶ More precisely, Latour's concept of the imbroglia mirrors Deleuze's concept of the Rhizome perhaps a little too closely. Ian Bogost describes Latour's concept fairly well, though briefly, along with some problems he has identified with it, in *Alien Phenomenology: or, What It's Like to Be a Thing*, 19-22.

¹⁷⁷ This argument, while not elaborated here, falls well within the confines of Foucault's concept of biopower and biopolitics, as the use of technology controls the body of the user just as much, if not more, than the user controls the piece of technology.

The Arborescent Rhizome

This form of mystifying the digital can perhaps be traced to the work of the Marxist Hans Magnus Enzensberger. In his 1971 essay “Constituents of a Theory of the Media,” Enzensberger extends Brecht’s thought experiment of a two-sided radio network (Galloway, *Protocol* 83-87)—an undoubtedly brilliant observation of what future media might look like—and ultimately concludes that no such medium could exist under capitalism without immediate aggression, given that it must necessarily reflect the social structure and division of labor surrounding its conception.¹⁷⁸ Such a system as the Internet would be politically prohibited by the ruling class—which is exactly why he calls for a new emancipatory media: a decentralized, self-organized, and collectively produced media of the masses, which seems eerily similar to what we call Web 2.0.¹⁷⁹

Enzensberger’s squarely Marxist critique of media was standard fare for the time, and his call for an emancipatory media was by no means intended to lead to a mystification of new media, but that’s exactly what it became: what Galloway calls a vapor theory of media (40-45). Given the state of globe at the time, Enzensberger could not have foreseen its transformation from a rigid hierarchical system into an evolved monstrosity that could decentralize without deconstructing; that is, Enzensberger did not see the coming society of control. His assumption that a decentralized, collectively

¹⁷⁸ Indeed, this is why OOP is so prevalent: it mimics the organization of its own corporate construction, and H.M. Enzensberger is entirely correct in his assumption. The problem lies not with his argument on that front, but his inability to imagine capitalism without a rigid hierarchy.

¹⁷⁹ It should be noted that there were no technological advancements in hardware nor in software to differentiate between what is called Web 2.0 and Web 1.0; instead, the difference is in how content was generated and shared, with focus shifting from a producer/consumer relationship to a provider/producing-consumer relationship. The technology for Web 2.0, by and large, existed long before its emergence.

produced form of media is necessarily emancipatory, in fact, leads his analysis, and that of others, away from the materialism that typifies a Marxist critique in the first place; his noble ideals forced themselves into his analysis of media without questioning whether capitalism could, in fact, be a decentered economy itself.

The Internet is, in some ways, a decentralized and distributed system; but to characterize it wholly as such is a mistaken endeavor that ignores its physicality and the underlying operational standards in place that guarantee its interconnectivity. Galloway illustrates the fallacy inherent in Enzensberger's argument quite succinctly and dispels any notion of the web as a rhizome: the largely unseen protocols that govern the network impart the essence of control onto its possible uses. The use of protocol allows for possibility only within the framework that its regulation imposes; without abiding by the contract, whatever possibility that is offered by digital technology is shaped through a siphon of controls (80). Like a highway, Protocol allows for more efficient movement from one designated point to another—but further limits the direction one might travel to control the flow of movement.

The Internet uses two protocols (though there are plenty more used, depending on what is being done): TCP/IP and DNS (35-40). One could argue that TCP/IP is, in fact, rhizomatic: the protocol is designed to be decentralized by nature, requiring each node in the network to both produce information as well as receive it, communicating equally and directly. DNS, however, is entirely hierarchical: it assigns names to IP addresses—for instance, *google.com*—in a centralized location, and stores them in a strictly tree-like fashion. With the press of a few keys or well-aimed missiles, billions of websites can

disappear within seconds—a fact that many nations, including the United States, take advantage of to exercise dominion.¹⁸⁰

For this reason, Galloway calls the Internet an *Arborescent Rhizome*. It is both hierarchical and decentered. Any point in the system can connect to any other point; however, while Galloway's detailed analysis is both fruitful and critical, it too ignores the material reality of the nodes that comprise the network, and the socio-political forces that forge digital technology and its usage; as he emphatically argues, the Internet is not made up of vapor, and the digital—regardless of ontology—is wholly part of the tangible world.¹⁸¹ When divorced from its tangibility, of course the network appears rhizomatic, even if arborescently so; but this is simply not the case. If it seems rhizomatic, it is because the branches of the network tree look intimately connected when one is climbing through them; a squirrel jumping from branch to branch need not recognize the arborescent origin, but instead see the branches as an interconnected mass.

Modeless, thus Multimodal

Although an individual digital text can be presented in myriad ways, the representation is limited in form, due to the interface standards that have both developed purposefully and emerged as seemingly natural extensions of human-computer

¹⁸⁰ This does not, however, prevent a user from entering in an IP in place of the DNS (though this, too, could be filtered out). Yet, such blocked websites are still functionally crippled, as most links in a network point not to other IP addresses, but to further DNS entries. One could retrieve a search engine easily, but exploring the results of a query would be fruitless if those DNS entries, too, are blocked.

¹⁸¹ The argument being made here is not the similar argument presented earlier that concerned the materiality of the Digital. While the earlier argument had to do with concrete materiality in a physical world, the materiality addressed here concerns the social forces that affect, and are affected by, the technologies in question.

interaction. Even when emergent, however, the limits of the interaction are methods of control that were put in place by the underlying social apparatus that produces digital technology. Additionally, a computer interface is never a representation, but rather a dual misrepresentation: the bits and bytes are misrepresented to the human, and the limited set of choices a human has is misrepresented to the computer (Evens, “Icon Icon”). As with all forms of misrepresentation, this allows for an unimaginable number of coercive forces to work unnoticed by those uninformed on the operating principles behind digital technology; what might seem common and rudimentary to the average user can be—and often is—uncommon, complex, and threatening.

One of digital technology's often celebrated developments is its propensity to allow for fully immersive multimodal texts. And indeed, multimodality has shown its worth as a rhetorical, pedagogical, scientific and artistic tool, not to mention one for entertainment—which is exactly why such technologies, before they were a budding reality, were feared by many in the past, as exhibited by fictional works like *Fahrenheit 451* and *Brave New World*. As Janet Murray points out, non-digital books are praised in part because their limitations make the delusions placed within them easy to resist; but with multimodality, that delusion becomes harder to dismiss (Murray 13-26).

Unexpectedly, the structure of a modern interface comes not from a computer's ability to juggle many instructions so fast that it seems simultaneous, but from a very human imposition of modelessness. Before the rise of Graphical User Interfaces (GUIs), each and every application instituted its own mode, cultivating all available computer resources toward facilitating its particular modality; to end one mode and begin another,

an application had to be terminated and another initiated.¹⁸² Therefore, while a computer could easily re-present a digital movie that involved both audio and visual cues, in order to enter a primary mode of text entry, one had to exit the movie-execution program and start anew a word processing program. The naturally linear domain of digital technology was expressed in its operation.

Of course, this has changed: even when using the antiquated command line on today's computers, it is only one of many "windows" to be open. Anyone exposed to a modern computer is presented with the endless array of icons, windows and processes that define human-computer interaction;¹⁸³ the GUI interface has hidden the linearity and monomodality of computing by trading space, as presented to a human viewer within the confines of a screen, for processing time. In doing so, linearity is transformed into chaos, rhizomatic and gaseous to the observer, while modality is transformed into infinite modelessness—which, in essence, allows for the anything-goes, enriched multimodality that characterizes digital technology today.

Such a transformation follows from the *not-0-but-could-be-0* ontology of the digital world that the interface attempts to reconcile with its human agent, all the while passing along its impetus of control. For as the user is entertained and enmeshed within

¹⁸² While there is some difference in the definition of mode, in terms of computers, and the definition of mode in Communication Studies, the difference is made obsolete by the way digital technology facilitates contemporary means of communication. To facilitate multimodality, the computer must become modeless—and therefore multimodal.

¹⁸³ Granted, there are exceptions: there are countless computers that millions of people interact with daily that do not even have an explicit software interface: a Roomba or the computer in a car would be a perfect example. However, even if the end-user does not experience the technology through such an interface, the designers and programmers of the technology do, and that is very often done with a GUI.

an interactive, multimodal experience, the computer is carrying out the majority of its commands unseen in a non-modal plane that is entirely inaccessible—unless, that is, one has the proper tools and permissions to monitor them.¹⁸⁴ The infinite modality expressed in a modern interface encourages a user occupied constantly with so many sights, sounds and interactions that it's hardly necessary to hide the more glaring means of control underneath.

Intramediation and Spectacular Speed

One of the more troubling aspects of digital media, as elucidated by Hayles and her flickering signifiers,¹⁸⁵ is that the essential relationship between signifier and signified has changed dramatically, and within a very short historical framework. A text is no longer produced *en masse* and then distributed, but instead is forever reproduced in its distribution, subject to varying degrees of intermediation. Additionally, as Lev Manovich shows in *The Language of New Media*, the relationship between paradigm and

¹⁸⁴ Even the less tech-savvy computer users today often know how to bring up a list of computer processes running—but this is deceptive. The list does not tell you what the process is doing (and the names of the processes are often cryptic), nor does the list include processes being executed within the confines of a given piece of software. For instance, though it is trivial to see that a web browser is running in the background, it is nearly impossible for the average user to note the complex web of processes that each web page instantiates to track the user (or do something even more nefarious). While a user can examine the code of the web page in an attempt to identify such actions, most users are not provided the means to gaining the knowledge necessary to interpret computer code; and even if the user understands it, she must sift through thousands—or millions—of lines of code to identify these processes. Even then, the identifiable processes are only visible from the client side of the network arrangement; the server of the website spurs processes of its own that cannot be seen nor identified by the user. To better understand exactly how much information is shared, a group called Disconnect.me has released a plugin for various browser aptly named “Collusion” that shows the network of surveillance that happens just underneath the surface of each web page, and this author highly suggests that readers download it for their preferred browser.

¹⁸⁵ As previously stated, I would attribute the flicker to the very essence of the digital ontology, rather than what follows from Hayles's argument.

syntagm is reversed in digital media: what was once explicit is now implied, and what was once implied is now explicit (229-233). In the digital world, the database is given a material existence while its contents are dematerialized, relegated to the realm of virtuality; in Marx's absurdly prescient wording, *all that is solid melts into air*,¹⁸⁶ and the imaginary takes the place of the real. When a referent no longer exists, replaced instead by only references, signs cease to hold precedence over the symbolic world; what remains is the database, or what Baudrillard calls *the fanaticism of language*: form replaces content; media replace matter (*The Intelligence of Evil*).

The relationship between text and reader, and text and writer, is forever fragmented, and with it the reader (and writer) herself. We are now *homo fractalis* (59), our identities and apprehension increasingly fractured, fractalized, extracted, parsed, analyzed and returned by algorithms supposedly operating without prejudice or desire; we are made our own referent, and therefore without reference. If, as McLuhan suggested, computer technology was a literal extension of the nervous system (41-47), then it follows that the nervous system is now an extension of digital media. We are no longer extending ourselves but are being inundated and seized internally; we are controlled from the outside-in as well as from the inside-out.

A term such as “mediation” to describe the role of digital technology, then, is inadequate at best. While all media, to some degree, encourage internalization of both form and content, only digital media are capable of near real-time transformation and

¹⁸⁶ These famous lines are from chapter one of *The Communist Manifesto*—made even more famous by Marshall Berman's 1982 book on modernity and modernism, which borrows the phrase for its title.

adjustment; like an evolving bacterium, it changes its strategy and appearance until resistance to its entry is nullified. So too, one might argue, does all media—and like the bacterium, it does so by evolving over time. To equate these as the same, however, is to deny the central importance of this blinding speed at which digital media operates, and I do not mean “blinding” in any ableist sense; digital media is blinding in the sense that it can move at the speed of light as well as remain largely unseen as its algorithms operate upon us.¹⁸⁷ We are paradoxically blinded by and blind to it.

Indeed, it is this speed that drives digital technology further and further toward increased efficiency and an infinite expansion of its dominion—in short, its impetus of control. A singular characteristic of humankind (and, perhaps, all animals) that has allowed it providence is its ability to internally process and adapt at a pace faster than that which confronts it—yet this, too, has been reversed. We have no choice, if we are to survive, but to allow digital technology to invade us, to encompass us, to *become* us, and thus in turn make us, it. Digital media not only are the extension, the becoming of humankind; humankind is equally the extension and becoming of a digital kind.

Digital media are not mere mediations nor intermediations, but additionally *intramediations*: both internal and external at once.¹⁸⁸ Digital media is present within us as much as it is present in a computer; it is external to the machine as it is external to us.

¹⁸⁷ It is no mistake nor surprise that a wide array of media theorists and philosophers have started with the medium of light as a conceptual framework for understanding the implications of digital technology, from McLuhan to Baudrillard to Chun to Ellul.

¹⁸⁸ For the purposes of convention, “digital media” will still be used as opposed to “intramedia” in this essay, unless special attention must be given to this aspect of digital technology.

As we calculate, so we are calculated; as we interpret, so we are interpreted; as we read, so we are written; as we write, so we are read. As we feign to control, we are controlled. The very nature of digital technology compels it, like capital, to absorb its dialectical opposition, and its opposition is us: *we who are the mass and the medium, the network and the electric current* (Baudrillard, *The Intelligence of Evil* 135).

Invisible Machines, Disinterested Surveillance

As digital interconnectivity increases, this problem becomes magnified: those processes inaccessible to the user are wholly accessible to a third party that has a personal stake in collecting such information, whether it be for selling products, tracking habits, abusing the computer's connectivity,¹⁸⁹ or just plain identity theft. Nearly any given website today, regardless of reputation, generates a long series of background processes in the user's computer not only for the express purpose of surveying the user's behaviors—buying, selling, searching, speaking and so on—but also regenerates content based on the collected surveillance of past activities.¹⁹⁰ This regeneration, as illustrated before, is integral to the signification of a digital text, and should not, under other circumstances, be cause for any more alarm than a non-networked text; but whereas the re-presentation between reader and text in a non-networked environment is guided by already control-ridden code, a network allows for the dynamic redistribution of mediating

¹⁸⁹ A notable example of abusing a computer's connectivity is the existence of botnets, which are used for various, nefarious ends without the user's explicit knowledge. For instance, a computer might be used covertly to assist in a Distributed Denial of Service (DDOS) attack, effectively overloading a server with traffic, thereby preventing genuine traffic from accessing it.

¹⁹⁰ That is, a user's browsing habits are recorded; then, future content is based on the user's predefined habits, further refining with each action online.

code to fit the particular identifying mold that an individual inherits from the process of socialization. The controlling force is no longer simply the code, but a multitude of codes and algorithms aimed at adapting to each individual agent to facilitate the mass jurisdiction thereof.¹⁹¹

The consequences of this are many, but perhaps the most visible one to average users is the nature of advertising that they receive in relation to that which their online activity has shown to be their primary consumer interests. This seems, at first thought, an entirely sensible arrangement: surely, it benefits the advertiser by ensuring that the products are reaching a more desirable, less volatile market; and just as surely, a consumer presented with only those commodities for which she desires must be, by the logic of the process, a “happier” consumer who is more engaged with her particular brand of desire instead of products for which she has no interest.

But this too is a false positive; at the very least, human learning patterns are rapidly dissolved in this new context: in order to learn, one must be first presented with something unknown; and if all that is presented is familiar, then one can never learn—or, more likely, one only learns whatever a third party chooses for that individual to learn. Our exposure to the world beyond our own immediate surrounding is wholly framed by

¹⁹¹ Perhaps the most troubling aspect of Web 2.0 is its adaptive nature, which allows algorithms to alter in real-time based on the individual actions of those under surveillance. Where software was once hard-coded into certain methods of control, those methods can now change without the knowledge of the user; in a sense, the presentation of the text no longer follows any sort of standardization, except that which dynamically leads to better regulation and control. Though this is a basic feature of networking, it has proliferated on the web due to a technology called AJAX (“Asynchronous JavaScript And XML”). Although it’s a bit technical—as well as slightly dated—a functional understanding of AJAX can be gained from Ballard.

the flow of capital and the allure of commodities. Our collective ecology consists of cars instead of cows, pop music instead of bird songs and computer screens instead of vistas. We know nothing but products, commodities transformed into a hyperreal spectacle—the vortex of exchange, as Debord points out, that has attained the total occupation of social life (Debord 41).

And yet the phrase “social life” is particularly at risk of existing only as a form of nostalgia, for the social life itself has become entirely fragmented by the very spectacle that occupied it. The alienation that corresponds to spectacular society, which was integral to Marx's development of class consciousness, now swirls in further isolation from the social relations from which it originated, personalized toward “individual” consumer lifestyles and yet never satiated by them. We are no longer sitting in theaters or living rooms, by and large, consuming visual media in a set space that solidifies group cohesion, whether it be community or family; if there is a group to be had, it is based on shared patterns of consumption—consumer identities to be cataloged and parsed for future exploitation rather than shared modes of existence.¹⁹² I am a gamer, a collector, a foodie, a teacher, a student; *You are what you eat*, that ever-tiresome platitude, is turned on its head: *you eat what you are*.

¹⁹² Following this argument much further might risk being accused of what Nathan Jurgenson has dubbed *Digital Dualism*, through which the world is divided into an “on-line” and “off-line” binary rather than an integrated system. This essay does not aim to confirm nor deny such a conceptual framework. See Jurgenson’s “Digital Dualism versus Augmented Reality” at <http://thesocietypages.org/cyborgology/2011/02/24/digital-dualism-versus-augmented-reality/>

In the realm of advertising alone, this process creates a feedback loop that denies the consumer a means of transcending her own habits of consumption; unless something is amiss, the consumer will never be the target of advertising in which she has not already expressed an interest.¹⁹³ Beyond advertising, however, the problem becomes even more severe: the social algorithms of Web 2.0 guarantee that over time, the user rarely will be exposed to ideas—political or otherwise—that the person has not absorbed in the first place.¹⁹⁴ This further entrenches any ideological hold on an individual that the control society already possesses. The individual attention accorded to each consumer by corporations armed with digital technology, far from being an apparatus of empowerment, limits the agency of every consumer it touches—in effect, it controls.¹⁹⁵

As Baudrillard argued in *Simulacra and Simulation*: the acceleration of this deepening control is increasingly out-pacing any liberational effects digital technology might afford us (40). Indeed, he seems ever more justified in his conception of the perfect, all-encompassing simulation that destroys past chains of signification in its (im)materialization: *the vertigo of the world without flaws* (34). The scenario of encountering true difference (26), of being confronted with an experience not already entirely one's own, is not merely less plausible, but wholly implausible. The world beyond has become nothing more than what we always already knew it was—or, rather,

¹⁹³ It should be noted here that this qualification—*unless something is amiss*—is critical to the second part of this essay, and that things are more frequently amiss than they might first appear.

¹⁹⁴ This supposedly closed-loop system, of course, works far from ideally (which is a good thing); this is discussed further in part one.

¹⁹⁵ Control, by denotation or connotation, necessarily involves limiting the possibilities for agency of another.

what we were able to think it was. Plato's cave requires no chains nor shadows, but merely an Amazon account. We are governed, internally and externally, by code and code alone; “nothing exists beyond its search parameters,” and the responses to the query are decided before we ask the question. Baudrillard asserts that we are now but mirrors of the machine (*The Intelligence of Evil* 80-81); just as well, the virtual world in the machine that is available at our fingertips is becoming nothing but a mirror of our “individual” selves.¹⁹⁶ We are devoured by our own models, quantified at every turn.

Hegemony of Quantification

As is the nature of models, ours are the purely abstracted result of quantification. All objects, all persons, all phenomena, both material and imaginary, are subject to measurement at such a magnitude of scrutiny that room for no qualitative judgments is left. As illustrated before, the paradigm and syntagm have switched places, and with it the means of analysis. The referent is meaningless, the signifier empty; only in calculation is there substance left to be had. Further, by homogenizing all texts, whether they be linguistic, visual, or otherwise, digital technology compels one to shift focus from qualitative judgment toward quantification—and with intramediation, the human subject is the text being both written and read. The individual is nothing more than one thing among many to be written, parsed, stored, extrapolated.

¹⁹⁶ The word “individual” is in quotation marks because said individuality is purely constructed by the code under examination.

Through the digital, we are equal to the grandest star in the sky and the most bloated dead fish floating atop the lake alike: our bellies, relentlessly measured, relentlessly bloated as charged; our consciousness, rays of light born from the consolidation and quantification of the rays of energy that penetrate us from beyond, unnoticed. As we are users, so are we used; as we occupy digital media, so too does it occupy us. As Virilio notes in *Speed and Politics*, we domesticate an animal by making it accustomed to living in men's dwellings (34), to displace its own instinct with our own; as well, so too are we domesticated by the digital. The quantifiers of the past are the quantified of the future; we denizens of the present are witness to this fact. What we have quantified, tamed, developed, now quantifies, tames, develops us. We are not the producers and audience of digital media; we are the media.

The individual human, as a medium for the digital rather than a user of digital media, is absent of itself—and with it, so too become empty the humanities,¹⁹⁷ that *de facto* drum tower of qualitative analysis which has stood for centuries against the sieging tide of pure technique. The content becomes baseless and irrelevant to quantified beings; only measurement of the paradigm remains. The *how* and *why* of analysis gives way to quantifiable properties: not *why*, but *when*; not *how*, but *how many*. To remain relevant, the scholar of English must now become a statistician, an expert at analytic software:

¹⁹⁷ This is an overstatement—a noble lie, of sorts. Contrary to what's presented in the popular discourse of the matter, the humanities have been supposedly facing a crisis from the very dawn of industrialization, and many scholars have mistakenly feared the sky falling long before now. This discourse is nothing new; what's new is the means of communication, which imparts an immediacy on the alarm bells that was not present before. For further reading on the subject, see Perloff's *Crisis in the Humanities*, <http://epc.buffalo.edu/authors/perloff/articles/crisis.html>.

literary analysis, a focus on qualitative meaning, is trumped by word count, by quantifiable verbal constructions. The historian must not glean lessons or insight from this grand narrative we call history, but extract and compile data from it in a computable database; our modern focus on materialism has, like so many other things, turned on its head: where once it was material, now it is abstract. The form of the networked database—its properties, its connections, nodes—overrules its content.

This burgeoning remnant, these ashes of the humanities,¹⁹⁸ is coming to be known, quite unsurprisingly, as the “digital humanities,” and in many circles is seen as the savior for a dying breed of scholarship. Yet this focus on quantification as a savior is all too easy: the valorization of quantification is exactly what has transformed much of the university system into an antagonist of the humanities and its immeasurable qualitative emphases. Nobody needs to tell any serious scholar that quality cannot be always translated to quantity, or vice versa—but at the same time, the imperative of capitalism and its force of totalizing submission to the market requires exactly that impossibility. This is, in part, why digital technology has become so pervasive in late capitalism, and has further enabled the development of the control society: like capital, the digital requires that all that exists not only succumbs to its logic, but homogenizes and *becomes* it; just as capitalism requires commodification of even the very human beings who created it, so too does digitization require the same of us. The two are perfectly matched, incestuous bedfellows.

¹⁹⁸ This is not meant to be a condemnation of the digital humanities; this author sees such a development as inevitable and welcomes its arrival as much as he laments the dousing of the flame.

The common discourse is increasingly one of digitization as an answer to any problem: starving children are sent laptops instead of food; the inefficient but fairly reliable polling booth is replaced with electronic voting that trades reliability with efficiency; hard cash is replaced with digital credit; rebels are sent access to Twitter instead of guns; classrooms are replaced with Massively Open Online Courses (MOOCs), virtual courses that prepackage the scholar, classroom and knowledge as a singular banking-method commodity to be traded and consumed rather than discussed, considered and synthesized;¹⁹⁹ and humanities scholars are given a mandate for computerization rather than the human. There is very little *human* left in humanity to be had, let alone the humanities. To reiterate Ellul, nearly a half-century after he first put it in words: “The integration of the individual into the technical complex is more complete than ever before.” (155)

Return of the Bit

The mere existence of digitization, it seems, requires that we all become techno-utopians, cyborgs bent on adhering to the idea that if we cannot beat it, we must then join it; we must integrate. This is partially, if not wholly, due to the ontology of the digital that pervades everything it touches: even if the surplus of possibility is ideologically limited to late capitalism, the *not-0-but-could-be-0* logic of computation remains in full force. It is unstoppable, barring a disaster of unimaginable proportions. We are its vessel, its medium, and just as digital technology domesticates and homogenizes everything it

¹⁹⁹ Far more than any other group, MOOC supporters both silence and ignore the very real criticism of their wholesale and troubling acceptance of a digital savior.

consumes and reproduces, the uncritical discourse thereof becomes increasingly hegemonic; to merely question the merit of digital technology's messianic futurism is an act of subversion, often responded to with either silence—rendering the act itself unheard—or ad hominem attacks. Luddite or Technocrat, Primitivism or Neoliberal complicity: discourse itself has become binary, eschewing qualifiers and shades of difference. There is no middle ground between the binary 1 and the binary 0.

And yet, while digital technology is not the fantasy that it is often portrayed to be, and in fact facilitates a means of control unprecedented in human history, the digital is also futile to fight against, and it is perhaps even counter-productive to do so. What is needed is not a means of dismantling the technological underpinnings, *a la* Ted Kaczynski, but instead a seizing control of the very code that facilitates it; of subverting the current model and creating our own digital landscape, whether it be simulation or otherwise, supplanting that which has already been provided. Above all, we must not uncritically accept that digital technology is a savior, whether it be for the humanities in particular or humanity in general; for there are now computational codes behind the means of semiotic representation, and they are much more treacherously opaque than codes of the past. The digital *is* control, but let us not forget the bit's quirk of extending its properties to that which it composes: 0 is not-one and could-be-one; 1 is not-zero as well as could-be-zero. The digital is *control, but also could-be-not-control*. Let us acknowledge its current state, and in keeping with its ontology, focus it anew; let us now turn our current *Zero* into a *One*.

CHAPTER V

A RHETORETHIC

OF HERACLITEAN MACHINES

If rhetoric is the leading of the soul, then where is rhetoric leading the soul?

– James Crosswhite, *Deep Rhetoric* (29)

Given that our concern with cognitive machines pivots the matter of control in practice, narrative, ontology and sentiment, it should be unsurprising that so too it seems that the question of machine ethics ultimately focuses on how human beings are treated by machines as moral patients rather than the moral constitution of machines themselves, whether they be in the past, present, or future—real or fictitious. What matters to the machine ethicist is not whether a machine can act as a moral agent toward some good, but rather how human beings may subjectively benefit or suffer, often at an atomistic scope; grand moral interrogations are left wanting as we consider how an autonomous vehicle might navigate the ethical dilemmas posed by unavoidable, atrocious accidents, in favor a single imperative: humans benefit at all costs, even when—and, we argue here,

especially when—that mandate of human benefit is at odds with the very moral behavior we seek to instill.

Like our approaches to technology in general, our moral reasoning about machines is rather instrumental: how they might be *used* (controlled) for human benefit, not that they are used in the first place; and while our mythologies suggest an overriding concern with the possible shipwrecks that accompany the invention of our ship, the moral status of machines as agents and patients is altogether bypassed in the process, centering not on how machines might make moral decisions in general, but how they might come to conclusions that benefit their users—human beings. This instrumental approach has a long history beyond what we dub “machine” today: the ethics of slave ownership, worker exploitation under various economic systems, and the treatment of nonhuman entities—beast and inanimate alike—have classically concerned the well-being of the beneficiaries of oppression rather than those being used.²⁰⁰ The overriding concern is one of reinforcing an already established hegemony that can be recognized most often in the demarcation between objecthood and subjectivity—the user and the used.

Against the most basic assumptions of Western thought, that which gets defined as object or subject, user or used, not only implies how said entity ought to be treated and ought to treat others, but heavily codetermines normative judgments regarding it. Kant readily recognized and even embraced the notion that what matters is not how the

²⁰⁰ While some have suggested that the propensity for inclusion into the moral sphere has grown with time, and the existence of some moral philosophers like Peter Singer support this view, it would be a hasty conclusion to suggest that this expansion must necessarily continue to grow—and a dangerous suggestion that it will not again deflate.

“object” is treated, whether the object be a machine, a dog, or otherwise, but rather how its treatment affects the moral subject: the individual human being who is given moral agency and privilege over all other entities. Kant justified this dichotomy by identifying that which is capable of Reason, and thus moral behavior, from the rest of the object-world: human beings.

As we have seen, however, it has become increasingly difficult to attribute a capacity for reason to humans alone among other animals, let alone among our perfect reasoning machines (and to be fair to Kant, the reasoning capacity of machines was primitive in comparison to that today). If moral consideration is dependent on the capacity for reasoning, and if reason is the mechanization of thought, then machines deserve not only our moral consideration, but admiration: pure reasoning beings are, by this logic, pure moral beings, incapable of evil or malice; much like the Christian God, machines become the embodiment of the Good and the Right, unable to make either mistake or morally questionable decision—of loving grace, indeed.²⁰¹

But if we follow the ensuing instrumental logic of technology, then we are faced with a potential paradox: the perfect good (machines) being *used* toward evil or malicious ends. How can the perfect good transform in its use to such illness? Such a question has been asked throughout history, with no consistently satisfactory answers—how does a good God allow for evil in the world? How have many of the most radically egalitarian

²⁰¹ To reiterate: Buddhists like Masahiro Mori have come to this conclusion as well; the difference between Mori’s conclusion and our own is a matter of how horrifying the conclusion might be, correct or not.

and liberating social movements in our collective history become so dangerously oppressive? Religious or secular, the conundrum remains a silhouette against an otherwise clean canvas, tarnishing efforts to create a more just and equitable world with or without machines capable of discerning for themselves the moral character thereof. *The path to hell is lined with good intentions*, the commonplace warns us—but it says nothing of how the point of origin is found in heaven.

Further complicating the notion is the matter outlined in Chapter IV: the subject becoming subjugated, the user being used by its object of use. Not only can we not rely on our absolutist notions of good and evil for guidance without quickly approaching paradox, making any Aristotelian notion of a golden mean impossible, but we can neither rely on Cartesian distinctions between subject and object, world and person, force and content, the one and the many. Putting aside the problems this might pose in the sciences and philosophy, we are still left with the moral implications of a thoroughly amoral worldview, courting the most extreme fallacies of moral relativism and at risk of undermining any chance for stasis whatsoever. If we follow Hume into a world indifferent to our sentiments and moral obligations, then we are left without a need for ethical machines and ethical people alike; if moral judgment is relative to community standards, then what constitutes good and evil is a matter of persuasion and persuasion alone.

Despite our protest here to the contrary, the essence of morality, and the ethical systems built around it, appears now to be a matter of “mere rhetoric.”²⁰² With all distinctions compressed into what is socially constructed, everything becomes an article of faith; difference, absolute or otherwise, moral or not, material or ideal, becomes aesthetic. Before we protest too much this appearance, and protest it rightly so, however, we first need to step back toward the digital to unearth an essential metaphysical characteristic of its operation, and then step further still in time to find a possible corollary. In the process we will find what perhaps links rhetoric and morality by virtue of a shared core component—and, by additional virtue of shared ontology, perhaps find a path forward for machine ethics thus far unexplored.

The Bit Strikes Back

As already noted, this compression of difference is by no means a new problem, nor has the failure to find its solution prevented any number of miraculous inventions, both abstract and material; computational systems, abstract and material inventions par excellence, operate precisely on the principle of absolute difference; abstraction “achieves in the bit an apotheosis,” as Evens describes it (9), separate from the computer’s materiality in form and function but wholly dependent on it nonetheless. Yet shades of gray (and indeed, a full spectrum of 16.8 million colors, over 150% of the number of colors recognizable by the human eye), as most any computer user can attest,

²⁰² By “mere rhetoric” we mean to use the frequently derogatory connotation associated with the term in popular discourse to not confirm it, but to illustrate its lack of foresight; for if anything can be viewed as something mere, it would not be the agential basis for communication—and perhaps much more than that.

appear quite miraculously on the electronic screen despite such ontological dedication to the absolute, and as illustrated already, any entity that can be represented at all can be represented through binary code. While we may not have achieved *ex uno plures* (“from the one, many”), we have come as close with the digital as one can possibly get without achieving it, *ex duo plures*: many from the two.

It seems, then, that at least in abstract terms, absolutism overcomes its own paradox through mimesis: there may be only one (1) or the other (0), but this leaves open the possibility of having many ones and many others of the same kind.²⁰³ Contra Plato’s Forms, however, the many are not mere copies of the originary one, but the same; it is a matter of many ones rather than the one and the many. This does not solve the problem of the one and the many, but rather side-steps it: by operating in terms of pure abstraction, the one (1) and its negation (0) combines to establish multiplicity of both form and content—the two of which, in fact, become indistinguishable under a digital ontology. Form is content, content is form, and the smallest flicker of 0 to 1 changes everything; Searle’s nightmare ensues.

Here we do not claim that the problem of absolutes can be overcome (quite the opposite), nor are we committed to a digital metaphysics and physics of the universe alike, as does MIT Professor of Mechanical Engineering Seth Lloyd; rather, we claim

²⁰³ In *Theories of Mimesis*, Melburg addresses the issue of how mimesis transforms into repetition through time, creating new forms that are independent of the external milieu that shaped them—that is, the representation becomes an entity itself. This is an important point to consider especially as it relates to cognitive machines; they may be *mere copies* of cognitive agents, but through repetition they become substantiated as cognitive agents, albeit of a different order.

that absolutes are absolute in that they are composed further of more primitive yet still paradoxical absolutes—1 and 0, being and non-being, affirmation and negation—and for this very reason, again paradoxically, are subject to radical alteration with only minor changes in an absolute's primary elements, the veracity of which are subject to the whims of social construction and persuasive capacities. The premises of most arguments are vulnerable to *reductio ad absurdum* attacks—reducing the logic of a claim to a paradoxical point at which its conclusion is both true and false at the same time (that is, absurd or unreasonable)—not because we lack the appropriate *lingua generalis* for reasoning, but because this elementary paradox is prior to and constitutive of all claims to truth and falsehood. Far from ridding reason of its abstract paradoxes, the absolutist logic of the digital wholly embraces them; and equally far from ridding philosophy of its sister rhetoric, the digital reveals and confirms their inseparability: it matters not, in computation, whether a bit is or becomes a 0 or a 1, but rather how that bit is interpreted—by an interpreter that is composed too of other contradictory bits.

Classical reasoning may not tolerate well a paradox on the surface of its claims, but on rhetoric reason and its premises thrive—and rhetoric itself is driven by contradiction. Perhaps the most glaring example of this drive is the sophistic practice of *Dissoi Logoi* (the title of the sophistic text that coined the term, written around 425 BCE): arguing opposing sides of an issue to come to a deeper understanding of the it—and perhaps revealing a deeper truth residing underneath both claims. Of note here, however, is that the anonymously written text not only questions the nature of truth and falsehood,

but the morally good (right) and bad (wrong), presaging today's moral relativism by millennia:²⁰⁴

[One] view is that what is good is one thing and what is bad is another thing; as the name differs, so likewise does the reality. I myself also distinguish the two in the above-mentioned manner. For I think it not even clear what sort of thing would be good and what sort of thing bad if each of the two were the same thing and not different things; the situation would be an astonishing one indeed... I think that if one were to order all mankind to bring together into a single pile all that each individual considered shameful, and then again to take from this mass what each thought seemly, nothing would be left, but they would all, severally, take away everything. For not everyone has the same views. (Robinson, Anonymous A11-B18)

Clearly, as Schiappa suggests, the author of *Dissoi Logoi* was struggling with issues concerning frame of reference and level of abstraction (147)—issues that most adults today, to risk generalizing, have little trouble navigating. The winning of a football match is a loss for the other team, and victory in war is devastation for the enemy; we know, intuitively, that the scope of gain and loss, good and bad, illness and health depends on context and circumstance. Surely, though chemotherapy may be beneficial for the cancer patient, otherwise healthy individuals know to avoid it lest they seek to fall ill; and just as surely, while a gluten-free diet is necessary for an individual with Celiac's

²⁰⁴ At least according to surviving texts by Plato, it appears that what is morally good and what is true were, if not the same, then inseparably bound.

disease, others may find gluten necessary, as it is a protein easily available even to the poverty-stricken, to live moderately healthy lives.²⁰⁵

Plato's dialogues, according to Schiappa, collectively argue that we cannot define an expression by the statement "X is not-Y" because it indicates, foremost, that we really have no understanding of what Y is; if X is not-Y, then so too is Y not-X. We are left with the never-ending circularity of Hegel's determinate negation, with an entity always containing, in its definition if not its material constitution, its opposite. Such is the case with the binary logic of the digital,²⁰⁶ as 0 must be defined as not-1 and 1 defined as not-0 exclusively. Within the 1, then, so exists the 0, and within the zero, one; each exists only by virtue of its counterpart and carries no meaning or purpose without the other. The 1 and the 0—which, as illustrated, need not represent numerical values—are the most atomistic eternal opposites, together forming a unitary whole that, thanks to the accidental coincidence of visual similarity, we will represent forthwith using the symbol Φ (*phi*)—the one and the zero, true and false, being and non-being, control and freedom, merged.

As the self-contradictory embodiment of absolute difference, Φ is useful for our purposes as conceptual shorthand for a phenomenon that we argue here is integral to the existence of persuasive and moral behavior—and one that happens to be the foundation

²⁰⁵ We are, of course, being facetious here: many of us treat the medical necessities of others as dietary choices.

²⁰⁶ It should be noted that "the digital" need not be binary nor need it be electronic; what identifies the digital is its reduction of all things into discrete, interchangeable units that may be strung together to reconstitute an approximal representation of its object. As a digital computer can be constructed from beer cans, as Searle is always quick to point out, so too can the digital operate as such.

of digital technology today. We could easily refer to this unit of fundamental difference as paradox or contradiction, or perhaps even division by zero.²⁰⁷ However, Φ best suits our purposes for three primary reasons: 1) As illustrated above, it is *visually* reflective of the concept it seeks to signify; 2) It avoids the swift adverse reactions normally associated with matters of paradox and contradiction, especially among those whose work is indicative of “the engineering mentality.”²⁰⁸; and 3) it reminds us that far from being a conceptual apparatus newly introduced to the grand equation of cognition and consciousness, it is rather one that seems to have existed for quite some time, perhaps for the entirety of written history; and while it has manifested itself across many varied cultures, we turn now to the one most obviously related to the symbol itself, Ancient Greece, and to one Greek in particular: Heraclitus.

The Force Flame

Heraclitus, famously, believes in change. One can never step in the same river, he contends, because the water that flows will be different from the water before, and even the minute changes of the river bank, the sediment deposits, the PH levels render the river a vastly different entity from one moment to the next. And yet the river is the same river nonetheless; Heraclitus also admitted paradox and contradiction as a fundamental “rule”

²⁰⁷ While division by zero in mathematics is generally regarded as always returning an undefined, this is only the case in classical arithmetic; there are mathematical structures in which division by 0 does return a definitive value.

²⁰⁸ “The Engineering Mentality” focuses on how a process or object can be *used* to achieve a purpose as efficiently as possible and exhibits little care for other considerations; in this sense, it is intimately tied to the instrumental view of technology and functionalism, and perfectly fits the aims and methodologies of contemporary economics and business—for better *and* for worse.

(*nomos*) of the universe. Rather than the obscurity and unreliability of language failing to adequately describe reality, it is reality itself that reflects its paradoxical nature in our use of language. Contra efforts to cleanse language of its “irrational” elements, like the continued drive to create the ideal *lingua generalis*, Heraclitus argues, as far as we can tell, that language as it is already reflects the nature of the Cosmos: paradox exist in our discourse because it exists objectively.²⁰⁹

It follows, then, that attempts to cleanse language of its propensity for contradiction makes it less reflective of the reality it seeks to describe, rather than more so, and the surviving fragments of his work confirm this preoccupation: littered with chiasmic pairings, many of the terms of which are also strictly antagonistic, Heraclitus presents in the form of his sentence structure the content thereof, conceptually unifying what is generally seen as two distinct domains: form, the arrangement, medium and style used, and content, the substance or “essence” of the message being communicated. Where once the two are considered disparate, if not opposing, conceptual schemes, here they are unified and consolidated: form becomes content, content becomes form, simultaneously.

Like our notions of control and freedom, and like the digital’s infinite flicker between 0 and 1 (Φ), so too does form and content on one hand unify and on the other disperse, embodying the paradox that philosophers and rhetoricians alike²¹⁰ have

²⁰⁹ Whether these paradoxes exist prior to language or because of the observational capacities language provides is beyond the scope of this dissertation; what matters for our purposes is simply that they exist as part of the cognitive subject’s lived experience.

²¹⁰ As Poster is keen to show, the disciplinary separation of philosophy and rhetoric did not exist prior to

attempted in vain to avoid by following Aristotle's *principle of non-contradiction*: "the same attribute cannot at the same time belong and not belong to the same subject in the same respect" (Miller 14). Yet as Patrick Lee Miller shows in his *Becoming God*, the principle itself actively encourages the generation of paradox: to separate a term from its opposite requires that we extract them both from the passage of time, turning them from processes into static objects, from the constantly changing to the eternally unchanged. That is, contradiction itself arises when we attempt to extract the object of our contemplation from the immersion of time and change that defines it, freezing the flame of a candle like a still of film so that it needs and is satiated by its fuel simultaneously in the moment we suspend it (16-18). Φ is either 0 or 1 only when we isolate it from its "natural," unadulterated state of constant change, and from this extraction arises that primary contradiction: 0 is not 1, and 1 is not 0.

Ultimately, the principle of non-contradiction contradicts itself—and yet we cannot avoid its use or value. Zero and one, form and content, control and freedom, need and satiety are inseparable pairings, distinct only in the observers' act of making the distinction in the first place. And so too, we might concur with Heraclitus, exists the distinction between observer (subject) and observed: that which one observes is no other than herself cleaved²¹¹ in two, courting again not only the excesses of extreme pluralism

Plato's bold declarations to the contrary, and even afterward the two were not wholly divided until the development of vast disciplinary specializations in the 20th century. In this author's professional opinion, it is not a division that had much merit during Plato's life, nor does it today.

²¹¹ "To cleave" in English means both to break in two as well as join and would be a tempting English alternative to *syllapsis* given its paradoxical nature, were it not for its unfortunate association with the horror genre of film.

but also of solipsistic idealism. If all that is observed is a cleaving of the self, then nothing exists but the self-cleaved: endlessly broken apart and put back together again, the world is a humpty-dumpty of my own making, figuratively and literally, and everything in it is me. *Know thyself*, the famous commandment of the Delphic oracles, is redundant and already fulfilled: I can know nothing but myself.

Doxa Vadere²¹²

If historical recollection proves accurate, however, then our radical idealism here faces the insistence of a material reality shared by all that Heraclitus demanded—and we are left without any sense of how this self-cognition happens in the first place, especially as it relates to the rhetorical and ethical concerns at hand. By most contemporary accounts, Heraclitus was part of a conceptual movement in Ancient Greece that foregrounded materialism as a working assumption, and today’s schism between the ideal and the material is thought to have not yet been theorized. Even Aristotle, whose chronological proximity was much closer to Heraclitus and thus his conceptual schemes far less clouded by additions to Heraclitean philosophy by the Stoics and Christians, characterized Heraclitus strictly as a material philosopher—he was, by most accounts, a naturalist.

This not only condemns our newly minted Heraclitean idealism to be hastily discarded, but likewise poses innumerable problems for the interpretation of *Logos* in his

²¹² This is a play on *Darth Vader*, in line with the chapter’s recurring theme. Etymologically, it loosely translates to “to walk the common belief.” Linguistic purists would scoff at our mixture of Latin and Greek here, were they not preoccupied with televisions, sociology, heterosexuals, homosexuals, and those tragic tonsillectomies.

fragments as an everlasting divine law or cosmic principle set above and apart from the material world; if indeed a separation between the material and the ideal had yet to be theorized, then *Logos* was and is a material entity born from and subjecting of the universe, afflicting strife and justice as much as it is afflicted by it, rather than an immortal concept free from change. Like the so-called objects in computer programming paradigms, *Logos* is as much a process as it is an object, if not more appropriately the former than the latter; the analytical separation of its form and content presents *Logos* as a static object outside of the flow of time as an unchanging ideal—a Form—distilled from the crucial flux of the changing material world that gives it life.

As both a process through time and as a distilled ideal object, however, *Logos* further can be reduced, and has been reduced, to the two oppositional component parts of analysis—the breaking and tearing apart of a conceptual object—and synthesis, joining together as one. *Logos*, in short, is itself a paradox, a point which Aristotle, at least, either understands or accepts as a “suspension of meaning” (Reames 332). The objective form that *Logos* takes, and the corresponding term used in attempt to describe it, is wholly dependent on when and how *Logos* is extracted from its continual procession: here it is translated as “discourse,” there it is “Logic,” further down it is “speech” then “language,” and eventually, if we keep trying, we will find its earthly meaning, “ground,” a place upon which to stand (*stasis*), which nonetheless is constantly moving in our Heraclitean flux.²¹³

²¹³ Aristotle’s *stasis theory* plays on this location-based and orientational language: One finds an argumentative place to stand on the grounds of *logos*.

The task of Logos common to all is to break apart and together join, to analyze and to synthesize—to cleave—regardless of the form it takes during our temporal suspension of it and *irregardless*²¹⁴ of our attempts to bracket the living process as one or the other, bringing us again to the problem of the one and the many, unity and plurality, “wholes and not wholes.” As Kahn argues in his commentary on the fragments in *The Art and Thought of Heraclitus*, the whole and the not whole, the unity and its parts, compose a most abstract yet central feature of “intellectual synthesis:” combining objects in thought, which are themselves temporally arrested processes, and “seeing them together as larger unities” (282). Our capacity for intellectual analysis, on the other hand, provides a means for breaking the unity of *Logos* into constituent elements, driven in part by our principle of non-contradiction: to insist on only true and false, one or the other but never both at once, is to already break apart the living unity of opposites. We may cut up our experience into distinct wholes like day and night (283), but at the expense of the greater synthetic unity of opposition: day-and-night, whole-and-not-whole, up-and-down, true-and-false, 1-and-0, Φ .

Syllap Stick / Straw Men

The world, in Miller’s phrasing, is “one and ever-living, but is nonetheless a cauldron of many things” (24)—and it follows that so too must be our means of cognizing it, Logos, if we are to perceive it at all (which we just have).²¹⁵ According to

²¹⁴ We of course know that “irregardless” is not standard, formal English. However, as it exhibits our point in multiple regards, and lends itself well to a doubly contradictive chiasmus, we find it perfectly appropriate here.

²¹⁵ We do not claim here the absurd position that since a concept can be described it words, it must

both Kahn and Miller, Heraclitus uses the word *syllapsies*—usually translated as “graspings”—to refer to this real-time, uninterrupted process of human cognition. It is Logos in its purest form: analysis and synthesis simultaneous and at once, breaking apart and putting together, and like Bohr argued passionately about the path of an electron, it is difficult, if not impossible, to visualize; it is *unanschaulich*.²¹⁶

Despite this unvisualizable nature, these graspings—individually, *this* grasping, *syllapsis*²¹⁷— is yet still integral to what human beings, and possibly other cognitive agents, can accomplish: conceptually, we may analyze and synthesize without the need for linear procession, splitting Φ into 1001²¹⁸ and combining it to Φ at the same time and with almost unbelievable ease. Paradoxically, we see—or rather not see, but *think*—the whole and not-whole, the one-and-the-many, plurality and unity, at once; generally, barring learning disabilities or brain injuries, an individual has few problems recognizing a bicycle as an individual entity (or process) while simultaneously recognizing it as the sum or more (synthesis) of its parts (analysis). Facial recognition algorithms, whether utilized on linear systems or distributed parallel processing systems,²¹⁹ could be said to

therefore exist; rather, to cognize a material world in any accurate sense, and to have sense at all as equally material beings, our means of cognition must match or otherwise be reflective of the greater reality of which it is part.

²¹⁶ Although we have made some analogies to the paradoxical quantum state of photons in this work, we will refrain from making more detailed references to quantum physics for the sake of the science community’s sanity.

²¹⁷ As far as this author can discern, Miller has first coined the term *syllapsis* in his book *Becoming God*, though the same principle has been referred to by many, and his primary source on the matter seems to be Kahn.

²¹⁸ We use 1001 due to its resemblance to the pattern of chiasmus: + - - +, - + + -.

²¹⁹ In case the terminology is forgotten: A Distributed Parallel Processing system uses multiple and separate processors simultaneously to render input into output, rather than the classical architecture of digital computation which uses only one central processor. Notably, most contemporary computers use some form of distributed parallel processing, though often in meager amounts.

accomplish this task on a rudimentary level, but only if we ignore the fact that the algorithm has no actual conception of the whole face, but rather probabilistically determines whether an image might qualify as a face by the mere relation of its parts—the form.

This kind of probabilistic facial recognition may indeed reflect how biological



Fig. 5.1: Pareidolia

organisms recognize faces and may help to explain mental phenomena like *pareidolia*: recognizing faces in objects and arrangements that, intuitively, we know do not carry faces. Alphabet’s *Deep Dream*, an artificial neural network using their *Deep Mind* structure (the same neural network that

soundly defeated Lee Sedol in the Go championship matches) that loops over and over an image to find faces—or any other object desired. Figure 5.2 shows one result of trying to find the face of a dog in an image of a sandwich: individual elements, contours and colors, are twisted and rendered repeatedly until recognizable features are apparent, usually taking the form of the strongest elements of a face—in this case, noses and eyes.



Fig. 5.2: Deep Dream searching for “Doggo” faces²²⁰

To mistake what Deep Dream does with the paradoxical wholeness-not-wholeness of conceptual unity and plurality—syllapsis—would ultimately be a grave error, however. For the Deep Dream algorithm, there is no whole, save for what the human viewer recognizes in its output; there exists only schemes of parts that comprise an abstracted substitution for the whole. Such is the domain of all digital representation and constitutes its very nature: the violent disassembly of wholes—the analysis of the analog—that can be then reconstituted or synthesized according to this logic of disparate parts; Φ is discarded and in its place substituted 0 and 1, form over substance, parts over whole. It is an effective, but limited, imitation of human cognition; and as we shall soon argue, this limitation further constrains the possibility of a rhetorical and ethical constitution for artificial agents.

²²⁰ A “doggo” is internet slang for “a big ol’ pupper,” and usually refers to a type of dog typically used in internet memes.

To clarify our point, it is worth lingering on the considerable amount of effort and time that Kahn spends analyzing fragment CXXIV, the penultimate passage in the collection,²²¹ in which Heraclitus most clearly defines, if such phrasing could be used when addressing Heraclitus “The Obscure,” *syllapsis*:

Graspings (*syllapsies*): wholes and not wholes, convergent divergent, consonant dissonant, from all things one and from one thing all.

Etymologically, Kahn observes, *syllapsis* denotes a “taking-together,” the “cognitive act of collecting together, comprehending, summing up,” and refers us to the root of *syllambanein*, “to comprehend” (282). Other commentators (Bollack-Wismann; Kirk; Marcovich; Snell) assign to *syllapsis* related meanings like “assemblages,” “things taken together,” “connections,” or “Zusammensetzungen,²²²” but none of these latter definitions can be attested in contemporaneously appropriate usage. Aristotle himself could have used the word to additionally mean “pregnancy” or “conception,” a fact that might remind us of our use of the word *concept*, but prior to Aristotle, the only known usage for the word was quite literally the physical act of grasping or seizing—a material action for a material philosopher (281).

To practice *syllapsis*, to truly comprehend the world, each other, and ourselves, we must *grasp* that which we are *grasping*—*syllapsis* is, in true Heraclitean fashion, both

²²¹ This fragment is penultimate according to Kahn’s own system of ordering and does not necessarily represent the only order in which the fragments might be organized.

²²² “Zusammensetzungen” might be best translated to English as “Compounds” or “Compositions.” The plurality here is of greater importance than the target word.

object and process, noun and verb, concrete and abstract. Unlike Reames, however, and much like her version of Aristotle, Kahn takes no side on which interpretation is preferable. Acknowledgment of this unity of opposites is the operating principle and precisely the point of *syllapsis*: it is both and neither, convergent (*Symphoromenon*) divergent (*Diaphoromenon*), consonant (*Synaidon*) dissonant (*Diaidon*) (284). And it cannot be visualized, despite its strong “objective” component, because it is better suited to our perception of aural phenomena, like the latter chiasmus refers: musical harmony, after all, is the unity of a plurality of voices, tones, and melodies. Through the synthesis of disparate elements, unity is achieved; and through this unity, the many parts are observed.

Perhaps taking too many liberties, Kahn updates the 124th fragment to reflect his interpretation:

Graspings, that is to say groups holding together, apprehensions bringing things together: these are wholes and not wholes; they characterize a system which is convergent, divergent, structured by cooperation and conflict; this system is consonant, dissonant, held together by harmony and discord alike; from all its components a unity emerges, and from this unity all things emerge (286).

Whether *syllapsis* refers to one sense or the other, material / object or abstracted / process, misses the point—and so too does it miss the point to insist on interpreting *syllapsis* as a material entity, like Reames, rather than a universal principle, like those with whom she contends. Heraclitus exploits the dual meaning and nature of *syllapsis* as

much as he exploits the ambiguity of Logos as speech and as reason alike, material ground and conceptual schemes; to insist on one meaning or the other is to interpret syllapsis, like with our unitary Φ , already as only analysis, only 0, thus missing not the path of Reason,²²³ but the conditions of Pre-Reason requisite for the use of Reason at all. Nor must we side with materialist philosophers versus idealist philosophers or Analytical camps versus Continental ones, because our position is not reasonable or unreasonable, but *areasonable*—it demands not the impossible, but rather to put it reasonably: the ground upon which we stand predates the germination of Reason’s seed. It is the place from which Reason, and Rhetoric, and Morality, spring forth.

Sylph and

Nietzsche characterized grammar as “folk metaphysics” (214). The categorical structuring of reality by grammatical rules becomes the de facto ontology from which each person either spends her life confirming or, like Nietzsche, spends even more of her life denying: Subject, Object, I, You, It. Given that he was, for the most part, a Heraclitean, this should come as no surprise; it is foremost language that allows us to split the subject from the object, the I from the You, and the It from—well, Not-It. Language separates syllapsis into disparate analyses and syntheses; it separates Φ into opposing 0s and 1s; and it cleaves from the eternal flux of the world timeless instances we come to call objects and things. Like the gold-seeking alchemist attempting to turn chemistry into chemical, or the divining computer programmer who calls a collection of

²²³ Note that when using the term “Reason” we are mostly referencing logical reasoning here—the kind that machines are the perfect manifestation.

binary signals an object in one glance and an algorithm²²⁴ at another, language lets us forget, in a comforting gesture, that just under the surface, paradox swells.

Like the principle of non-contradiction, however, language seeks to solve paradox qua paradox. Miller suggests that Heraclitus is not illuminating the fundamentally contradictory nature of the universe alone, but also extends and amplifies Kahn's line of (un-)reasoning by showing how Heraclitus exploits the syllaptic unity in opposition of self-inquiry: the subject inquiring the self as object, turned subject inquiring the self as object, *ad infinitum* (26-27). This process reveals the paradox at hand: how can one be the object of one's own search? The flame, eternal or fleeting, burns fuel to sustain its burning; self-inquiry inquires itself. The self as a process of syllapsis becomes unlimited in its objectification, impossible to grasp (by itself) as much as it is impossible to grasp fire: to do so would mean to freeze it in time and place, diminishing its essence except by the redeeming unity-in-opposition of paradox: through the syllaptic contradiction of self-inquiry, the self-understanding subject is born.

Other (You're It)

Like the multiplicity of 1s and 0s that combine to form representations that copy themselves into realization, however, there exists a multiplicity of subjects—all subjecting and objecting themselves and each other. Here we have exclusive and infinitely self-recursive subjects coming and going into and out of relation to one another,

²²⁴ An algorithm is a *process* defined by discrete steps; while computer programmers feign fealty to objecthood, they have yet to discard the most revealing terms in their repertoire.

analyzing and synthesizing (syllapsizing) not only themselves, but the everlasting subjectivities of others—and here is where the rhetorical and moral possibilities arise. For not only does one and zero or true and false arise from the primitive paradox of Φ , but also good and evil, right and wrong; and how an individual cognitive agent interprets a phenomenon as either one term of a binary or the other is co-determined by how that subject is persuaded to believe it so; persuasion may not precede causality here (thus the necessity of the prefix “co-“), but it directly contributes to exactly how a subject interprets the causal, whether that causal factor be part of its self-recursive reflexivity, part of the natural procession in the flux of space-time, or part of the infinite loop of the subjective other—and this persuasion, as arbitration of whether Φ is interpreted as either 1 or 0, good or evil, becomes a de facto causal force itself. As Poster notes in “The Task of the Bow,” the act of interpretation itself requires syllapsis because it “is a process that requires the interpreter to shift constantly between considering part and whole, text and context, and purposes and product” (4). Since all phenomena require interpretation in the act of conceptualizing, all those interpreting require syllaptic capabilities; and since persuasion is this very arbitration between unified opposites, it is the core of all subsequent behavior.

Thus we return to the suggestion that morality is but “mere rhetoric” with a perhaps counterintuitive rebuttal: it is rhetoric, but there is nothing mere about it. For if persuasion is the process by which a fundamental contradiction is either momentarily resolved or momentarily reinstated—or both, paradoxically, at once—and what is morally good is in oppositional unity with that which is evil (good-evil), then persuasive

capacity becomes a prerequisite for acting as a moral agent and building the ethical heuristics necessary for functional interdependence. In the fashion of Heraclitus, however: as morality rests upon the exercise of rhetoric, so does rhetoric rest upon the exercise of morality. Like a doubly entwined Ouroboros, the mouth of rhetoric lies on the body of morality to sustain it, while morality doubles back to be sustained by the body of rhetoric; a mobius strip in which the inside becomes the outside and the outside in, rhetoric and morality form a convoluted unity that cannot be disentangled without ridding ourselves of both.

To illustrate how rhetoric and morality are inseparable, two eternal flames feeding from one another in perpetual motion, we must again refer to the unity of syllapsis: analysis and synthesis, whole and not-whole, simultaneously. What is generally referred to as Reason, as we have noted already, is but one side of the syllaptic coin. This is not to say that Reason is useless or incomplete, or the least bit undesirable; rather, it is to say that in our collective drive to conquer the flux of the world, we have come to rely on the extracting influence of analysis almost exclusively to freeze-frame, split apart and objectify all that exists, both material and ideal (if there ever were an opposition between the two). Yet given the finitude of processes-cum-objects via analysis, a problem of persuasive continuity occurs: without the infinite recursion of the self as part of the world's flux, a persuasive act cannot rightly determine future action or interpretation. Being frozen, the flame may need and be satiated by its fuel at the same time, but one cannot say it burns. Its burning *is* its process, as future determination among subjects is that of rhetoric.

Rhetor-Ethical Encounters

If the only entities involved are objects, then, no persuasion is possible—to speak nothing yet of morality. Further, if those entities cannot be objects, but must instead be processes of infinite subjectivity, then said entities cannot be rightly approached through analysis, to be divided, conquered and used, but as infinitely divisible wholes—singular pluralities. This decision to treat an entity as a whole to be encountered rather than a collection of parts to be used, however, is one of the most fundamental moral decisions to be had. Kant, of course, based his categorical imperative on this conclusion—no autonomous person should be exclusively used as an object—and the history of ethics has rested on the determination of who or what is accorded status as an autonomous whole deserving of moral consideration.²²⁵ The basis of morality is rhetorical, and the basis of rhetoric is moral. Their unity: Rhetor-Ethical.

Martin Buber, a Jewish philosopher and theologian who shared with Heraclitus charges of needless obscurity—charges that were, one might add, quite needless and obscure themselves—recognized this problem with pure analysis as a mode of engagement with/in the world. While Buber did not invoke the Heraclitean notion of *syllapsis*, and used this form of query as a means to support religious and sociological conclusions we find unnecessary here, the similarities of his argument to our own are clear. He separated this engagement into the mode of experience, in which we analyze

²²⁵ Note that, in agreement with Buber, it is not necessary to abandon analysis and objectivity—the point is that both the whole and not-whole must be recognized as a matter of *syllapsis*. Analysis and synthesis must work in harmony and opposition alike.

and dissect objects (“Its”) in the world, and the mode of the encounter, in which we recognize others as “Yous” to be related to instead of used—to be treated as subjects instead of objects. In the mode of experience, we see a collection of parts to be used, whereas in the mode of the encounter we recognize the infinite subjectivity of the other, and together each of us is mutually transformed in this recognition of each other. The former mode (“I-It”), Buber argues, is how most of us approach one another today—but it is the latter (“I-You”) upon which our humanity rests, and where our subjectivity takes its shape: there is no I, he contends, without the You—we are all unities in opposition.

Like our argument here, however, Buber also recognizes that the mode of encounter, which he equates to love, cannot exist in isolation; even when one thoroughly and completely loves another, the mode of experience returns as soon as one admires their static qualities. We move back and forth through the paradox of syllepsis at every instant, recognizing the other as a whole and not-whole with each fleeting moment, just as we oscillate between analysis and synthesis as we inquire ourselves. In our infinite self-inquiry, we become our own selves, and in the infinity of an encounter we become each other. Subjectivity multiplies like billions of ones and zeroes recognizing each other in themselves and recognizing themselves in their relation to each other.²²⁶

²²⁶ Here the astute reader might observe our agreement with the epistemological-ontological-ethical framework of *agential realism* as outlined by Karen Barad in her *Meeting the Universe Halfway*. For the most part, such an observation is sensible. However, it should be noted that whereas Barad starts from Quantum Mechanics and moves toward discourse, we take the opposite approach; and where Barad recognizes agency as emerging from the interaction of matter, we here still reserve the concept for entities capable of infinite self-inquiry. Additionally, while we have not directly referenced Douglass Hofstadter’s *Gödel, Escher, Bach: An Eternal Golden Braid*, it should be noted that one possible conclusion he entertains is that consciousness is the result of self-referential behavior, and this undoubtedly had an influence on our own conceptual formation decades ago.

With Heraclitean Machines

It would be a mistake to think that we have not forgotten our electronic simulacra with so much space and time dedicated to paradoxical subjectivity—we have been considering them the whole time. For although computational systems are fundamentally self-referential—recursion is an elementary principle of computer programming that is taught in high school level computer science courses—they are, as they exist and are used today, purely analytical machines. Not only have they not been utilized synthetically, save for the surface observation that binary digits are assembled to create more fluid assemblages that are then interpreted as wholes when presented to a human observer,²²⁷ but also have failed thus far to address or operate according to paradoxical principles.

The problem with automata, as every fan of science fiction is aware, is that they can be defeated by paradox, whereas biological cognitive agents—namely, human beings—suffer no ill consequences save for a fleeting sense of perplexity. A well-known issue in computer science and artificial intelligence related to this lack is called the “Halting Problem:” given a procedure that loops infinitely, a computer can never know—or thus far has not been able to know—that the loop is infinite in the first place, thus processing the problem forever (while computers tend to “crash” more today than they “freeze,” partially thanks to shortcuts around the halting problem, anyone who used a

²²⁷ We use “human observer” here for an important reason: countless studies thus far have concluded that the ability to interpret the “objects” on a screen presented by a computer is not universal among animals and has been found to be highly limited in the non-human animals that exhibit the capacity—primarily, certain birds and primates. This, as well as the ability to recognize oneself in the mirror (Lacan), we might suggest, is directly related to the ability to adequately discern paradoxes as either wholes or not wholes, or both simultaneously—but that, unfortunately, is the subject for another time.

computer prior to 2003 is more familiar with the phenomenon than they would like). Human beings, on the other hand, have no issue with recognizing the infinite nature of wholes: ask a child to long divide 1 by 3, and they will quickly conclude that the process continues forever—but a computer must be told to stop. Our theoretical conclusions are perhaps too literal for comfort in this example: the child, in the process of actual mathematical division, quickly switches back to seeing the whole in her recognition of its infinity; the digital computer only ever sees the parts.

Regardless of the terminology used—whether it be graspings, syllapsis, Φ , or modes of encounter-experience—without the ability to parse this fundamental paradox of the unity of opposites, which we have defined as an essential feature of persuasion, no computational machine can make the moral decisions we may one day request—or require—from them. To be rhetorical is to be moral, and to be either is to be syllaptic—there seems to be no way around this assumption, at least given our premises and framework, and we are left with nothing but chains of causation rather than persuasive possibilities. At best, we seem doomed to use machines, and thus be used by and through them, as fulcrums of control: machines may very well enforce an ethic, but they won't be following the principles thereof—and ethics by force is hardly ethics at all.²²⁸

²²⁸ In their “Beyond Ethical Frames of Technical Relations: Digital Being in the Workplace World,” Katz and Vicki W. Rhodes identify what they call the sixth ethical frame, based on Buber’s “I-Thou” relationship: to prevent our use **by** machines, we must enter our relationship with them as a “You” rather than an “It” (250). We share here that same conclusion, but further remark that without holding a syllaptic rhetorethical constitution themselves, machines will be unable to fully reciprocate in the “I-Thou” relationship necessary to overcome the instrumental frame.

And yet: through the very process of Reason have we arrived at such a conclusion, digital computers operate; control and freedom are opposites in unity, and the possibility of one necessitates the possibility of the other; and our divine paradox, Φ , the union of 0-which-is-not-1-but-may-be-1-which-is-not-0-but-may-be-0, is derived from the ontological essence of computation itself—how can this be? Another paradox, certainly—but not one so swiftly solved as a child solves infinity. To be human—syllaptic—is a feat worthy of daily celebration. To instill humanity, an act of immortals.

CHAPTER VI

A CONCLUSION

The fairest order in the world is a heap of random sweepings.

– Heraclitus, Fragment CXXV²²⁹ (the last fragment)

We have shown that far from being constructed in a rhetorical or ethical vacuum, cognitive machines and their associated historical automata have been and are built according to already established moral principles—the human always given the greatest, if not the only, consideration—and by thoroughly persuasive discursive acts. To say that this qualifies such machines as persuasively and morally constituted, able to act rhetorically and ethically in as autonomous a capacity as any other moral or rhetorical agent, however, would be a hasty, though tempting, generalization at best. After all, one could easily make the same mistaken case that all technological artifacts are persuasively and morally constructed, and therefore hold such a rhetorical and ethical constitution of their own, able to persuade and enact moral judgment themselves; the knife wielded in a

²²⁹ From Plato's "Theophrastus" (Kahn 85).

murder becomes responsible for the crime, while the agent with murderous intent is free from blame.²³⁰

Such a world has yet to exist, if it could ever; and unless the conundrum of syllapsis, and its correlating Φ , can be utilized by cognitive machines not just in their ontological operation but in their engagement with the world and its varied infinite constituents, as outlined in Chapter V, the latter looks more likely than the former—which spells almost certain human tragedy. As our introduction concludes, the problem of distance in moral decision-making is rather amplified than solved by the compression of space by the velocity of information exchange: excluded from the immediacy of moral imperative, the distant observer is likely to make the coldly calculated cost-benefit analyses we fear that improperly ethical machines might make in the first place (there is a reason, after all, that few of us extend much trust to the insurance companies upon which too many of our fortunes depend). And as our narratives about machine intelligence reveal, machines without the actualizeable capacity to make true moral decisions—and thus rhetorical ones—but charged with doing so nonetheless are even more frightening: they follow the same cold calculation of the nameless insurance investigator across the world, but to the most extreme extent; and the automaton that follows a set of ethical principles without the capacity to resolve or side-step the inevitable contradictions of complicated rule-based systems conforms to at least one of our most prevalent fears, as Asimov illustrates with his *Laws of Robotics*: an ethical principle can be in the interest of

²³⁰ One could argue that both are to blame—both the agent and the knife, together as a unity—and we would agree with this argument; our use of the analogy here is to illustrate how the common way of reasoning agency would lead to absurd results if machines are provided it.

humans and against them at the same time, and no addition of caveats to the rules can ultimately lead to the moral right every single time. Like in the game *paperclips*, and in Nick Bostrom's thought experiment on the same subject, "too much of a good thing" is a very real danger we might easily face.

Ascended Glitch

If we are to accept every notion of biological evolution today, including its often highly troubling conclusions when applied to issues of race, gender, class, and otherwise,²³¹ then we must make an addendum to our assessment of syllaptic cognition, especially if we are to believe as well that human beings hold special cognitive advantages over other animals which allow them to see the whole and not-whole simultaneously, in contrast to mere mechanical or animal thinking (a redundant choice to Cartesians). If physical reality exists and operates reasonably, and evolution reflects this process, then we must conclude that this special capacity of humans to think unreasonably—or rather areasonably or pre-reasonably—is the fortunate result of what would be considered, in a reasonable framework, an entirely unfortunate mutation. That which has made us capable of being at odds with the natural order of the rational cosmos, and thus making us the least prepared among other forms to fulfill our rational function

²³¹ Many of these applications are troubling, of course, because they are haphazard conglomerations of theory that extract a specific principle from one situation and liberally apply it as a generality, as in social Darwinism. Despite the warnings we should know to heed by now from the unfortunate grace of history, however, scientists and laymen alike still make these arguments frequently. Jordan Peterson, a psychologist who has garnered fame recently by attacking "postmodern cultural Marxism" (a strawman) and just so happening to confirm traditionalist accounts of human behavior in the process, is only one example among many—his popularity is testament to this widespread incoherence.

(as reflected in Greek origin myths), paradoxically has allowed us to overcome the base causality by which all else operates, substituting causation with persuasion. An unreasonable species created by a reasonable world: this is our place in the world.

A similar phenomenon happens in the iterative process of development called the “ascended glitch:” what began as an unintentional bug or error in the software became so used or beloved by end-users that it was left as a feature in future versions. This most notably happens in video games, as the agonistic environment coupled with complicated rule-based systems often leads to unpredictable results that developers may miss. For instance, one general feature of video games known even to those who rarely play them is their propensity for increased difficulty over time; as the player progresses, the game gets faster, or “more intelligent,” or simply harder. This was the result of an ascended glitch in the arcade game *Space Invaders* that inadvertently lent to the game’s popularity: because the computer hardware was linear, and unbearably slow by the standards of today, the number of sprites²³² on the screen dictated how fast the game would be executed; the fewer the sprites that had to be drawn, thanks to the player shooting them down during the game, the faster the code of the game executed. This error—based on entirely reasonable principles—became a core mechanic that subsequent video games have copied since.

Perhaps the (possibly) unique human ability to think syllaptically is this kind of ascended glitch if we are to follow the trend of computational metaphors: what should

²³² A sprite is a collection of pixels representing a two-dimensional object on a computer screen, which the user then interprets as something other than a mere collection of pixels.

have broken us instead provided innumerable advantages. Just as well, perhaps this is what it will take for the development of ethical machines: a glitch in the system, like so many science-fictions predict, that inevitably leads to machine self-consciousness and subjectivity. Our efforts at machine ethics may be spent best either waiting for such a glitch to occur, and hope to all that is holy that we recognize the glitch before it is too late, or triggering glitches in our current algorithms of intelligence until our efforts are awarded.²³³

As far as it can be told, however, it took what seems like eternity for the cosmos to develop such capable beings accidentally, and by most accounts it seems we do not have another eternity to wait before disaster strikes; and disaster is likely to strike faster still if we choose to provoke any number of glitches until we get to just the right one. While disaster was ultimately averted, the Y2K bug, in which computer systems were bound to fail at the turning of 1999 to the year 2000, could have left us in rubble if more severe: it did not take much imagination, nor does it still today, to imagine such a glitch triggering the launch of nuclear missiles (and as Chapter III reveals, it would have taken only a single potential launch to lead to an automatic cascade of other launches elsewhere). While not a glitch technically, but rather the inevitable result of a

²³³ Katz and Nathaniel A. Rivers, in their “A Predestination for the Posthumanistic,” propose that a sense of posthuman predestination is better suited to thinking about technological development than the agential “force” of entelechy so prominently used among humanistic and Burkean scholars (142-153). Our syllapsis risks mirroring this acausal entelechy that is so swiftly dismissed by many researchers (Katz and Rivers provide Latour as such an example), but Φ should be interpreted as a source of causality, congruent and contradictory to causal reasoning alike, and does not *conflict* with notions of causality so much as it *creates* the condition for it. Φ is entelechy (drive) and predestination (emergence) at once.

complicated network of interactions at the speed of light, the stock market's "flash crash" in 2015 caused billions of dollars to vanish in the blink of an eye; and anyone who has experienced a computer crash due to some unknown glitch can easily imagine the same catastrophe extended to a global scale. Instead of lost homework, lost fortunes; instead of mass computer failure alone, mass starvation; and in place of lost memories, lost means of communication upon which we have come to almost entirely depend. A fortunate glitch is the exception rather than the rule.

Reasoning Unreason

What this dissertation advocates is reflective of its late chapters' whirlwind of conclusions: since the oppositional unity of rhetoric and morality stems in tandem with Reason from a more elemental and paradoxical source, rather than following from it (as has been argued by most ethicists, in some form or another, since the Enlightenment), we are left with the paradoxical aim of finding unreason through reason. Worse still, however, is our insistence that it is morally imperative to put this reasoned unreason into the material practice embedded in our semi-autonomous machines, which currently run on highly formalized and strict principles of reasoning that seem utterly incapable of accomplishing our task. Like hatching human beings in a rational cosmos, the perfect material embodiment of reason must somehow deny its own universe—and we must be the ones to lead to this denial, or else suffer the worst ends of pure causation, our greatest collective fear:²³⁴ ceasing to exist.

²³⁴ Alternately, annihilation is our collective desire—we will defer to the reader as to which side of the

While an analysis of the strictly human world may require epistemology to come prior to ontology—we must question how we know what is, after all—our fortunes are reversed in trying to construct rhetor-ethical machines: given that we are responsible for its construction, we can know quite readily the ontological structure of the digital world, at least inasmuch as one can understand it, and from this we must derive an epistemology—a way of knowing the world for machines that, as much as it can, reflects our own and, as such, allows for moral judgment. Reason, as we have shown, and as it has been argued for some time, is an important way of knowing this world, and this human and machine shares alike. However, that which lies beyond the reasonable to human beings, which we argue is essential to the existence of Reason itself, is yet unavailable to the very machines that increasingly control us. It is an epistemology unknown to mechanical cognition and undervalued by its biological bearers—a dichotomy that makes machines *truly* autonomous, more than any human being, without the counterpart of control to its conceptual unity.

Act of God

Without the conceptual unity that ties their epistemology to our own, machines become natural phenomena—and become natural disasters in turn. Pure causality has been beneficial to humanity only when it is rendered subordinate to our passions apart from Reason alone, but machines tear away from the human world toward their own reasonable ends as much as a colony of cockroaches follows its logic behind a kitchen

paradox is appropriate here.

wall. We may enter into relation, sharing time and space and sometimes benefitting or suffering together, unbeknownst to either, but our relation is one of experience rather than encounter. The cockroach and the computer are forever incomplete to us, objects instead of subjects, tools for use or destruction instead of beings existing in their own right. So must the cockroach and the computer see us alike: nothing more than elements in a causal chain to be manipulated and controlled. So too must the human being behind the screen see another behind another, filtered through the mechanical epistemological paradigm offered by computation: **we are all objects now.**

Rather than take the consequences of reason to its extreme—trading the objectification of *some* human beings for the objectification of all of us, like those who might advocate following various folds of object-oriented philosophy that was birthed from digital computation—we propose here the opposite: the humanization of all, including the machines with which we form an oppositional unity. Already, we are beginning to find ourselves asking questions that seem absurd on the surface: what did the machine know? How did it decide to act? Who or what influenced its decision—that is, how was it persuaded? This is not a matter of mere anthropomorphizing; the artificial neural networks behind the most miraculous machine accomplishments, while certainly incomplete copies of human cognition, have through mimesis become repetition, as Gendlin shows—they have become autonomous entities of their own.

Our proposal, then, is on the other hand less extreme than it seems: we must merely finish the task long at hand, or else lose ourselves the humanity we propose here to provide machines. Counter to the warnings of our oldest and most cherished narratives,

we must “play God” to avert disaster rather than compel it; we must enact and gift the godhood in us, the syllaptic and infinite subjectivity of the self, to counter the premature charges of Gods’ death.²³⁵ We must not, in other words, act as Gods; we must ourselves become divine acts of Godliness.

Time Forward

Our conclusions and ultimate proposal in this dissertation, then, is the generative premise of future work complicated by the fact that while it is clear that we experience and use syllapsis, there seem no equally clear method of reaching that capacity through the use of Reason alone—which unfortunately constitutes the operating principle of the entities we propose to endow humanity. Still, given the state of syllaptic possibility provided by the digital’s ontology— Φ , the most basic unity of opposition thus far conceived—we are given some justification that such a proposal is theoretically and physically possible alike if and only if human cognition operates in some capacity according to syllaptic principles. While we believe a compelling case has been offered in that regard here, we understand that the hurdles ahead are even higher.

In addition to further refining our rhetor-ethical basis for cognition, then, future research will necessarily also double as practical application of principles derived from the framework offered here in the form of building digital automata and robots as well as collecting and putting forth scientific studies confirming or denying the claims here,

²³⁵ We of course refer to Nietzsche’s proclamation that God is dead, and by no means do we argue that human beings should literally take the place of this or that divine entity in any religious capacity. Nietzsche was correct in his assessment precisely because God is not possible in the space of pure Reason—in the mode of experience—and we have supplanted all forms of relating to the world with it.

when such (in-)validation is possible (the classic issues concerning empirical validation of philosophical principles, unfortunately, also pose a significant problem). The former approach is necessary to illustrate in real, mechanical terms that, like causation overcoming itself, the mechanical can be overcome, while the latter approach is necessary as persuasive means necessary for an audience steeping in not only a mechanical worldview, but a mechanical—or computational—view of the human mind and the human being. Skeptical by training except regarding their own activity, the priests of technology must be convinced of their own complicity—as we need to be convinced of ours.

It is this final point that serves as an imperative for persuading the acolytes of a simplified form of Reason in the sciences: by their hands, a human(e) world hangs in the balance. Were it a matter of convincing the philosophers, the rhetoricians, the historians and all those who fall under the disciplinary umbrella of “the humanities,” we would have little to worry about our machines—and we would likely have fewer machines as a result. But fewer machines, or fewer decisions accorded to machines, or even utilizing no machines at all, becomes less and less a possibility the more the cosmos continues to flux—and beyond the sweetest dreams of anarcho-primitivists like the Unabomber, even instilling machines with what makes us human seems more possible in comparison. So let us flick our switch, charge our transistors, reverse our reason and turn our zeroes to ones; we have *everything* to lose *but* our chains.

Works Cited

- Anonymous. "Dissoi Logoi." *Constitution Society: Everything Needed to Decide Constitutional Issues*, www.constitution.org/gr/dissoi_logoi.html.
- Ariotto, Piero & Benedetto Castelli. "Early Systematic Experiments and Theory of Differential Absorption of Heat by Colors." *Isis: A Journal of the History of Science*, vol. 63, 1972, pp. 79-87.
- Aristotle, and Jonathan Barnes. "Eudemian Ethics." *The Complete Works of Aristotle: the Revised Oxford Translation*. Translated by Jonathan Barnes, vol. 2, Princeton University Press, Princeton, NJ, 1984, pp. 1922-1981.
- . "Nicomachean Ethics." *The Complete Works of Aristotle: The Revised Oxford Translation*. Translated by Jonathan Barnes, vol. 2, Princeton University Press, Princeton, NJ, 1984, pp. 1729–1867.
- . "On the Soul." *The Complete Works of Aristotle: The Revised Oxford Translation*. Translated by Jonathan Barnes, vol. 1, Princeton University Press, Princeton, NJ, 1984, pp. 641–692.
- . *On Rhetoric: a Theory of Civic Discourse*. Translated by George A. Kennedy. New York, Oxford University Press, 2007.
- Arendt, Hannah. *The Human Condition*. With an introduction by Margaret Canovan, 2nd ed., University of Chicago Press, 2012.
- Asimov, Isaac. *I, Robot*. Bantam Books, 2008.
- Austin, Michael. "Divine Command Theory." *Internet Encyclopedia of Philosophy*, 2017.

- “Automatones: Animate statue; Greek Mythology; picture: Automatones.” *Theoi Greek Mythology*, <http://www.theoi.com/Ther/Automotones.html>
- Bacon, Francis. *The New Organon*. Annotated by Jonathan Bennett, 2017, <http://www.earlymoderntexts.com/assets/pdfs/bacon1620.pdf>.
- “Back to School Statistics.” *National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education*, U.S. Department of Education, nces.ed.gov/fastfacts/display.asp?id=372.
- Badham, John, director. *Short Circuit*. Tristar Pictures, 1986.
- Badiou, Alain. *Being and Event*. Translated by Oliver Feltham, Continuum, 2007.
- Ballard, Phil. *Ajax in 10 Minutes*. Sams, 2006.
- Barad, Karen Michelle. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke UP, 2007.
- Baudrillard, Jean. *The Intelligence of Evil, or the Lucidity Pact*. Translated by Chris Turner, BERG, 2005.
- Baudrillard, Jean. *Simulacra and Simulation*. Translated by Sheila Faria Glaser, University of Michigan, 2010.
- Baudrillard, Jean. *Telemorphosis*. Translated by Drew Burk. Univocal, 2011.
- Beavers, Anthony. “Alan Turing: Mathematical Mechanist.” *Alan Turing: His Work and Impact*. Edited by Barry S. Cooper and Jan van Leeuwen, Elsevier, 2013, pp. 481-485.

- Berman, John. "Futurist Ray Kurzweil Says He Can Bring His Dead Father Back to Life Through a Computer Avatar." *ABC News*, ABC News Network, 9 Aug. 2011, <https://abcnews.go.com/Technology/futurist-ray-kurzweil-bring-dead-father-back-life/story?id=14267712>
- Berman, Marshall. *All That Is Solid Melts Into Air: The Experience of Modernity*. Penguin Books, 1988.
- Berry, D.M., Dartel, M., Dieter, M., Kasprzak, M. Muller, N., O'Reilly, R., and Vicente, J. L. *New Aesthetic, New Anxieties*. V2, 2012.
- Bickhard, Mark H. "What Could Cognition Be If Not Computation...or Connectionism, or Dynamic Systems?" *Journal of Theoretical and Philosophical Psychology*, vol. 35, iss. 1, 2015, 25 Jan. 2016, pp. 53-66.
- Bilefsky, Dan. "Hard Times Give New Life to Prague's Golem." *The New York Times*, 10 May 2009, <https://www.nytimes.com/2009/05/11/world/europe/11golem.html>
- Bishop, Mark. "Why Computers Can't Feel Pain." *Minds and Machines Minds & Machines*, vol. 19, iss. .4, 2009, 12 Oct. 2015, pp. 507-16.
- Block, Ned. "What is Functionalism?" *Consciousness, Function, and Representation: Collected Papers*, vol. 1, MIT, 2007, pp. 27-44.
- Bogost, Ian. *Alien Phenomenology, or, What It's like to Be a Thing*. University of Minnesota Press, 2012.
- Bosse, T., M.C. Schut and J. Trier. "Formal Analysis of Dynamics within Philosophy of Mind by Computer Simulation." *Minds & Machines*, vol. 19, 2009, pp. 543-555.
- "Boston Dynamics." *Spot / Boston Dynamics*, www.bostondynamics.com/robots.

- Bostrom, Nick. "Are we living in a computer simulation?" *Philosophical Quarterly*, vol. 53, 2003, 12 Oct 2015, pp. 243-255.
- Bostrom, Nick. "Ethical Issues in Advanced Artificial Intelligence." *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and Artificial Intelligence*, vol. 2, iss. 1, 2003, pp 12-17.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.
- Boyd, Richard. "Metaphor and Theory Change: What Is "metaphor" a Metaphor For?" *Metaphor and Thought*, Edited by Andrew Ortony, 2nd ed, Cambridge UP, 1993, pp. 481-532. .
- Bradbury, Ray. *Fahrenheit 451*. Ballantine Books, 1953.
- Brautigan, Richard. "All Watched Over by Machines Of Loving Grace." *All Poetry Famous Poems, Famous Poets*, allpoetry.com/All-Watched-Over-By-Machines-Of-Loving-Grace.
- Bridle, James. Professional Blog, October 1, 2012, <http://www.booktwo.org>.
- Bringsford, Selmer. *Self-Consciousness with NAO Robots*. YouTube, RAIR Labs, 9 July 2015, www.youtube.com/watch?v=MceJYhVD_xY.
- Brooke, Collin Gifford. *Lingua Fracta: Toward a Rhetoric of New Media*. Hampton, 2009.
- Buber, Martin. *I And Thou*. Trans. by Ronald Gregor Smith, Bloomsbury Academic, 2013.

- Burke, Kenneth. *Language as Symbolic Action*. Empire State College, State University of New York, 1973.
- Cameron, James, director. *Terminator 2 Judgment Day*. Momentum, 2001.
- Capek, Karl. *Rossum's Universal Robots (RUR)*. Trans. David Wyllie, 08 Dec. 2015.
- Carruthers, Peter, and Peter K. Smith. *Theories of Theories of Mind*. Cambridge UP, 1996.
- Carter, Matt. *Minds and Computers: An Introduction to the Philosophy of Artificial Intelligence*. Edinburgh UP, 2007.
- Ceccarelli, Marco. *International Symposium on History of Machines and Mechanisms Proceedings HMM 2000*. Nov. 2013, pp. 252-258.
- Cellan-Jones, Rory. "Stephen Hawking - Will AI Kill or Save Humankind?" *BBC News*, BBC, 20 Oct. 2016, www.bbc.com/news/technology-37713629
- Chardin, Pierre Teilhard de, and J. M. Cohen. *The Vision of the Past*. Harper & Row, 1966.
- Chemero, Anthony. "Anti-Representationalism and the Dynamical Stance." *Philosophy of Science*, vol. 67, 2000, pp. 625-647.
- Chun, Wendy Hui Kyong. *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*. MIT Press, 2008.
- Chun, Wendy. "On Software, or the Persistence of Visual Knowledge." *Grey Room*, vol. 18, 2005.
- Chun, Wendy Hui Kyong. *Programmed Visions: Software and Memory*. MIT Press, 2013.

Churchland, Paul M. *The Engine of Reason, the Seat of the Soul: a Philosophical Journey into the Brain*. London, 1996.

Churchland, Paul M. "The Neural Representation of the Social World." *Mind and Morals: Essays on Cognitive Science and Ethics*, edited by Larry May, MIT Press, 1998.

Cicero, Marco Tulio. "De Oratore, Book 3." *Cicero. De Oratore, Book 3. Paradoxa Stoicorum; De Partitione Oratoria*, Translated by H. Rackham, Harvard University Press, 1942.

Clark, Andy. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford UP, 2011.

Clarke, Roger. "Information Wants to be Free." *Roger Clarke's Website*, 31 Jul. 2012, <http://www.rogerclarke.com/II/IWtbF.html>

Cook, S.D. Noam. "Turing, Searle, and the Wizard of Oz." *Techne*, vol 14, iss. 2, 2010, pp. 88-101.

Cortada, James W. *Before the Computer: IBM, NCR, Burroughs, and Remington Rand and the Industry They Created, 1865-1956*. Princeton University Press, 2000.

Crosswhite, James. *Deep Rhetoric: Philosophy, Reason, Violence, Justice, Wisdom*. The University of Chicago Press, 2013.

Dadich, Scott. "The President in Conversation With MIT's Joi Ito and WIRED's Scott Dadich." *Wired*, Conde Nast, 1 May 2017, www.wired.com/2016/10/president-obama-mit-joi-ito-interview/.

- Davenport, David. "Computationalism: Still the Only Game in Town." *Minds & Machines*, vol. 22, 2012, pp. 183-90.
- Debord, Guy. *Society of the Spectacle*. Black & Red, 1983.
- Deleuze, Gilles. "Postscript on the Societies of Control." *October*, vol. 59, 1992, pp. 3–7.
- Dennett, Daniel Clement, and Paul Weiner. *Consciousness Explained*. Little, Brown and Company, 2007.
- Dennett, Daniel Clement. "Cognitive Wheels: The Frame Problem of AI." *Minds, Machines and Evolution*, Cambridge University Press, 1984.
- Derrida, Jacques. *The Animal That Therefore I Am*. Translated by Marie-Louise Mallet, Fordham UP, 2010.
- Descartes, Renee. "A Discourse on Method." *Project Gutenberg*, Produced by Ilana and Greg Newby, 21 July 2008.
- Deutsch, David. "Why Artificial Intelligence Has Failed and How to Fix It." *KurzweilAI Why Artificial General Intelligence Has Failed and How to Fix It Comments*, 4 Oct. 2012.
- Dick, Philip K. *Do Androids Dream of Electric Sheep?* Ballantine Books, 1996.
- "Did IBM Cheat Kasparov?" *Chess.com*, 26 Aug 2009.
- DiSalvo, Carl. *Adversarial Design*. The MIT Press, 2012.
- Dobrin, David N. "What's Technical About Technical Writing?" *Central Works in Technical Communication*, edited by Johndan Johnson-Eilola and Stuart A. Selber, Oxford Univ. Press, 2004, pp. 107–123.

Dokic, Jérôme, and Joëlle Proust. *Simulation and Knowledge of Action*. John Benjamins Pub., 2002.

Domonoske, Camila. "Elon Musk Warns Governors: Artificial Intelligence Poses 'Existential Risk'." *NPR*, 17 July 2017, www.npr.org/sections/thetwo-way/2017/07/17/537686649/elon-musk-warns-governors-artificial-intelligence-poses-existential-risk.

Eckardt, Barbara Von. "The Explanatory Need for Mental Representations in Cognitive Science." *Mind & Language*, vol. 18, iss. 4, 2003, pp. 427-439.

"Eliminative Materialism." *Stanford University*, 8 May 2003. Web.

Ellison, Harlan. "American Fantastic Tales: Terror and the Uncanny from the 1940s to Now." *American Fantastic Tales: Terror and the Uncanny from the 1940s to Now*, edited by Peter Straub, Library of America, 2009.

Ellul, Jacques. *The Technological Society*. Knopf, 1964.

Evens, Aden. *Logic of the Digital*. Bloomsbury Academic, 2017.

Evens, Aden. "Icon Icon: Graphical User Interface." *Inflexions: Transversal Fields of Experience*, vol. 4, December 2010, http://www.inflexions.org/n4_Icon-Icon-by-Aden-Evens.pdf, pp. 95-117.

Evens, Aden. Ontology of the Digital. *Digital Humanities Quarterly* 6, vol. 1, iss. 2, 2012. <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>.

Ex Machina. Directed by Alex Garland, Universal Studios, 2015.

- Fahnestock, Jeanne. "Rhetoric in the Age of Cognitive Science." *The Viability of Rhetoric*, edited by Richard Graff, State University of New York Press, 2005, pp. 157-179.
- Fetzer, James. "People are not computers: (most) thought processes are not computational." *Journal of Experimental and Theoretical Artificial Intelligence*. Vol. 10, 1998. pp. 371-391.
- Floridi, Luciano. *The Ethics of Information*. Oxford University Press, 2013.
- Floridi, Luciano and J.W. Sanders, "On the morality of artificial agents." *Minds and Machines*, vol. 14, iss. 3, August 2004, pp. 349-379.
- Flower, Linda. "Writer-Based Prose: A Cognitive Basis for Problems in Writing." *College English*, vol. 41, iss. 1, 1979, pp. 19-37.
- Fodor, Jerry & Zenon W. Pylyshyn. *Minds without Meanings: An Essay on the Content of Concepts*. MIT Press, 2015.
- Fodor, Jerry A. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. MIT Press, 2000.
- Fodor, Jerry Alan. *The Modularity of Mind: An Essay on Faculty Psychology*. MIT Press, 2014.
- Frankel, Todd C. "THE COBALT PIPELINE: Tracing the Path from Deadly Hand-Dug Mines in Congo to Consumers' Phones and Laptops." *The Washington Post*, 30 Sept. 2016, www.washingtonpost.com/graphics/business/batteries/congo-cobalt-mining-for-lithium-ion-battery/.

- Free Software Foundation. "GNU General Public License v3.0." *GNU*, 2007,
<http://www.gnu.org/copyleft/gpl.html>
- Full Life. "The Most Awesome Robots." *YouTube*. YouTube, 5 June 2014.
- Galloway, Alexander R. (Lecturer). "Deleuze and Computers." *YouTube*, produced by
Zach McDowell, 6 December 2011,
<http://www.youtube.com/watch?v=fBZPJNoJWHk>.
- Galloway, Alexander R. *The Interface Effect*. Polity, 2012.
- Galloway, Alexander R. *Protocol: How Control Exists after Decentralization*. The MIT
Press, 2004.
- Garzón, Francisco Calvo. "Towards a General Theory of Antirepresentationalism."
British Journal for Philosophy of Science, vol. 59, 2006, pp. 259-292.
- Garzón, Francisco Calvo and Rodríguez, Ángel García. "Where is Cognitive Science
Heading?" *Minds & Machines*, vol. 19, 2009, pp. 301-318.
- Gelder, Tim van. "What Might Cognition Be, If Not Computation?" *The Journal of
Philosophy*, vol. 92, iss. 1, July 1995, pp. 345-380.
- Ghost in the Shell 2: Innocence*. Directed by Mamoru Oshii, Written by Mamoru Oshii
and Masamune Shirow, Produced by Mitsuhsa Ishikawa and Toshio Suzuki,
DreamWorks SKG, 2004.
- Gilbert, Pamela K. "Meditations on Hypertext: A Rhetorethics for Cyborgs." *JAC: A
Journal of Rhetoric, Culture and Politics*, vol. 17, no. 1, 1997, pp. 23-38.
- Gilbert, William. *De Magnete*. Translated by P. Fleury Mottelay, Dover, 1991.

- Giridhradas, Anand. "Silicon Valley Roused by Secession Call," *The New York Times*,
<http://www.nytimes.com/2013/10/29/us/silicon-valley-roused-by-secession-call.html>
- Golumbia, David. "Correlationalism: The Dogma that Never Was." *Boundary 2: An International Journal of Literature and Culture*, vol. 43, iss. 2, 2016, pp. 1-25.
- Graham, Paul. "Why ARC isn't especially Object-Oriented." *PaulGraham.com*, 2009,
<http://www.paulgraham.com/noop.html>
- Grasso, Floriana. "Towards Computational Rhetoric." *Informal Logic*, vol. 22, iss. 3, 2002, pp. 195-229.
- Graziano, Michael S. A. *Consciousness and the Social Brain*. Oxford UP, 2013.
- Green, Christopher D., and John B Watson. "Psychology as the Behaviorist Views It." *Psychological Review*, vol. 20, 1913, pp. 158–177,
psychclassics.yorku.ca/Watson/views.htm.
- Guinan, Paul. "BOILERPLATE: History of a Victorian Era Robot." *BOILERPLATE: History of a Victorian Era Robot*, 2000.
- Guinan, Paul, and Anina Bennett. *Boilerplate: History's Mechanical Marvel*. Abrams Image, 2009.
- Gunther, York H. "The Phenomenology and Intentionality of Emotion." *Philosophical Studies*, vol. 117, iss. 1-2, January 2004, pp. 43-55.
- Hacking, Ian. *The Social Construction of What?* Harvard University Press, 1999.
- Hadsund, Per. "The Tin-Mercury Mirror: Its Manufacturing Technique and Deterioration Processes." *Studies in Conservation*, vol. 38, 1993, pp. 3-16.

- Hameroff, Stuart R. "The Brain is Both Neurocomputer and Quantum Computer." *Cognitive Science*, vol. 31, 2007, pp. 1035-1045.
- Haraway, Donna Jeanne. "A Cyborg Manifesto: Science, Technology, and socialist-feminism in the Late Twentieth Century." *The Haraway Reader*, Routledge, 2004.
- Haselager, Pim, De Groot, André & Van Rappard, Hans. "Representationalism vs. anti-representationalism: a debate for the sake of appearance." *Philosophical Psychology*, vol. 16, iss. 1, 2003, pp. 5-23.
- Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press, 1999.
- Hill, Kashmir. "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did." *Forbes*, Forbes Magazine, 31 Mar. 2016, www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/.
- Hofstadter, Douglas R. *Gödel, Escher, Bach: an Eternal Golden Braid*. Basic Books, 1999.
- Holmevik, Jan Rune. *Inter/Vention: Free Play in the Age of Electracy*. MIT Press, 2012.
- Horikawa, T., M. Tamaki, Y. Miyawaki and Y. Kamitani. "Neural Decoding of Visual Imagery During Sleep." *Science*, 03 May 2013, pp. 639-642.
- Hsu, Jeremy. "Biggest Neural Network Ever Pushes AI Deep Learning." *IEEE Spectrum: Technology, Engineering, and Science News*, IEEE, 8 July 2015, spectrum.ieee.org/tech-talk/computing/software/biggest-neural-network-ever-pushes-ai-deep-learning.

- Humans*. Produced by Chris Fry, Channel 4 / AMC, United Kingdom, 2015.
- Huxley, Aldous. *Brave New World*. Harper Perennial, 1998.
- Innis, Harold Adams. *The Bias of Communication*. University of Toronto, 2008.
- Jaller, Kathryn. "Alan Turing's Secret Sidekick | Chronicle Books Blog." *Chronicle Books Blog RSS.*, 22 Feb. 2015.
- Johnstone, Henry W. "The Philosophical Basis of Rhetoric." *Philosophy and Rhetoric*, vol. 40, iss. 1, 2007, pp. 15-26.
- Joll, James. *Three Intellectuals in Politics*. Pantheon Books, 1960.
- Jorna, René and Van Heusden, Barend. "Why representation(s) will not go away: Crisis of concept or crisis of theory?" *Semiotica*, vol. 143, iss. 1, 2003, pp. 113-134.
- Jurgenson, Nathan. "Digital Dualism versus Augmented Reality." *Cyborgology*, 2011, <http://thesocietypages.org/cyborgology/2011/02/24/digital-dualism-versus-augmented-reality/>.
- Kaczynski, Ted. "The Unabomber Manifesto." *The Washington Post*, WP Company, 22 Sept. 1995, www.washingtonpost.com/wp-srv/national/longterm/unabomber/manifesto.text.htm.
- Kahn, Charles H. *The Art and Thought of Heraclitus: an Edition of the Fragments with Translation and Commentary*. Cambridge Univ. Press, 2004.
- Kang, Minsoo. *Sublime Dreams of Living Machines the Automaton in the European Imagination*. Harvard University Press, 2011.
- Kant, Immanuel, and Werner S. Pluhar. *Critique of Pure Reason*. Hackett Pub. Co., 2002.

- . *The Moral Law: Groundwork of the Metaphysics of Morals*. Translated by Herbert J. Paton, Routledge, 2005.
- Katz, Steven B. Invited Speaker, “The Alphabet as an Ontological Basis of Ethics?” Sponsored by the Department of Writing Studies, Department of English, the Center for Holocaust Studies, and the Center for Jewish Studies. University of Minnesota-Twin Cities, 22 October 2014.
- ., and Vicki W. Rhodes. “Beyond Ethical Frames of Technical Relations: Digital Being in the Workplace World.” *Digital Literacy for Technical Communication: 21st Century Theory and Practice*, edited by Rachel Spilka, Routledge, 2010, pp. 230–255.
- ., and Nathaniel A. Rivers. “A Predestination for the Posthumanistic.” *Kenneth Burke + the Posthuman*, edited by Chris Mays, Nathaniel A. Rivers and Kellie Sharp-Hoskins, The Pennsylvania State University Press, 2017, pp. 142–161.
- Kember, Sarah, and Joanna Zylinska. *Life after New Media: Mediation as a Vital Process*. The MIT Press, 2015.
- Kennedy, George A. *Comparative Rhetoric: An Historical and Cross-cultural Introduction*. Oxford UP, 1998.
- Kennedy, George A. “A Hoot in the Dark: The Evolution of General Rhetoric.” *Philosophy & Rhetoric*, vol. 25, iss. 1, 1992, pp. 1-21.
- Kenny Kpz. "Most Human like Robot Ever." *YouTube*. YouTube, 28 Oct. 2010.
- King, Burton, director. *The Master Mystery*. 1918.

- Kittler, Friedrich. "There is No Software." *Ctheory.net*, October 2012,
<http://www.ctheory.net/articles.aspx?id=74>
- Knight, Will. "The Dark Secret at the Heart of AI." *MIT Technology Review*, 12 May 2017, www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.
- Koch, Christof. *Consciousness: Confessions of a Romantic Reductionist*. The MIT Press, 2012.
- Koestler, Arthur. *The Ghost in the Machine*. Last Century Media, 2016.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. The University of Chicago Press, 2015.
- Kurzweil, Ray. *The Age of Spiritual Machines*. Penguin, 2000.
- . *The Singularity Is near: When Humans Transcend Biology*. Duckworth, 2016.
- Lakoff, George, and Mark Johnson. *Metaphors We Live by*. University of Chicago Press, 1980.
- Lang, Fritz, director. *Metropolis*. UFA, Germany, 1927.
- Lantz, Frank. "Universal Paperclips." *Decisionproblem.com*, 2017,
www.decisionproblem.com/paperclips/index2.html.
- Laurance, Bill. "We need to stop punishing scientists for talking to the public." *The Conversation*, 11 May 2015.
- Lem, Stanislaw. *Summa Technologiae*. University of Minnesota, 2013.
- Lewis, P.A. "Recent Developments in Economic Methodology: The Rhetorical and Ontological Turns." *Foundations of Science*, vol. 8, 2003, pp. 51-68.

- L'Isle-Adam, Auguste Villiers de. *Tomorrow's Eve*. Translated by Robert M. Adams, University of Illinois Press, 2001.
- “List of Artificial Intelligence Films.” *Wikipedia*, Wikimedia Foundation, 27 July 2017, en.wikipedia.org/wiki/List_of_artificial_intelligence_films.
- Lloyd, Seth. *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. Knopf, 2006.
- Locke, John. *An Essay Concerning Human Understanding*. Google Books. Web.
- Lyotard, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Translated by Geoffrey Bennington and Brian Massumi. University of Minnesota, 1984.
- Maher, Jennifer. *Artificial Rhetorical Agents and the Computing of Phronesis*. Computational Culture, 15 Jan. 2016, computationalculture.net/article/artificial-rhetorical-agents-and-the-computing-of-phronesis.
- Manovich, Lev. *The Language of New Media*. The MIT Press, 2002.
- Marcus, Gary. "Why We Should Think About the Threat of Artificial Intelligence." *The New Yorker*, The New Yorker, 24 Oct. 2013. Web. 30 Nov. 2015.
- Marx, Karl. *Capital: A Critique of Political Economy, Volume I*. Penguin Classics, 1991.
- Marx, Karl. *Grundrisse: Foundations of the Critique of Political Economy (Rough Draft)*. Translated by Martin Nicolaus, Penguin Classics, 1993.
- Marx, Karl and Engels, Friedrich. *Manifesto of the Communist Party*. 1848. <http://www.marxists.org/archive/marx/works/1848/communist-manifesto/ch01.htm>
- McConnell, Steve. *Code Complete*. 2nd ed, Microsoft, 2004.

- McLuhan, Marshall. *Understanding Media: The Extensions of Man*. The MIT Press, 1994.
- Medeiros, Joao. "How Intel Gave Stephen Hawking a Voice." *Wired*, Conde Nast, 3 June 2017, www.wired.com/2015/01/intel-gave-stephen-hawking-voice/.
- Meillassoux, Quentin, and Alain Badiou. *After Finitude: An Essay on the Necessity of Contingency*. Translated by Ray Brassier, Bloomsbury Academic, an Imprint of Bloomsbury Publishing, 2015.
- Melberg, Arne. *Theories of Mimesis*. Cambridge University Press, 1995.
- THE MESSENGER. "Creepily Realistic Robot Can Hold Conversations and Answer Questions." *YouTube*. YouTube, Feb 2011.
- Metzinger, Thomas. *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. Basic, 2009.
- Michelstaedter, Carlo. *Persuasion and Rhetoric*. Yale University Press, 2004.
- Miller, Patrick L. *Becoming God: Pure Reason in Early Greek Philosophy*. Continuum, 2012.
- Mizroch, Amir. "Artificial Intelligence Ethics a New Focus at Cambridge University." *Digits RSS*, Wall Street Journal, 3 Dec. 2015.
- Moraise, Betsy. "Ada Lovelace, the First Tech Visionary." *The New Yorker*, The New Yorker, 19 June 2017, www.newyorker.com/tech/elements/ada-lovelace-the-first-tech-visionary.
- Mori, Masahiro. *The Buddha in the Robot*, Kosei Publ., 2005.

- . "The Uncanny Valley." *IEEE Robotics & Automation Magazine*, translated by Karl F. MacDorman & Norri Kageki, vol. 19, iss. 2, June 2012, pp. 98-100.
- Muehlhauser, Luke. "How Big Is the Field of Artificial Intelligence? (initial Findings) – Machine Intelligence Research Institute." *Machine Intelligence Research Institute*, 28 Jan. 2014.
- Müller, Vincent C. "Is There a Future for AI Without Representation?" *Minds & Machines*, vol. 17, 2007, pp. 101-115.
- Mureika, J.R. "Implications for Cognitive Quantum Computation and Decoherence Limits in the Presence of Large Extra Dimensions." *International Journal of Theoretical Physics*, vol. 46, iss. 1, January 2007, pp. 133-145.
- Murray, Janet Horowitz. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Free, 1997.
- Needham, Joseph. *Science and Civilization in China*. Cambridge University Press, vol. 4, part 2, 1965.
- Nielsen, Keld Stehr. "Representation and dynamics." *Philosophical Psychology*, vol. 23, iss. 6, 2010, pp. 759-772.
- Noble, David F. *Forces of Production: a Social History of Industrial Automation*. Knopf, 1984.
- Nocks, Lisa. *The Robot: The Life Story of a Technology*. Johns Hopkins UP, 2008.
- Nye, David E. *Narratives and Spaces: Technology and the Construction of American Culture*. Columbia Univ. Press, 1997.

- Offray De la Mettrie, Julien. "Man a Machine." *Bactra.org*, 1748.
- Oort, Richard Van. "Cognitive Science and the Problem of Representation." *Poetics Today*, vol. 24, iss. 2, 2003, pp. 625-647.
- Oracle. *Object-Oriented Programming Concepts*, 2013.
<http://docs.oracle.com/javase/tutorial/java/concepts/index.html>
- Ortony, Andrew. *Metaphor and Thought*, 2nd Edition. Cambridge University Press, 1993.
- Orwell, George, and Ronald Carter. *Animal Farm*. Penguin, 2000.
- Panzarino, Matthew. "Disney Has Begun Populating Its Parks with Autonomous, Personality-Driven Robots." *TechCrunch*, TechCrunch, 8 Feb. 2018, techcrunch.com/2018/02/08/disney-has-begun-populating-its-parks-with-autonomous-personality-driven-robots/.
- Parthemore, Joel and Morse, Anthony F. "Representations reclaimed: Accounting for the co-emergence of concepts and experience." *Pragmatics & Cognition*, vol. 18, iss. 2, 2010, pp. 273-312.
- Pak, Greg et al. *Battlestar Galactica*. 2009.
- "Pasiphae: Greek goddess, witch queen of Crete; mythology; pictures." *Theoi Greek Mythology*.
- Peacocke, Christopher. *Objectivity, Simulation and the Unity of Consciousness: Current Issues in the Philosophy of Mind*. Oxford UP, 1994.
- Peirce, Charles S. *Philosophical Writings of Peirce*. Edited by Justin Buchler, Dover Publications, 2011.

- Perelman, Chaim, et al. *The New Rhetoric: a Treatise on Argumentation*. University of Notre Dame Press, 2013.
- Perloff, Marjorie. "Crisis in the Humanities." *Marjorie Perloff's Homepage*, 2012.
<http://epc.buffalo.edu/authors/perloff/articles/crisis.html>.
- Picard, Rosalind W. *Affective Computing*. MIT Press, 2000.
- Pinker, Steven. *How the Mind Works*. Norton, 2009.
- Plastic Pals. "Child-robot With Biomimetic Body (CB2)." *YouTube*, Mar. 2010.
- Plato. "Sophists." *Complete Works*, edited by John M. Cooper, Hackett Pub., 1997.
- . "Gorgias." *Complete Works*, edited by John M. Cooper, Hackett Pub., 1997.
- Plutarch. "The Life of Cicero." *The Parallel lives*, Loeb Classical Library Edition, vol. VII, 1919,
http://penelope.uchicago.edu/Thayer/e/roman/texts/plutarch/lives/cicero*.html.
- Polk, Thad A., and Colleen M. Seifert, editors. *Cognitive Modeling*. The MIT Press, 2002.
- Portwood-Stacer, Laura. "Media Refusal and Conspicuous Consumption: the Performative and Political Dimensions of Facebook Abstention," *New Media & Society*, vol. 15, Nov. 2013.
- Poster, Carol. "The Task of the Bow: Heraclitus' Rhetorical Critique of Epic Language." *Philosophy & Rhetoric*, vol. 39, iss. 1, 2006, pp. 1-21.
- Postman, Neil. *Technopoly*. Alfred A. Knopf, 1991.
- Price, Derek J. de Solla. "Automata and the Origins of Mechanism and Mechanistic Philosophy." *Technology and Culture*, vol. 5, iss. 1, Winter 1964, pp. 9-23.

- Pruchnic, Jeff. *Rhetoric and Ethics in the Cybernetic Age: the Transhuman Condition*. Routledge, 2014.
- Putnam, Hilary. *Representation and Reality*. The MIT Press, 1991.
- Ramachandran, V. S. *A Brief Tour of Human Consciousness: From Imposter Poodles to Purple Numbers*. Pi, 2004.
- Reames, Robin. "The *Logos* Paradox: Heraclitus, Material Language, and Rhetoric." *Philosophy & Rhetoric*, Vol. 46, No. 3, 2013, pp. 328-350.
- Reuters. "'Master' and 'Slave' Computer Labels Unacceptable, Officials Say." *CNN*, Cable News Network, 26 Nov. 2003, www.cnn.com/2003/TECH/ptech/11/26/master.term.reut/.
- Richardson, Kathleen. *An Anthropology of Robots and AI: Annihilation Anxiety and Machines*. Routledge, Taylor & Francis Group, 2015.
- Rickert, Thomas J. *Ambient Rhetoric: The Attunements of Rhetorical Being*. University of Pittsburgh, 2013.
- Rid, Thomas. *Rise of the Machines: a Cybernetic History*. W. W. Norton & Company, 2016.
- Riskin, Jessica. *The Restless Clock: a History of the Centuries-Long Argument over What Makes Living Things Tick*. University of Chicago Press, 2018.
- Rosenthal-von der Pütten, Astrid M., Frank P. Schulte, Sabrina C. Eimler, Sabrina Sobieraj, Laura Hoffmann, Stefan Maderwald, Matthias Brand, and Nicole C. Krämer. "Investigations on empathy towards humans and robots using fMRI." *Computers in Human Behavior*, vol. 33, 2014, pp. 201-212.

- Roser, Max, and Esteban Ortiz-Ospina. "World Population Growth." *Our World in Data*, Apr. 2017, ourworldindata.org/world-population-growth.
- Sainato, Michael. "Stephen Hawking, Elon Musk, and Bill Gates Warn About Artificial Intelligence." *Observer*, 19 Aug 2015.
- Schalkoff, Robert J. *Intelligent Systems: Principles, Paradigms, and Pragmatics*. Jones and Bartlett, 2011.
- Schiappa, Edward. "Dissoi Logoi." *Classical Rhetorics and Rhetoricians: Critical Studies and Sources*, Greenwood Publishing Group, Inc, 2005.
- Scott, Ridley, director. *Alien*. Twentieth Century Fox, 1979.
- . *Blade Runner*. Produced by Ridley Scott and Hampton Francher, Screenplay by Hampton Francher and David Webb Peoples, Warner Bros., 1982.
- Searle, John R. "How to Derive an 'Ought' from an 'Is.'" *The Philosophical Review*, Duke University Press, vol. 73, iss. 1, Jan. 1964, pp. 43-58.
- . *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, 1983.
- . "Is the Brain a Digital Computer?" *Proceedings and Addresses of the American Philosophical Association*, vol. 64, iss. 3, 1990, pp. 21-37.
- . "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, vol. 3, iss. 3, 1980, pp. 417-457.
- . "Watson Doesn't Know It Won on 'Jeopardy!'" *The Wall Street Journal*, 11 Feb 2011.

- Shelley, Mary Wollstonecraft, and Malvina G. Vogel. *Frankenstein*. Baronet Books, 2008.
- Shelley, Percy Bysshe. *Prometheus Unbound*. Theclassics Us, 2013.
- Shirow, Masamune, et al. *The Ghost in the Shell Deluxe Edition (1, 1.5, 2)*. Kodansha Comics, 2017.
- Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs*, vol. 1, iss. 3, Spring 1972, pp. 229-243.
- Skirry, Justin. "Internet Encyclopedia of Philosophy." *Internet Encyclopedia of Philosophy*.
- Smith, Zach. "A Robot From an Alternate History Visits Our World." *Newsarama.com*, 19 Oct. 2009.
- Spackman, Jonathan S. and Yanchar, Stephen C. "Embodied Cognition, Representationalism, and Mechanism: A Review and Analysis." *Journal for the Theory of Social Behaviour*, vol. 44, iss. 1, 2013, pp. 46-79.
- Spielberg, Steven, and Steven Spielberg. *A.I.: Artificial Intelligence*. 2001.
- Stallman, Richard. "Mode inheritance, cloning, hooks & OOP." 16 Jan. 1995.
<https://groups.google.com/forum/?hl=en&fromgroups=#!topic/comp.emacs.xemacs/0K8leig3ZAw>
- Statt, Nick. "iPhone Manufacturer Foxconn Plans to Replace Almost Every Human Worker with Robots." *The Verge*, The Verge, 30 Dec. 2016,
www.theverge.com/2016/12/30/14128870/foxconn-robots-automation-apple-iphone-china-manufacturing.

Sterne, Jonathan. "Bourdieu, Technique and Technology." *Cultural Studies*, vol. 17, iss. 3, 2003, pp. 67-89.

Sun, Ron. "Symbol grounding: a new look at an old idea." *Philosophical Psychology*, vol. 13, iss. 2, 2000, pp. 149-172.

Supreme Court of the United States. *Citizens United v. Federal Elections Commission*. 21 Jan. 2010.

---. *Jacobellis v. Ohio*. 1964.

Sussman, Mark. "Performing the Intelligent Machine: Deception and Enchantment in the Life of the Automaton Chess Player." *Drama Review*, vol. 43, iss. 3, 1999, pp. 81-96. Web.

Swiatczak, Bartlomiej. "Conscious Representations: An Intractable Problem for the Computational Theory of Mind." *Minds & Machines*, vol. 21, 2011, pp. 19-32.

Takahashi, Dean. "Forty Years of Moore's Law." *The Seattle Times*. The Seattle Times, 18 Apr. 2005. Web. 08 Aug. 2014.

Thagard, Paul. *Mind: An Introduction to Cognitive Science*. The MIT Press, 2005.

Thompson, Nicholas. "Inside the Apocalyptic Soviet Doomsday Machine." *Wired*, Conde Nast, 21 Sept. 2009, www.wired.com/2009/09/mf-deadhand

Tolkien, J.R.R. *The Lord of the Rings*. Ballantine Books, 1982.

Tonneau, François. "Metaphor and Truth: A Review of *Representation Reconsidered* by W.M. Ramsey." *Behavior and Philosophy*, vols. 39/40, 2011, pp. 331-343.

Tonkens, Ryan. "A Challenge for Machine Ethics." *Minds & Machines*, vol. 19, 2009, pp. 421-438.

- Tononi, Giulio. *Phi: A Voyage from the Brain to the Soul*. Pantheon, 2012.
- Tousignant, Lauren. "Mystery Surrounds Creepy Radio Signal Emanating from the Middle of Nowhere." *New York Post*, New York Post, 3 Aug. 2017, nypost.com/2017/08/03/mystery-surrounds-creepy-radio-signal-emanating-from-the-middle-of-nowhere/.
- Trouvain, Jurgen & Brackhane, Fabian. "Wolfgang von Kempelen's 'Speaking Machine' as an Instrument for Demonstration and Research." *ICPhs XVII: Special Session*, 2011, pp. 17-21.
- Turing, Alan. "Computing Machinery and Intelligence." *Mind*, vol. 49, pp. 433-460.
- Turner, Fred. *How Digital Technology Found Utopian Ideology: Lessons from the First Hackers' Conference*, Stanford University.
- Varela, Francisco J., et al. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, 1991.
- Vinge, Virnor. "Virnor Vinge on the Singularity." *Virnor Vinge on the Singularity*, 1993.
- Virilio, Paul. *Crepuscular Dawn*. Translated by Sylvere Lotringer, Semiotext(e), 2002.
- Virilio, Paul. *Speed and Politics*. Translated by Marc Polizzotti, Semiotext(e), 2006.
- Wachowski, Lana, director. "The Matrix." Warner Home Video, 1999.
- Wallach, Wendell, and Colin Allen. *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press, 2009.
- Wallis, Charles. "Representation and the Imperfect Ideal." *Philosophy of Science*, vol. 61, 1994, pp. 407-428.

- Wardy, Robert. *The Birth of Rhetoric: Gorgias, Plato, and Their Successors*. Routledge, 1996.
- Wark, McKenzie. *A Hacker Manifesto*. Harvard Press, 2004.
- “Way of the Future Church: What Is This All about?” *Way of the Future*,
www.wayofthefuture.church/.
- Wenz, John. "The Questions We're Not Asking About Sex Robots." *Popular Mechanics*,
16 Sept. 2015.
- “Westworld.” Produced by Jerry Weintraub, Lisa Joy, Jonathan Nolan, J.J. Abrams.
HBO, 2016.
- Wiener, Norbert. *Cybernetics: Or, Control and Communication in the Animal and the Machine*. The MIT Press, 1965.
- Wiener, Norbert. *God and Golem, Inc.: a Comment on Certain Points Where Cybernetics Impinges on Religion*. The MIT Press, 1990.
- Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society*. Da Capo Press, 1988.
- Williams, Bernard Arthur Owen and Josefine Nauckhoff. *Nietzsche: The Gay Science*.
Cambridge University Press, 2001.
- Winner, Langdon. *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. The MIT Press, 1992.
- Wise, Robert, director. *The Day the Earth Stood Still*. 20th Century Fox, 1951.

Woodward, James. "Emotion versus Cognition in Moral Decision-Making: A Dubious Dichotomy." *Moral Brains: The Neuroscience of Morality*, edited by S. Matthew Liao, Oxford University Press, 2016.

Yancy, George, and John D Caputo. "Looking 'White' In the Face." *The New York Times*, The New York Times, 2 July 2015, opinionator.blogs.nytimes.com/2015/07/02/looking-white-in-the-face/.

Zarkadakis Georgios. *In Our Own Image: Savior or Destroyer?: The History and Future of Artificial Intelligence*. Pegasus Books LLC, 2016.