The Creative Partnership of Humans and Technology

PLATO AND THE NERD

EDWARD ASHFORD LEE

Plato and the Nerd

The Creative Partnership of Humans and Technology

Edward Ashford Lee

DRAFT FOR REVIEW
Version 0.08
To be published by MIT Press, Fall 2017.

This book was set in Times Roman using LaTeX.
Printed and bound in the United States of America.

This book is dedicated to my muse, Rhonda Righter, with thanks for many dinnertime conversations that shaped my thinking.

Contents

Preface

What This Book Is About

When I was young, my father wanted me to become a lawyer or get an MBA and take over the family business. Engineers were the people who worked for him. The brightest young minds, at least those of white Anglo-Saxon stock in the United States, went to law school, business school, or medical school. Today, engineering schools are much harder to get into, but that was not true when I was going to college. Yes, my father was profoundly disappointed in me when I double majored in Computer Science and what Yale called "Engineering and Applied Science." I made it worse when I went to MIT for graduate school in engineering and then went to work as an engineer at Bell Labs, and worse still when I went to Berkeley for a PhD and then became a professor. This book is perhaps my last-ditch attempt to justify those decisions.

When I started writing the book, I really didn't know who my target audience would be. As it has turned out, this book is targeted toward readers who are either literate technologists or numerate humanists. I'm not sure how many such people there are, but I'm convinced there must be a few. I hope you are one of them.

This book is my attempt to explain why the process of creating technology, a process that we call engineering, is a deeply creative process, and how this explains why it has become so hot and competitive, making geeks out of the brightest young minds. The book is about the culture of technology, about both its power *and* its limitations, and about how the real power of technology stems from its partnership with humans. I like to think of the book as a popular philosophy of technology, but I doubt it will be very popular, and I am not sure I have the qualifications to write about philosophy. So really, the only guarantee I can make is that it is about technology and the engineers who create technology. And even then, it is limited to the part of technology that I understand best, specifically, the digital and information technology revolutions.

This book is *not* about the artistry and creativity that is unleashed by using technology as a medium. For that topic, I recommend the wonderful book by Virginia Heffernan, *Magic and Loss* (Heffernan, 2016). Heffernan claims that "the Internet is a massive and collaborative work of realist art," but she is referring to the *content* of the Internet. In my book, I claim that Internet *technology itself*, and all of digital technology that shores it up, is a massive and collaborative creative work, even if not an artistic work.

Digital technology as a medium for this latter sort of creativity has enormous potential, well beyond what has been accomplished to date. In the first part of this book, I explain exactly why this technology has been so transformative and liberating. I study how engineers use models and abstractions to build inventive artificial worlds and give us incredible capabilities, such as the ability to carry around in our pockets everything humans have ever published.

But this is not to say that digital technology has no limitations. Pursuing a yin and yang balance, in the second part of the book, I attempt to counter a runaway enthusiasm among some thought leaders about digital technology and computation. Driven by the immense potential of computers, this enthusiasm has led to unjustified beliefs that go as far as to assert that everything in the physical world is in fact a computation, in exactly the same sense as in modern computers. Everything, including such complex phenomena as human cognition and such unfamiliar objects as quasars, is software operating on digital data. I will argue that the evidence for such conclusions is weak and the likelihood is remote that nature has limited itself to only processes that conform with today's notion of digital computation. And I will show that this digital hypothesis cannot be tested empirically, and therefore can never be construed as a scientific theory. Because the likelihood is remote, the evidence is weak, and the hypothesis is untestable, these conclusions are an act of faith. My argument here will likely get me into trouble because I'm swimming against a considerable current.

Also bucking much current thought, I argue that the goal of artificial intelligence to reproduce human cognitive functions in computers is misguided, is unlikely to succeed, and vastly underestimates the potential of computers. Instead, technology is coevolving with humans, augmenting our own cognitive and physical capabilities, all the while enabling us to nurture, evolve, and propagate the technology. We are seeing the emergence of symbiotic coevolution, where the complementarity between humans and machines dominates over their competition.

But most of the book is very much swimming with the current, upbeat about the enormous potential of technology to improve our lives. But more than just utilitarian, one of my main messages is that engineering is a deeply creative and intellectual discipline, every bit as interesting and rewarding as the arts and sciences. In areas where the technology is less mature, the creative contributions reflect the personalities, aesthetics, and idiosyncrasies of the creators. In areas that are more mature, the work can become deeply technical and opaque to outsiders. This happens in all disciplines, so this is hardly surprising.

Like the sciences, engineering is built around accepted paradigms that provide frameworks for thought. Also like the sciences, engineering is punctuated by paradigm shifts, to use the words of Thomas Kuhn (Kuhn, 1962). Unlike the sciences, however, the paradigm shifts are frequent, even relentless. I argue, in fact, that the pace of technological progress in our current culture is more limited by our human inability to assimilate new paradigms than by any physical limitations of the technology. I attempt in this book to explain why this is.

Like the arts, the evolution of the field of engineering is governed by culture, language, and cross-germination of ideas. Also like the arts, success or failure is often determined by intangible and inexplicable forces, such as fashion and culture. And in an observation that may take many readers by surprise, also like the arts, the creative media used to engineer new artifacts and systems today, particularly digital media, have become astonishingly versatile and expressive. In my opinion, this latter property, the versatility and expressiveness of digital media, accounts for the attractiveness of the field to bright young minds, more even than the lucrative job prospects.

Engineering is a broad field, encompassing everything from water supply systems to social networking software. Any individual, myself included, cannot have more than a superficial understanding of more than a few of its subdisciplines. My arguments in this book, therefore, are based on my experience with electronics, electrical engineering, and computer science. These arguments apply to digital and information technologies and may or may not apply to other technologies such as bridges and chemical plants. Nevertheless, I do know from experience that digital technologies have invaded nearly all other engineering disciplines. Modern chemical plants, for example, include substantial computer control and therefore become instances of cyberphysical systems, discussed in chapter 6. Such systems are most certainly subject to the potential, vagaries, and limitations of digital technology that I point out in this book.

I do not assume of the reader any particular technical background. In some sections of the book, I do dive more deeply than I probably should into technical topics that are near and dear to my heart, but I promise the reader that every such indulgence is short, and hopefully skipping the technical details will not seriously undermine the message. Please persist. The nerd storm will pass quickly.

I do assume a numerate reader. Against all advice, I have even included 12 equations in the book. They are not complicated equations. High school math and science is more than sufficient to fully understand them, but even then full understanding is not needed to get the message. My publisher has used this argument against me, saying that if it is true, I should remove them. But I like them. I have confidence that there are more numerate readers than there used to be. I have assured the publisher that, counting my friends and family, a few dozen book sales are assured.

The title of this book comes from the wonderful book by Nassim Nicholas Taleb, *The Black Swan* (Taleb, 2010), who titled a section of the prologue "Plato and the Nerd." Taleb talks about "Platonicity" as "the desire to cut reality into crisp shapes." Taleb laments the ensuing specialization and points out that such specialization blinds us to extraordinary events, which he calls "black swans." Following Taleb, a theme of my book is that technical disciplines are also vulnerable to excessive specialization; each speciality unwittingly adopts paradigms that turn the speciality into a slow-moving culture that resists rather than promotes innovation.

But more fundamentally, the title puts into opposition the notion that knowledge, and hence technology, consists of Platonic Ideals that exist independent of humans and is discovered by humans, and an opposing notion that humans create rather than discover knowledge and technology. The nerd in the title is a creative force, subjective and even quirky, and not an objective miner of preexisting truths.

I hope that through this book, I can change the public discourse so young people are more inclined to consider a career in engineering, and not just because of the job prospects. I am convinced that engineering is fundamentally a creative discipline, and the technical drudgery that prejudices many people is no more drudgery than found in any other creative discipline. Yes, hard work is required, but as a reward for that hard work, you can change the world.

Overview of the Chapters

Some readers like to be told what they will be told before they are told it. Putting aside the problematic self-referentiality, for those readers, I provide here a brief overview of the book. But honestly, I recommend skipping this and going directly to chapter 1. The story told in this book cannot be accurately summarized in a few paragraphs, and any such summary will necessarily make the book seem more dense than it is. Nevertheless, for those who really need it, here is my summary.

Popular perception of technology and engineering is often one of a dispassionate field dominated by logic and trading in colorless facts and truths. In chapter 1, I explore the idea of facts and truths in technology, showing that these are not just discovered but more often invented or designed. Rather than being built on timeless Platonic Ideals, technology is built on ideas that are more fluid and sometimes quirky. The notion of truth becomes more subjective; collective wisdom becomes better than individual wisdom; a narrative about how facts evolve becomes more interesting than the facts themselves; facts and truths may be wrong; and it can cost billions to show that facts are true. I then develop the idea that engineering and science, disciplines rooted in facts and truths, are complementary and overlapping, leveraging each others' methodologies. In this chapter, I try to understand the cultural phenomenon that engineering has been considered the "kid sister" of science.

In chapter 2, I focus on the relationship between discovery and invention. A key theme of this chapter is that models are invented not discovered, and it is the usefulness of models, not their truth, that gives them value. Note that the usefulness of a model need not be a practical, utilitarian sort of usefulness. A model may be useful simply because it explains or predicts observations, even if the phenomena observed have no practical application.

Models are useful to scientists when they are faithful to the natural system being studied, whereas models are useful to engineers when a physical realization can be constructed that is faithful to the model. These uses are complementary, and, in fact, are often applied in combination.

Chapter 2 is heavily influenced by Kuhn (1962). But Kuhn focused on science, not engineering. The engineering use of models results in more room for creativity in the construction of models because it is not necessary for the models to be faithful to some preexisting natural system. But the use of models can also slow technological change because models are built on paradigms that frame our thinking and therefore limit our thinking. Models can also get quite sophisticated, forcing increased specialization, which can also slow change by impeding communication across specializations.

In chapter 3, I dive into exactly how the engineering use of models enables creativity. I do this by illustrating the role that models have played in the development of digital technology, where models are stacked many layers deep, with the design of each layer affecting the designs both above and below it. Digital technology has, through this multiplicity of layers, mostly removed any meaningful physical constraints from a broad class of engineered systems. Each layer of models conforms with an established paradigm, a way of modeling and abstracting an engineered design. Innovation, therefore, is less limited by the physics of the technology than by our imagination and ability to assimilate new paradigms.

I argue that paradigms play a central role in digital technology because without them, no human could possibly comprehend the complexity of the systems we routinely build today. But these paradigms are human constructions, governed by culture and language. In many cases, the paradigms that have emerged are idiosyncratic, reflecting the personality and aesthetics of their creators.

A notable feature of digital technology is that paradigms are layered one on top of another. Semiconductor physics gives us the ability to make transistors, which we can use as electrically controlled switches that have two distinct states: "on" and "off." This enables a digital abstraction that turns out to be just the first of many layers, building up eventually to the programming languages that enable us to build databases, machine learning systems, web servers, and so on. Each of these layers forms through coalescing of competing paradigms.

In chapter 4, I explore the layered paradigms that make up much of today's digital technology hardware. I show that the physical substance of the hardware is not durable, but the paradigms are. The hardware is routinely discarded every few years as it wears out and becomes obsolete, but the principles on which the hardware is designed, with all their warts and idiosyncrasies, persist for decades.

In chapter 5, I explore the layered paradigms that make up much of today's *information* technology. These paradigms define how we construct software, and software, it turns out, endures much better than hardware. Paradigms, like human culture, change slowly, particularly compared with the speed with which technology changes. Although Kuhn's scientific paradigms are strictly human constructions, the paradigms of software are encoded in the software. In an orgy of self-referentiality, software builds its own scaffolding. The self-scaffolding of software makes it much more durable than hardware, despite its ephemeral nonsubstantive existence. It could even outlast humans.

Chapter 6 explores the structure of technology revolutions, with a particular focus on digital technology. This chapter is also heavily influenced by Kuhn, but it strives to identify how technology revolutions differ from scientific revolutions. One key difference is that technology paradigms appear and disappear much more rapidly probably because, compared with scientific paradigms, they are relatively unconstrained by the physical world and are layered one upon another many layers deep. Like scientific paradigms, new technology paradigms do not necessarily replace old ones. They may instead overlay the old ones, building new platforms on top of existing platforms. The ability to do this depends on the transitivity of models explored in the three previous chapters. Unlike scientific paradigms, the crises that trigger new technology paradigms do not arise so much from the discovery of anomalies but from increasing complexity and technology-driven opportunity.

To balance the enthusiasm, the next few chapters look at what we *cannot* do with digital technology, at least not today. This requires explaining three classic concepts that emerged in the 20th century: Shannon's information theory, the Church-Turing thesis, and Gödel's incompleteness of formal models. In the later chapters, I consider the concept of determinism and examine how we can build models that embrace uncertainty using the notion of probability. Along the way, I need to confront another paradigm that emerged in 20th century called digital physics and a view that human cognition is software.

This part of the story begins in chapter 7, where I examine the concept of information — what it is and how to measure it. In this chapter, I introduce Claude Shannon's way of measuring information and show that his notion of information often cannot be represented digitally. I define an "information-processing machine" more broadly than what can be realized using software and computers, as they exist today.

In chapter 8, I explain what software cannot do. I point out that the number of information-processing functions is vastly larger than the number of possible computer programs. I introduce Alan Turing's undecidability result, which shows that useful information-processing functions exist that are not realizable by software on today's computers. But it does not follow that if a function is not realizable by software, then it is not realizable by any machine.

I caution against getting carried away by enthusiasm, marveling at what has already been accomplished with software, and caution against predicting that natural phenomena such as cognition and understanding are realizable in software. Here, I am forced to confront a belief that some people call "digital physics": that the physical world is somehow software or equivalent to software. I argue that this idea is unlikely to be either true or useful as a way of understanding the physical world, at least in its more extreme forms, and I show that this thesis is not falsifiable and therefore not scientific.

In chapter 9, I go beyond the countable world of computing and argue that computers are not universal machines and their real power comes from their partnership with humans. I explain the notion of a continuum, a concept that is out of reach for software and rejected by digital physics but seemingly essential for modeling the physical world. I examine the fundamental limitations of formal models that underlie the world of software, and I argue that the partnership and coevolution of humans and computers is much more powerful than either alone. In this chapter, I explain Kurt Gödel's famous incompleteness theorems, which impose fundamental limits on any modeling formalism that is capable of self-reference. We need to be humble, but we also need to recognize the as yet vast unexplored potential that still waits for us to catch up.

In chapter 10, I consider determinism, a property of software and many mathematical models of nature. I argue that determinism is a property of models not of the physical world. But it is an extremely valuable property, one that has historically delivered considerable payoffs in engineering and science. However, determinism also has its limits. Even deterministic models may not be usefully predictive because of chaos and complexity. Also, families of deterministic models that embrace both discrete and continuous behaviors are incomplete. There are unavoidable holes where determinism breaks down, and deterministic models have their limitations. In many cases, nondeterministic models are simpler and better reflect what we do not know. Nondeterministic models, used explicitly and judiciously, play an essential role in engineering.

In chapter 11, I finally confront the meaning of randomness and its measure, probability, which quantifies the likelihood of nondeterministic events. I argue that probability is fundamentally a model of *uncertainty* about something and not directly a model of that something. It models what we do not know. I examine the long-standing debate between the frequentists and the Bayesians, coming down solidly on the side of the Bayesians. I show that the philosophical difficulties presented by randomness vanish when using models in the engineering sense rather the scientific sense and when interpreting probability in the Bayesian sense. In this chapter, I also reconsider continuums and argue that probabilistic models over continuums reinforce the conclusion that digital physics is extremely unlikely. As a consequence, we should demand incontrovertible evidence for digital physics before accepting it.

In the final chapter, I tie things together by examining the epistemic role that models have in technology and the relationship between models and the physical systems they ultimately model. I leverage the previous arguments in the book: At least with digital technology, so many layers of abstraction exist between the models and the physical reality that the connection between the two becomes tenuous indeed. Moreover, the self-scaffolding that software paradigms have, described in chapter 5, allows these models to stand on their own, almost but not completely independent of physical reality. I argue that this does not lead to a Cartesian mind-body dualism, but it does emphasize the need to insist, with great determination and discipline, on separating the map from the territory. Models are best viewed as having a separate reality from the physical world, despite existing in the physical world.

The most expressive modeling paradigms are capable of self-reference, which enables them to build their own scaffolding but also makes them necessarily incomplete. This incompleteness is fundamentally what enables creativity and ensures that what we can accomplish with technology is limitless. So what holds us back? In this final chapter, I consider both the obstacles to progress and the threats that technology, when misapplied, can have on society.

Acknowledgments

The author gratefully acknowledges contributions and helpful suggestions from Christopher Brooks, Malik Ghallab, Thomas Henzinger, Madeline Johnson, Hokeun Kim, Gil Lederman, Marten Lohstroh, Dave Messerschmitt, Mehrdad Niknami, Rodion Rathbone, Rhonda Righter, Bernhard Rumpe, Naresh Shanbhag, Joseph Sifakis, Marjan Sirjani, Kimball Strong, David J. Stump, and Eli Yablonovitch. I would also like to thank three anonymous reviewers commissioned by the publisher who were extremely helpful. Several of these people disagreed with major points that I make in the book, and they thereby helped me to understand where my arguments needed to be strengthened or reworked. All remaining errors and opinions that I have stubbornly stuck to are entirely my own, not those of these contributors.

Most especially, however, I would like to thank two very special people who played a major role in the development of this book. The first is Heather Levien, who, unlike me, really knows how to write and without whom this book would be a disorganized pile of random ideas. The second is my mom, Kitty Fassett, a professional musician with an aversion for mathematics but a true intellectual and also a great writer. Without her help, this book would be unreadable to nonspecialists. She was my guinea pig, telling me each place where a nonspecialist might get lost.

I also thank the staff at MIT Press and Heather Jefferson for her superb copy editing. In addition, I thank the many unwitting contributors who have offered their thoughts through largely anonymous media such as Wikipedia and the contributors who have generously posted images online that I can (and have) reused because of their choice of creative commons licenses.