

**Edward Ashford Lee.** *The Coevolution: The Entwined Futures of Humans and Machines.* MIT Press 2020. 384 pp. \$34.95 USD (Hardcover ISBN 9780262043939).

A book touching on topics of artificial intelligence, computer science, and evolution may initially sound like it is meant for niche or expert audiences; however, *Coevolution* defies such constraints through careful attention to defining concepts and clear progression through complex themes. The typical axiom is that culture shapes the kind of technology developed. In *Coevolution*, Edward Ashford Lee asserts that technology is developing alongside humans. The most likely outcome of this coevolution is a symbiosis, rather than annihilation or integration, of machine and man.

Early in the preface, Lee explains to readers *why* the issue of technological development is relevant for everyone. Digital technology is disrupting the pillars of society, affecting the foundations that shape the world we live in, including economies, social relationships, and political structures (xi). Because *all* humans in contemporary societies inherently and necessarily interact with each of these pillars on a daily basis, no one can claim non-relevancy of the rise of digital technology for them. More than any other invention, digital technology is changing the ways we *act* as humans: the way we work, think, and interact with one another.

In response to concerns or fears that people may have about the book's fundamental tenet (that humans can no longer control the machines that are now so intertwined in their lives), Lee offers a parsimonious comparison to a relationship more familiar to readers. In considering humanity's connection to other living flora and fauna, he offers a less-disconcerting reality: humans affect *but do not control* the biological living things that surround us. In light of this, the relationship between humans and machines has a precedent, and may be less frightening.

Some readers may be off put by the way Lee initially characterizes machines. In an effort to liken machines to beings more familiar to humans (living beings), he uses metaphors to describe how they react to stimuli, how they are comprised of organisms, and how they require 'nourishment.' More skeptical readers may liken this to anthropomorphizing machines in a way that artificial intelligence skeptics and scholars warn against. Regardless, Lee says that the question of the most importance is *not* whether digital technology can be considered to be alive. Such a question would lead to a cascade of philosophical and technical questions that are equally challenging: Are digital technologies capable of living without humans? Will machines exceed human intelligence? Are machines capable of free will? Are they capable of ethics? Rather, Lee suggests that the most important question lies elsewhere: can the metaphor of technology as living beings help us better understand how they interact with humans and society?

*Coevolution's* fourteen chapters are structured to answer this question. Lee himself notes that attempts to summarize each chapter in a few paragraphs run the risk of making the book seem denser than it actually is. Yet the length of this review requires even briefer summarizations. Great pains have been taken to keep from portraying the *book's* content in a complex or difficult manner; however, if certain chapters or themes come across as not particularly friendly for general audiences, the reviewer defers to Lee's warning that perhaps a chapter or section has been summarized too succinctly.

Chapter 1 expands on a fundamental framing that Lee introduces in the preface: the metaphor of living digital beings that change, and are changed by, humans. This sets up a premise that carries the remainder of the book forward: by this point in time, humans and machines are now inextricably necessary to one another. Chapter 2 offers a challenge to the notion of machines as living beings by investigating the usefulness of the metaphor. The chapter depicts the fit of a living being's metaphor,

particularly as it speaks to our role in the human-machine relationship. The next chapter addresses a familiar debate in contemporary society: is technology making humans smarter or dumber? Lee takes a middle path approach, arguing that technology may limit our thinking *individually*, while ultimately expanding the thinking of humanity as a collective. In a similar process to globalization, society as a whole may benefit and be enriched, while individuals within the collective may experience negative effects or be ‘passed over’ in receiving benefits.

Chapter 4 takes the evolutionary concept of *feedback* and applies it to technology, focusing on AI software and deep-learning algorithms. Giving the example of an on-the-fly speech, Lee discusses the interactions between inspiration, speech, and conscientiousness. Words selected to represent thoughts spark new interpretive thoughts, such as a process called *noisy feedback* that is at the root of creativity. While machines are certainly capable of feedback, whether they are capable of *noisy* feedback is unknown. In particular, the *appearance* of creativity or human-like intelligence does not necessarily mean the *existence* of creativity or human-like intelligence.

Expanding on the previous chapter, chapter 5 considers the concept of *negative feedback*, or the ability to revisit and learn from past mistakes. Lee moves closer to the questions of artificial intelligence and surpassing humans in this chapter, considering the ability to learn from errors and correct them moving forward to be a necessary index for real intelligence. Chapter 6, which is shorter than most other chapters in the book, considers the ethical dilemmas inherently involved in ‘classifying.’ Humans are notoriously fallible in constructing explanations and defending reasoning. Therefore, seeking to have machines to make decisions the way humans do is likely to be a faulty aim.

Chapter 7 offers an alternative perspective to the fit of the ‘machines as living beings’ metaphor. Taking a stance of embodied cognition (the notion that ‘thinking’ requires more than a computing brain, and is necessarily tied to organic bodies), the chapter ensures balanced consideration is given to the fundamental metaphor upon which the book rests. The discussion is warranted, as the following Chapter 8 examines whether organic and cognitive beings can really be replicated by machines. In this chapter, Lee raises an important objection to a common premise in machine learning: that human thinking is itself digital and algorithmic. Calling this “faith, not fact,” Lee reckons with the uncomfortable uncertainty that results from a non-provable (and conversely, non-disprovable) premise.

Chapter 9 nuances a question that is at the core of many fears about machine learning and artificial intelligence: can machines be as smart as, or smarter than, humans? Lee posits that this is not the right question to ask. Instead, he argues that machines already demonstrate forms of intelligence different from that expressed by humans. Furthermore, he maintains that machines have, in many ways, exceeded humanity’s cognitive capabilities. The chapter thus dives deeper into what it meant by ‘intelligence,’ and challenges the notion that intelligence is a solely human (or organic) trait.

If machines can exhibit intelligence, the following concern is whether they can be held accountable for the actions that result from the realization of this intelligence. Chapter 10 explores the transitive principle in the context of machines. When a machine creates art, or conducts a crime, is its creator responsible for the outcome? Of all of *Coevolution’s* chapters, this one dives furthest into complex, abstract, and perennial questions about free will, ethics, and imagination. Chapter 11 continues the philosophical bent of the previous chapter to question the idea of causation. The debate has significant implications for whether machines can (or even should) be held accountable for the actions they undertake. Lee suggests that an understanding of causality is necessary to developing a first-person perspective, and is inherently subjective as a cognitive activity.

Lee writes that chapter 12 is, perhaps, the most challenging chapter in the text, as it expands on the foundation set by chapters 10 and 11 to explain highly technical concepts, including the power of interaction (particularly in comparison to observation), imperfect information presentation, zero-knowledge proofs, and biostimulation. While most (or all) of these terms may be foreign for lay readers, Lee explains each clearly and simply.

Chapter 13 shifts direction to consider practical recommendations for living in an era of machine and human coevolution. While the chapter does not paint a rosy outlook for humans as technology evolves, Lee cautions against a pessimistic outlook. Chapter 14 wraps up the book's complex subject matter by bringing the conversation back to evolution and biological processes. Not only are humans and machines symbiotes, Lee suggests, but they may be approaching a point of *obligate* symbiosis, where neither can live without the other. Rapid coevolution is inherently unpredictable. The lines are blurred as to whether humans are defining technology, or it is defining humans. Lee clarifies that a lack of control or agency is not a bad thing: evolution is an agency-less process, and it serves a purpose and has led to the world we live in today.

Lee's long career in electrical engineering and computer science shines through in *The Coevolution*, and his expertise is apparent, even for general and public audiences who may lack even introductory engineering, computer science, or information science backgrounds. He does not rely on a heavy use of jargon or scientific terminology. When field terms are used, their explanations and definitions appear to be determined through careful selection. In this way, *Coevolution* is even more accessible (though *never* by sacrificing depth of thought) than his previous general-audience publication, *Plato and the Nerd: The Creative Partnership of Humans and Technology* (MIT Press 2017). The relevancy of the book's topic in contemporary society means that *The Coevolution* will likely be of particular interest to curious undergraduates, trade readers, and engineers from various backgrounds.

**Ellen A. Ahlness**, University of Washington