R_TYPE/ISTOCK/THINKSTOCK

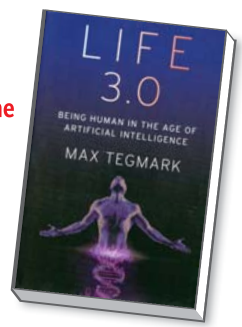
The dangerous allure of advanced AI

Life 3.0

Being Human in the Age of Artificial Intelligence

Max Tegmark

Alfred A. Knopf, 2017. $28.00



Albert Einstein famously quipped, "I know not with what weapons World War III will be fought, but World War IV will be fought with sticks and stones." In *Life 3.0: Being Human in the Age of Artificial Intelligence*, MIT professor Max Tegmark makes a persuasive case that artificial intelligence (AI) may be the weapon—or the enemy—that reduces us to those sticks and stones, unless we humans and our AI partners act with extreme care.

Tegmark is also the author of *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality* (2014). He clearly has found a formula for a successful book: a provocative subject, a well-reasoned argument, an easy-to-understand classification scheme (for example, a four-level multiverse in *My Mathematical Universe*; three forms of life in *Life 3.0*), whimsical illustrations, and personal anecdotes that humanize science. Tegmark also throws in something that's become close to a personal trademark—a vivid fictional opening scene that reads as if it had been ripped from a screenplay.

Tegmark defines life rather narrowly, as activities that can process information and can retain said information's complexity after replication. By that definition, it may not be obvious how a bacterium, which most people would say is alive, actually qualifies as life. In contrast, an electronic thermostat clearly processes complex information and so should qualify under his definition. Tegmark builds on the idea of life as information processing to construct a three-form schema. Life 1.0 is the most basic biological stage; Life 2.0 cannot design its own hardware but can alter its software—for example, by learning a new language. Tegmark then leads us to a life beyond human: superhuman AI, or Life 3.0.

Tegmark conceives of Life 3.0 as entities that not only can reinvent themselves via software enhancements but also can upgrade their hardware. Although we humans can install artificial prostheses, we cannot (as yet) make duplicates or create improved versions of ourselves at will. As such, we only qualify, at best, as "Life 2.1."

After dispensing with definitions, Tegmark takes us through the history of computing and AI. The book *Life 3.0* is less about the design or implementation of the third generation of life, though such topics are thoroughly discussed, than it is about Life 3.0's implications. Much has been made of the perils of runaway AI. Elon Musk, for example, called it "potentially more dangerous than nukes," and Stephen Hawking said it "could spell the end of the human race."

Tegmark is more optimistic than Hawking or Musk about the future of humans living with (or under) Life 3.0. Still, his is a cautionary tale, and he warns of the potential dangers should AI develop goals that are not aligned with humanity's. Consider, for example, what would happen if terrorists were to equip autonomous drones with weapons, creating weaponized AI. *Life 3.0* urges us to prepare for an onslaught of superintelligent AI now rather than waiting and, with luck, learning from our mistakes as humans typically do. His argument is clear: We can afford neither procrastination nor mistakes this time.

Tegmark presents a thorough account of massive flaws in several previous human–AI interactions that have led to disaster in part because of code errors, such as the tragic 1979 death of a Ford Motor Company worker struck in the head by a robotic arm. But he suggests that future AI will be superresilient, immune to software bugs. And what about competition? What happens when two AIs square off against each other in a battle for scarce resources?

The book is not without a few minor irritations. At first glance, the decimal place in his classification scheme seemed superfluous. Surely Tegmark doesn't mean there is an infinitude of different levels of life? But, in fact, he does seem to assert such a continuum, with Life 2.1 being humans who are able to learn languages and skills and thus improve their software and also make upgrades to their hardware with prostheses. Life 3.0

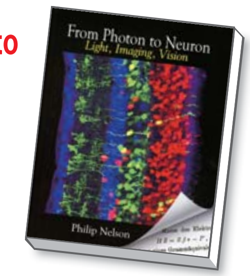
will be able to replicate itself—that is, the information it contains. We all know life can sometimes be irrational, but is there really any difference between Life 3.1 and Life 3.1415…?

There's also a bit too much hagiography of Musk; too many plugs for funding the Future of Life Institute, of which Tegmark is a founder; and too many photos and lists of people discussing AI at after-workshop dinners. Most of those could easily have been jettisoned to reduce the book's heft and to accentuate other material more essential to Tegmark's arguments.

Alas, Tegmark offers no solutions for AI life-forms that plague us today, such as AI chatbots or irritating Microsoft paper clips. But he does offer hope, if we act quickly and intelligently. Will there be a Life 4.0, or even a Life 3.1? It's hard to know. Perhaps if humanity takes Tegmark's call to action seriously, there may be more than just hope. There may actually be a chance for a future.

Brian Keating
University of California, San Diego

A pedagogical master class on biological physics

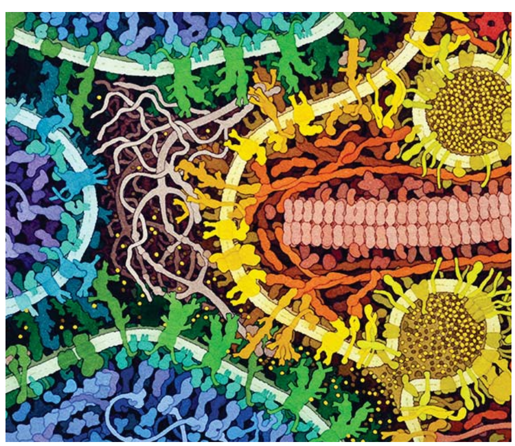


From Photon to Neuron
Light, Imaging, Vision
Philip Nelson
Princeton U. Press, 2017. $110.00

Philip Nelson's book *From Photon to Neuron: Light, Imaging, Vision* completes a trilogy begun by *Biological Physics* (2004) and *Physical Models of Living Systems* (2015). Those works establish Nelson as the preeminent author of textbooks at the intersection of physics and biology. All three books are aimed at upper-level undergraduates who already have studied a year of physics and calculus, but the texts are rich enough for the graduate level too.

From Photon to Neuron covers topics throughout biological physics. For instance, fluorescence microscopy is a theme Nelson introduces early and revisits often. He devotes one chapter to color vision and another to superresolution microscopy. My favorite chapter, "Imaging by X-Ray Diffraction," begins with Rosalind Franklin's iconic x-ray diffraction pattern of DNA and then develops enough theory to explain how James Watson and Francis Crick could, at a glance, obtain the key information they needed to derive their famous double helix structure.

Nelson presents enough electrophysiology to describe how rhodopsin's absorption of a photon causes a voltage signal across the neural membrane and enough physical optics to explain the iridescence of butterfly wings. The network diagrams of signaling cascades are a little dry, but that may reflect my own tastes rather than Nelson's presentation. Other topics include photosynthesis, fluorescence resonance energy transfer (FRET), and two-photon imaging. David Goodsell's beautiful drawings (right) further enhance the material.



A NEURAL SYNAPSE. Painting by David S. Goodsell, as published in *From Photon to Neuron: Light, Imaging, Vision* by Philip Nelson. Copyright © 2017 by Philip C. Nelson. Reprinted by permission of Princeton U. Press.

Instructors considering *From Photon to Neuron* may wonder if it is best suited for physicists interested in biology or biologists interested in physics. In my opinion, physics students will gain the most from this book, as they should be able to handle most of the mathematics. Biology majors will be challenged—Nelson includes, for example, the Fresnel integral—but the book will improve their quantitative skills. Instructors should also be aware that part 3 contains some advanced topics, such as the quantum mechanical analysis of the harmonic oscillator, that seem out of place in an undergraduate book. Students with a weak command of calculus and no desire to improve it may find Sönke Johnsen's excellent *The Optics of Life* (2011) more palatable.

The wave and particle properties of light are both crucial to biology. For instance, diffraction limits visual acuity, but a rod cell in the retina responds to a single photon. Nelson adopts a perspective like the one Richard Feynman presented in *QED: The Strange Theory of Light and Matter* (1985): Photons are governed by a probability amplitude that obeys a stationary-phase principle. That powerful point of view highlights the intimate relationship between quantum mechanics, probability, and vision. Physics students will appreciate it; I am not sure what biology students will make of it. For me, it works. Its disadvantage is that you must add a lot of $e^{i\varphi}$ terms to explain simple concepts like reflection and refraction.

Readers who are interested in vision but have little concern for light or imaging might prefer Robert Rodieck's masterpiece *The First Steps in Seeing* (1998). The books by Rodieck and Nelson share several characteristics: eloquent prose, outstanding artwork, and a quantitative approach that most biology textbooks lack. Nelson's book, however, is more useful for teaching; it includes homework problems, end-of-chapter summaries, and recommendations for additional reading.

Nelson's emphasis on computer code—or at least his insistence that the students write their own code—also sets his books apart. Many of his homework exercises