

The Sobering Limitations of Artificial Intelligence

July 25, 2022

by Laurence B. Siegel

Will artificial general intelligence (AGI) transform the experience of being human, opening up possibilities of knowledge, achievement, and prosperity that we can now barely conceive?

Or is AGI an existential threat to humanity, something to be feared and restrictively confined?

Erik J. Larson, in a fascinating book entitled *The Myth of Artificial Intelligence*, says “neither.”¹ I agree. AGI, if it is ever achieved, will be an illusion created by very fast computers, very big data, and very clever programmers. The promise or threat of AGI is hype. Lesser kinds of AI are real and need to be reckoned with. I’ll set forth a hierarchy of AI types in a moment.

Larson’s book is an exploration of aspects of philosophy, linguistics, intellectual history, computer science, and mathematical logic that bear on the assessment of AI and AGI. I’ve been obsessed with it, which is not my reaction to most books. Do I recommend it? It is not an “investment book,” but investors would benefit much more from learning about technology and the other fields I mentioned than from yet another investment book. Yes, I recommend it, but only for the intellectually adventurous. It is not easy.

That’s my review. The rest of this article is a collection of thoughts about AI, based on what I’ve learned from Larson and others.

Types of artificial intelligence

AI means many things, so we need a classification system to make the discussion clear. Because Larson does not present a typology of AI, I’ll use one I heard from an AI entrepreneur, based on work by DARPA, in order of least to most complex:²

- Algorithmic AI
- Statistical AI
- Causal AI
- Artificial general intelligence (AGI)

Algorithmic AI

Algorithmic AI uses a cookbook approach to solving problems. It is what a “smart” traffic signal does when it “sees” a car stop, despite the absence of cross traffic, and turns from red to green. The first time I observed this it seemed eerie, as though the traffic signal had a mind of its own; now we’re used to it.

This is just automation, the 19th century (or older) concept of providing machines with feedback so they can operate more efficiently. Automation has developed to a level that looks to the untrained eye like a modest degree of intelligence. Algorithmic AI is the use of computers, with their “if-then-else” or Boolean logic circuits, to implement automation.³

Statistical AI

The next level of complexity is statistical AI. It is what most tech-savvy people, including investment managers, are thinking of when they refer to AI. Bryan Kelly, a Yale professor and head of machine learning at AQR Capital Management, points out that the AI used in investment management mostly involves the application of basic statistical tools, developed by old-timers such as Thomas Bayes (1702-1761) and William Gosset (1876-1937), to very large datasets using very fast computers. It is not “intelligence.” Kelly refers to himself as head of machine learning, not head of artificial intelligence. The idea that machines can learn is not hype but reality; more precisely, machines can be *programmed* to learn.



Erik J. Larson

Larson echoes Kelly's thought: "What we now refer to as data science (or, increasingly, AI) is really an old field, given new wings by Moore's law and massive volumes of data, mostly made available by the growth of the web."

Statistical AI is what enables Amazon to recommend books that it "thinks" are likely to interest you, based to the books you've already bought. It is what helps airlines fill, but not overfill, their airplanes with passengers by constantly adjusting prices. It looks for patterns in stock returns. We have woven statistical AI into our lives so extensively that we barely notice it, despite it being a recent innovation.

Causal (contextual) AI

Causal AI, pioneered by the Turing Award-winning Israeli scientist Judea Pearl,⁴ is an enhancement to statistical AI that looks beyond correlations between variables for causal relationships. This task is tricky even for humans: the cautionary phrase, "correlation does not mean causation," applies in all applications of statistics.

Larson cites Pearl as "argu[ing] that machine learning can never supply real understanding because the analysis of data does not bridge to knowledge of the causal structure of the real world, essential for intelligence." Taking this as his starting point, Pearl then developed algorithms that look for causal relationships.

This is accomplished by teaching context to the computer. Because Larson is a specialist in natural language learning (by computers), some of the best material in his book is about teaching context and common sense to machines. Unsurprisingly, this is somewhere between difficult and impossible:

Common sense requires a rich understanding of the real world, which decomposes broadly into two parts: first, AI systems must somehow acquire everyday knowledge (and lots of it); and second, they must possess some inferential capability to make use of it.

Having identified the challenge, Larson describes the experience of meeting it:

The knowledge base of an ordinary person is unbelievably large, and inputting and representing it in a computer is a gargantuan task. Spoon-feeding a computer with common sense turned out to be a lifelong philosophical project, ferreting out commonsense knowledge like "pouring a liquid into a glass container with no cracks and only one opening will fill it up." Or that "living humans have heads," or that "a road is a pathway with a hard surface intended for vehicle travel."

Researchers were assuming computers would "get it" eventually, but...the project seemed unending.

It is unavoidable, then, that Larson, an expert on natural language, would be called to perform this task. He is saying that he and his colleagues failed. Much of what follows in Larson's book is a meditation on human language (not the stilted way we talk to machines but the flowing language we use to talk to each other) and the complexity of translating it to something a machine can process in a useful way.⁵

Artificial general intelligence

If all this "spoon-feeding" is needed to achieve a satisfactory level of causal AI, the next level, AGI, seems insurmountable.

And so it may remain. Biological intelligence, while subject to physical laws, has acquired characteristics over 230 million years of evolution that may not be replicable using logic circuits. Horses have horse sense; as we'll soon see, even turtles have a smidgen of it.

Computers have shown no evidence of common sense, which is one reason that science-fiction writers and philosophers, such as Nick Bostrom and Eliezer Yudkowsky, have had a field day speculating about evil AI robots taking over the universe.⁶

Larson believes these authors' wild speculations are unfounded. Using humor so dry it's hard to tell whether he's joking, Larson presents a parable involving an AI-enabled robot tasked with making paper clips; misunderstanding the instructions, it turns all the atoms in the universe, including us, into paper clips. But Larson is not worried about it:

The idea that the coming superintelligence will somehow be laser-focused and uber-competent at achieving an objective *yet have zero common sense* seems to cut against the grain of superintelligence itself – which is, after all, supposed to be human intelligence *plus more*.⁷

But what about DALL-E?

Recent events have boosted the hopes of AI researchers, who have developed a natural language-processing program called DALL-E (a nerdy pun on the artist Salvador Dali and the movie WALL-E, which is about sentient robots) that produces art based on verbal instructions from its user. Its developers write,

DALL-E 2...has learned the relationship between images and the text used to describe them. It uses a process called "diffusion," which starts with a pattern of random dots and gradually alters that pattern towards an image when it recognizes specific aspects of that image.

Here's my first piece of DALL-E art, using the version at <http://www.crayon.com>. I asked for a New York street scene drawn in the style of Caravaggio (1571-1610).

Exhibit 1

DALL-E Mini's "New York Street Scene Drawn by Caravaggio," compared with an actual Caravaggio painting



It's beyond terrible. DALL-E doesn't seem to know anything about Caravaggio even though his paintings are well known and available on the web. It is murky on the idea of a street scene (the street is sketched in poorly at the very bottom). I'll admit that it was unfair of me to (1) use the free public version of the software and (2) give it such challenging instructions, so I'll try something easier:

Exhibit 2

DALL-E Mini's Winnie-the-Pooh Flying a Spaceship



OK, I'm impressed. Given instructions that a six-year-old child could understand, it drew an image better than most adults. It even captured Winnie's typical bemused facial expression.

Isn't that AGI? No. It's AI, but it's not *general*. It's highly specific to producing visual images from natural-language instructions. The method is algorithmic and does not use anything resembling imagination. Winnie's facial expression gives the illusion of imagination because it is inferred from the many existing Winnies, drawn by people, to which DALL-E has access. If you asked DALL-E to clean your room, it would be baffled; it won't tell you that it's "physically handicapped" with respect to cleaning your room and that you should get a Roomba. DALL-E is not general intelligence in any way.

The AI sand dune

All four forms of AI - well, the three that already exist - give the illusion of being biological. The reason is that, until recently, we have not been accustomed to machines "thinking," and we are easily fooled when confronted with the unfamiliar. This illusion is produced by big data, just as the illusion that a sand dune is biological (with sensual curves) is produced by "big sand" - the immense number of grains of sand in Exhibit 3. The dune looks even more biological when the wind blows and it moves. Yet it is just a pile of rocks that, because they are so small and so many, give the momentary illusion of being alive.

Exhibit 3

A "living" sand dune



The logic behind the near impossibility of true AGI

We already have machines that can solve astrophysics problems and beat the greatest chess masters. These skills depend only the first three types of AI. "True" AGI, not just mimicking but equaling human cognition, requires that the machine be conscious and self-aware. This is a high bar: Dogs are conscious,

but they are not conscious that they are conscious. We call this last step *self-awareness*. Of all the animals, plants, and machines in the world, people are the only known example of it.

The Turing test

To qualify a machine as being capable of human-level cognition – “intelligent” – it would have to pass a Turing test. Devised in 1950 by the British computer scientist Alan Turing, the test consists of a human conversing with a counterparty (either another human or a machine). The machine passes the test if the human cannot tell which he is talking to. No machine has ever done so or come close.

It is also not fair for the tester to ask easy questions. Asked by a friend to give an example of a good Turing-test question, I suggested: “Whom do you find more inspiring, Shakespeare or Beethoven, and why?” A satisfactory answer involves:

- an understanding of what it means to be inspired;
- a sense that “you” are likely to have a different opinion than someone else, and that that’s OK – there is no “right answer”;
- the idea that literature and music are not the same, but similar enough that the question is not crazy; and
- contextual knowledge of both Shakespeare and Beethoven, the advantages they had, the obstacles they had to overcome, and the societies they lived in.

A computer that mined the world’s libraries for context would learn valuable details but still fail the test, because the question is mostly about feeling, which (so far) machines don’t have.

The evolution of true intelligence (not just the appearance of it)

But why is it so hard to imbue a machine with feeling? The answer is in biological evolution across deep time.

In a very real sense, our brains are just machines – they consist of billions of neurons that obey the laws of chemistry and physics. This fact has encouraged generations of AI researchers to try to build human-level intelligent machines. Why does this effort go nowhere?

The arrangement of the 80 billion or so neurons in the human brain is the result of 230 million years of evolution, which has solved – by random variation and natural selection – an optimization problem of incredible complexity.⁸ At each step in each individual’s struggle to survive and reproduce, certain arrangements of neurons were favored by natural selection (the organism lived long enough to reproduce) and others were disfavored (it didn’t).

Intelligence, evolution, and mistakes

This optimization “program” also makes mistakes; for example, the rods and cones in a human being’s eyes are mounted behind the retina, not in front of it where an optical engineer would put them. Our vision would be much better if some distant ancestor of ours had a mutation that flipped the retina around and enabled its owner to outcompete other individuals because of its better eyesight. Because that did not happen, we make do with pretty good eyes.⁹ But other mistakes, including those that resulted in intelligence, enhanced the probability of survival. We are built of mistakes.

In fact, it is possible that intelligence itself is a mistake. Intelligence is an extreme latecomer in evolution; animals and other organisms were doing fine for hundreds of millions of years without it. Intelligence is also very energy-intensive, as well as costly in terms of other brain and body functions forgone. It is thus possible that life is common in the universe and intelligence very rare.

Back to machines. We don’t have the luxury of making 230 million years of mistakes and random variations to produce an intelligent machine. Because we operate on a radically shorter time scale, with fewer degrees of freedom, we take a much more direct route to our objective, practically guaranteeing that the serendipitous mistakes that led to human intelligence will not occur with machines. For this reason, machine “brains” are much less complex than a human brain, therefore lacking the essential human characteristics of feeling, opinion, and curiosity.

Why did I mention 230 million years? Because turtles save each other’s lives by flipping them over when they become stuck in an upside-down position.¹⁰ Turtles emerged 230 million years ago, so I chose turtle kindness as marking the beginning of cognition.

Intelligence relies on abductive reasoning

I close by discussing Larson's argument, central to his book, that if AGI is ever to exist it will rely on *abduction*. No, not the seizure of someone against their will, but a form of logical thinking that you didn't learn about in high school.

Deduction

Classic logic relies on deduction and induction. Deduction is very simple: Socrates is a man; all men are mortal; therefore, Socrates is mortal. If the premises are true, the conclusion is true with certainty. (Surely you remembered *that* from high school).

Induction

Induction, which introduces uncertainty, is more subtle:

Two of 10 balls drawn from a bag are red, so the probability of the eleventh ball drawn from the bag being red is 20%.

The error bars around this estimate are wide: 10 balls are a small sample. There is no special reason to think that the color distribution of the first 10 balls is representative of the rest of them. That's why statistical inference, the special case of inductive reasoning most often used by investors, is so difficult to teach and learn (and should be supplemented with outside information, as Bayesian statistics recommends).

Inductive reasoning is not just about counting frequencies. It can bring in context and additional information:

Keisha, a high school senior, is six feet tall. Therefore, she is more likely to be on the volleyball team than a senior girl of average height.

There is plenty of context in this statement: the probability of a volleyball player being tall is high, and fewer than 2% of senior girls are six feet tall or more. These bits of information can inform Bayesian priors which would raise the perceived likelihood that Keisha is a volleyball player. The more outside information is brought into inductive reasoning, the more likely it is to be correct.

Abduction

While deduction and induction were known to the ancients, abduction is a recent invention or discovery due to the American philosopher Charles Sanders Peirce in the 1870s. It moves from an observation to a guess or theory that explains the observation. Peirce wrote, "Abduction is the process of forming explanatory hypotheses. It is the only logical operation which introduces any new idea."¹¹ (To abduct is to pull out, in this case a general truth from a specific one.)

According to Peirce, the fundamental syllogism of abduction is

The surprising fact, C, is observed. But if A were true, C would be a matter of course. Hence, there is reason to suspect that A is true.

This isn't induction and it certainly isn't deduction. It was, in Peirce's time, a wholly new concept in formal logic – yet it described perfectly what scientists and other thinkers had been doing since the idea of a hypothesis (obviously a Greek word) first arose in ancient times. It is a wonder Peirce is not more famous.

Larson asserts that abduction is the basis of all true intelligence. This is a bold statement. Larson backs it up (sort of) by commenting on the breadth of knowledge needed to engage in abductive reasoning: "We guess, out of a background of effectively infinite possibilities, which hypotheses seem likely or plausible."

Now, imagine a computer trying to decide what information to gather to narrow these hypotheses down from a near-infinite number to a manageable one. The first task is to understand the physical and social environment. This avoids wasting time on a million or a billion hypotheses that are ridiculous on their face if you know anything about that environment, but plausible if you don't. Anything relying on the moon being made of green cheese, the existence of a perpetual motion machine, or intervention by She-Ra (the



Charles Sanders Peirce
1839–1914

princess of power) can be ruled out.

It wasn't hard for me to think of three ridiculous hypotheses. Identifying *all* of them, however – leaving only a few plausible ones to evaluate – is beyond my or anyone's ability. It is beyond the ability of the hive mind consisting of everyone in the world. This collective common sense, then, cannot be spoon-fed into a computer because a machine-readable summary of it will necessarily leave out much more that it includes.

Considering the difficulty Larson had in teaching a computer even a smidgeon of common sense, it's unlikely that a computer will ever engage in abductive reasoning in the way that a person does.¹²

Advice for investors

Investors are always well advised to keep up with economic and financial news, and with the latest findings about asset allocation, diversification, and other tools of the trade. Those are the basics.

But many investors, hopefully most, will want to learn about the companies they invest in, now and in the future. To do that, they need to have more than a superficial understanding the technologies that make those companies valuable. As AI becomes more sophisticated and insinuates itself ever more into our lives, investors need to learn how to distinguish hype from reality and hope from achievement in this relatively new field. They should immerse themselves in the literature of technology, of which *The Myth of Artificial Intelligence* is an important, if challenging, part.

Laurence B. Siegel is the Gary P. Brinson Director of Research at the CFA Institute Research Foundation, the author of Fewer, Richer, Greener: Prospects for Humanity in an Age of Abundance, and an independent consultant. His latest book, Unknown Knowns: On Economics, Investing, Progress, and Folly, contains many articles previously published in Advisor Perspectives. He may be reached at lbsiegel@uchicago.edu. His website is <http://www.larrysiegel.org>.

¹This Erik Larson – Erik J. – is “a tech entrepreneur and pioneering research scientist working at the forefront of natural language processing,” according to Harvard University Press. The popular historian and journalist Erik Larson, who wrote the bestsellers *In the Garden of Beasts* and *The Devil in the White City*, is a different individual.

²DARPA, the military agency that pioneered AI decades ago, sets forth its own classification scheme in a very good YouTube video. The one I use is simpler and better suited to organizing this article.

³My Advisor Perspectives article on Claude Shannon, the inventor of logic circuits and many other key elements of computing and AI, provides more background.

⁴The Turing Award, given by the Association for Computing Machinery since 1966, is the computer science equivalent of the Nobel Prize.

⁵A detailed and eminently readable discussion of causal AI is at https://ssir.org/articles/entry/the_case_for_causal_ai.

⁶Bostrom's web site, worth reading, is <https://nickbostrom.com>. Yudkowsky's best known work is *Rationality: From AI to Zombies*. A brief introduction to his philosophy is in an interview by Scientific American's John Horgan at <https://blogs.scientificamerican.com/cross-check/ai-visionary-eliezer-yudkowsky-on-the-singularity-bayesian-brains-and-closet-goblins/>. It's a fun read. If you want to go more deeply into the evil-AI rabbit hole (where Yudkowsky sometimes wanders), read this. You've been warned.

⁷My emphasis.

⁸By “solved” I mean found an arrangement of neurons and synapses (connections between neurons) that serves the creature well. There is, obviously, no unique solution to the problem of how to build a brain.

⁹Some evolutionary biologists argue that the apparently backward placement of rods and cones is not a mistake.

¹⁰Some invertebrates have been observed to do this too, and they are even older, but turtle-to-turtle rescues are well known, even to children who keep turtles as pets.

¹¹The Peirce quotes are from “Peirce on Abduction,” *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/abduction/peirce.html>.

¹²Note that even primitive people who have never heard of hypotheses form and test them. Larson gives hunting as an example. Based on tracks, spoor, and other evidence, good hunters form theories (mental models) of where the prey is, what direction it's moving, and at what speed, then test the theory by trying to catch the animal.