BOOK REVIEW

The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do, Erik J. Larson, 2021, 320pp., $30 hardback, Belknap Press, Cambridge MA, ISBN 978-0674983519

The *Myth of Artificial Intelligence* is just asking for disagreement, its title a clickbait challenge to AI addicts. Indeed, it has more than a smattering of religious undertone. I would go further than the author, Erik Larson, and categorize the myth peddlers as artificial intelligence (AI) zealots. What is AI? Though not made explicit in this book, the question is explored through historical context, mathematical logic, technical implementation, and the dream of human and superhuman intelligence (whatever intelligence is). It's unclear what the author derides: is it the technologies under the banner of AI, big data projects or the idea of super intelligence?

Kitsch and exceptionalism

Let's look at the book's main assertions: AI is kitsch and the human mind is somehow exceptional. AI's boom and bust cycles are discussed as are the huge monetary cockups of big data projects, 'kitsch' encapsulating nicely Larson's stance on the subject: 'kitsch is a simplification of complicated ideas, the sweeping away with emotions and difficult questions of life'. Kitsch is an attack on the single-minded logic of AI's superhuman goal. It's also a fight against computational theories of mind (CTM), where minds, and by extension humans, must retain an element of complexity, an element of mysticism, of purity that is not undermined by reductionist computation.

For Larson, AI is a foolhardy destination that threatens to usurp everything in its path – money, scientific endeavour, and ultimately the mind. The myth makers, Larson claims, aim to solve everything with big data analytics and presume to be on the path to artificial general intelligence (AGI) by belief alone: 'simply saying "we're getting there" is scientifically and conceptually bankrupt'. Though I can completely agree that the hype machine and boom-and-bust cycles have plagued AI progress, I'm still not totally moved to admit the second premise of the book's title, which is that computers can't think the way we do. I find it hard to put faith in the exceptionalism of our minds, and am willing to take CTM seriously, unlike Larson.

This is truly a book that is best suited to an AI evangelist or perhaps someone sitting on the fence on such issues as technological determinism, singularity and CTM. Sceptics (I count myself one) will appreciate a thoughtful and engaging tour of mathematical and philosophical logic; a discussion of the origins of computing, the etymology of the word 'robot'; the Turing test and natural language processing (NLP) techniques (the author's interests really shine here). Each chapter is punctuated by the author's own critique, which usually involves pointing out some flawed logic or theory nit-picking. The book avoids the common approach of stating a claim then exemplifying it repeatedly and is worth reading through to the end.

So far, so good for an AI book. Where this one really stands out however, is in its debunking of the pervasive myths associated with AI. The author lets readers peek behind the curtain, reminding them that 'ideas have consequences', that 'the price we are paying for the myth is too high'. The first myth of AI starts with Turing and colleagues, where a concept of machine intelligence emerges as a reductive interpretation of human intelligence in the shape of problem solving. This sort of makes sense, given the limited computing resources and the abundant low-hanging fruit to which AI programs could be applied at the time. The field of AI was still to inherit a simplified model of cognition. Are we not all computers running on different substrates? In praise of Turing, though, the author does point to the degree to which Turing's test has been apt and enduring, as well as his ideas

on intuition and ingenuity. Humans have ingenuity and machines don't. One of the main thrusts of the book is that the path, and AI's final goal, are still unclear.

The author hopes to connect intuition, invention and inquiry. These anthropocentric terms and the discounting of machine intelligence expose a reverence for the human as creator. This reverence is almost antithetical to the assertion that humans, as inventors, would not be able to make artificially intelligent systems at some point. Perhaps Larson believes the problem is simply too Herculean. We can only assume that the author disagrees with the concepts of computational theory of mind wholesale – 'equating a mind with a computer is not scientific, it's philosophy'. However, 'the myth of AI insists that the differences are only temporary', the challenge of AI is in effect blind to the uniqueness of human cognition, it equates AI with the pursuit of intelligence but fails to encompass the fullness of human intelligence. This theme persists throughout, and is one to which I do not subscribe.

Later we are reminded that 'deep learning has been the hammer causing all the problems of AI to look like a nail'. But critically for the author, the renewed interest in AI from machine learning (ML) will be how problems are contextualized. The brain in the form of neuroscience, language in the form of NLP, even discovery and invention have been outsourced to big data systems. What will be left for us to do? AI is a set of tools, algorithms, logics, of means, as well as a destination, a philosophy, a dream. The issues of AI hinge both on myths and maths. In our zeal to adopt data *learning* systems we've fallen into a trap: 'society is about to experience an epidemic of false positives coming out of big-data projects'. And these failures will suck us all in, money and mind.

Logical inference

The author spends a great deal of time discussing three modes of logical inference: induction, deduction and abduction. These are critical to his argument of exceptionalism. The pitfalls of induction are exemplified in the author's reference to the parable of the turkey. Imagine you're a turkey. You're fed and treated well, you infer that life is good and you are well looked after day after day. Your mental model of the world can only strengthen this observational claim each day until your last morning … perhaps the day before Thanksgiving or Christmas. For the author, inductive logic fails to give context or real understanding, it's unclear how it can lead directly to AGI. Lawson never contextualizes understanding and knowledge so the book offers no definite answers about what they are and how they contribute specifically to intelligence. Inductive inference tends to view the future as looking like the past. It is dogmatic in nature, and can only ignore anything that sits outside its correlations. Inductive systems often produce spurious results and cannot be put to work creating formal logical claims of their own.

For me, ML techniques are still a toolbox worth exploring, especially for problem spaces where no other systems have previously succeeded. A human cannot, for instance, listen to thousands of sounds simultaneously and pick out coughs to detect a viral infection; or look at millions of images, labelling artefacts as they go. Machine learning systems can. However, these systems still need guidance: a system looking at images of polar bears or huskies may focus on the snow in the backgrounds rather than any particular differences in the animals. I agree with the author, we still have a big part to play in the dance of algorithmic progress, even if the importance of this role is often overlooked.

Those in expert systems or the symbolic AI camp will also find familiar arguments here. The book draws a line between deductive reasoning systems and issues of internal representation, data volume, ingestion and dogma. For example, a system that knows the road becomes wet when it rains could not deduce that the road is wet because it is raining: the posterior tells you little about the prior. It is also an extremely laboursome process getting these types of systems to utilize knowledge. Knowledge tends to be static, unchanged by new data in the environment (maybe a fire hydrant has gone off and is spraying the road with water). Though not discussed in the book, a small child would probably take updates in knowledge on board (and let out a gasp of surprise in the process), an update cycle we are still unable to emulate in software.

The abductive Holy Grail

Why not combine induction and deduction? Such systems as IBM's Watson are portrayed as hybrid systems, systems that are extremely good at doing one thing and doing it well. In the author's description of Watson, as well as the discovery of the Higgs boson at CERN, we are introduced to the novel abductive reasoning ability of humans, the third pillar of our logical tour. The intuition of human experts provides that abductive backbone for these systems, a linchpin in order to discover theorized phenomena. Humans provide direction and a goal. However, abduction does not yet have a good computational analogue and is thus described logically, mired in mystery. In this way, abductive reasoning is the Holy Grail, a missing ingredient for AGI. If we could compute abduction, we would be able to create human-like guessing, leaps of faith, intuition and invention. Here is where I feel the author runs afoul of his own mythos: abductive logic and the role it plays in human thought becomes magical and unknowable. The myth of AI becomes the myth of intelligence generally. Without a goal to measure our progress, AGI becomes slippery and diluted, a sea of dreams.

The first of these dreams is a return to the Turing test. Eugene Goostman, the AI chatbot, passed the test in 2014. What the chatbot, or rather the human programmers, were able to do is, in effect, game the test, working around the biases of the human judges without any real understanding. The hype surrounding the Turing test almost wills it to be passed each time. Passing the test properly will probably require feats of general understanding, of knowledge utilization, of common sense. But as the book points out, we're not there yet. Goostman's Turing test fudge and Watson's Jeopardy are good examples of human ingenuity coupled with deductive and inductive learning systems. An update to the Turing test is discussed in the form of Hector Levesque's Winogard schema.[1] The schema addresses possible cheating and makes it seemingly impossible. The changing nature of the language problem requires sophisticated semantic and grammatical understanding in order to complete the simple (for humans) common-sense multiple-choice test. To win, the agent requires prior general knowledge of things: why a crocodile couldn't win a steeplechase, why a pregnant woman might stop taking obviously dangerous pills, what things can be too big to fit in smaller things. It requires a compendium of priors, the semantic rules of language (in this case English) and abductive jumps of logic, a way of simulating plausible answers for the context.

The book's casual attacks on such projects and institutions as Google's Deepmind, IBM's Watson and the EU's Human Brain project makes me wonder if there isn't a false dichotomy being presented. Are all AI evangelists trying and failing to produce AGI at all cost? Do big data projects actually fail because of their hammer-and-nail reliance on ML to solve all problems? Are we actually replacing human intuition, invention, with pattern matching systems? The author puts together a good case, but maybe the situation is more complex than the threads of the book allow. Though I still see a misunderstanding between what we call AI today and what it can achieve (most prevalent in the business world), it seems excessive to throw out what progress has been made. Maybe we should moderate our appetite. I wholeheartedly agree with Larson's conclusion concerning ML being the latest hammer, but a binary division between AI proponents and opponents seems a little reductive.

Recipe for AGI

The book seems to deride AI progress and absconds from posing a counter direction to guide developments, instead opting to warn against the fallacies of an AI panacea. AI systems today seem to be brittle, idiot savants, incredible in constrained environments, useless in others. To get us to more competent AI, we will need ingredients. Inspired by Larson's work, here is a first best guess at what that recipe list should contain:

---

[1]http://commonsensereasoning.org/winograd.html

- an abductive layer
- self learning (and abductive enabled self-improving?)
- an ability to utilize symbolic knowledge which it has built itself, presumably using inductive and deductive strategies
- the ability to be wrong/uncertain.

The ability to be wrong is an extension of the proposal to counter AI running amok. In a world predicated on scientific discovery updating our prior assumptions (consider Newtonian physics and relativity), surely the uncertain AI is king, but hopefully not master!

Dreamers and experts

Staying the course, the book points out that our AI tools are idiot savants, able to look through and recall vast amounts of data, but without the ability to make general use of them. But what of the counter example? Consider the alternative. Let us call the ML systems 'experts' instead and the uncertain AIs 'dreamers'. It strikes me that in the future our lexicon will probably shift again, demoting ML experts to simply data tools, sophisticated data analysis, or similar. This leaves our dreamers as a continuation of the dream of artificial personhood. But is this better? Do we actually want to create strange beings composed of uncertainties? Or concentrate on solutions to wider issues such as climate change, poverty, equity, diversity and so on. You can insert your own preferred version of AGI. The future, of course, is unknown, but as the author points out, maybe the dream of AI is indeed costing too much in the present.

Human purpose

The last theme of the book asks what is left for us to do. Will scientific discovery continue to drive our societies, our interests, and in some ways, our purpose? Or have we already picked all the low-hanging fruit, leaving us to create increasing complex machinery to pluck the last remaining natural laws? Perhaps big data analytics really is the appropriate tool to sieve through the dirt in order to find nuggets of gold left for us, and our role is simply to move the sieve around until we see something shiny. If 'we're not out of ideas, then we must do the hard and deliberate work of reinvesting in a culture of invention and human flourishing', but, in a nihilistic sense, if we are out of new ideas, then we are faced with a greater problem: what purpose of our existence?

The Myth of Artificial Intelligence leaves us with a palpable sense that the global AI endeavour is bad and our whole concept of science might be an illusion. It feels like a cliff-hanger for a second series. Hopefully we can pick ourselves up, look at what is good about humans, our intuition and our ingenuity. I agree with much in this book, but not with the overarching premise that we are irreducible entities, innovative. Yet an unseen hand prevents us from knowing ourselves and creating anything in our likeness. This is too defeatist to contemplate. I've also been struck by the need for abductive software analogues, as well as head nodding through the idea of kitsch. Perhaps my main takeaway is that we should face the facts and forge forward in the knowledge that AI is not some messiah coming to save us – unless, of course, you believe it is.

Ben Byford
Ethical by Design
hello@benbyford.com