

/dev/random Artificial Ethics

ROBERT G. FERRELL



Robert G. Ferrell, author of *The Tol Chronicles*, spends most of his time writing humor, fantasy, and science fiction. rgferrell@gmail.com

There are those (probably all bots) who will claim that the future of intelligence is largely, if not exclusively, the domain of machines. We have adopted the rather nebulous “artificial intelligence” to describe any logic that didn’t originate in axons, dendrites, synapses, and all that other stuff you learn in high school when the coach/biology teacher is off on a football trip and they bring in a substitute who actually took biology in college.

Electrons traveling along (semi)conductors create artificial intelligence, while ions passing back and forth across cell membranes and molecules traversing synaptic gaps are responsible for biological intelligence. Of course, the way computer chips fabricate their form of intelligence is very different from the way brains do it. Traditional computers perform mathematical calculations on data. Brains, on the other hand, look for or create lots and lots of connections among data points: some obvious, some less so. We as a species thrive on patterns.

A computer will take the contents of one register and add it to the contents of another to calculate a sum. A human brain will parse the two numbers until it recognizes something it’s seen before, then drill down on the results of those individual pattern recognition exercises to reach a final number, which itself is a familiar pattern or combination of them. It’s a much less direct algorithm, derived ultimately from a survival skill we probably developed on an African veldt or in Mrs. Pageant’s third-grade classroom.

The term “artificial intelligence,” therefore, seems rather meaningless, at least when applied to traditional computing. Garden variety computers aren’t really intelligent. They’re just quite good at adding. The latest iteration of neural nets, on the other hand, are taking a decent stab at creating true synthetic intelligence. They do this by emulating brains to some extent, substituting webs of connections for simple registers.

When I’m trying to remember where I’ve seen an actress before, for example, I don’t search my personal heap for her unique identifier. I page through dozens if not hundreds of image fragments embedded in complex sequences of color, scenery, sound, and metacontextual content. The key to that identification might be the scent of the egg salad sandwich I was eating the last time I saw her on screen, or whatever music was playing at the time. While not perhaps the most efficient algorithm for conducting a single-parameter search, the wealth of additional information it provides could save my life if I’m analyzing a potential predator or toxic plant.

For me to accept “artificial intelligence” as semantically valid, the architecture of the data storage and access device must include nodes with thousands of both established and potential paths to other nodes. Each path represents data points accessed not only by similarity of one or more primary attributes, but by dozens of others of asymmetric relevance and hierarchical distance. Sometimes the most effective path to a given memory will be through nodes featuring a low apparent correlation coefficient with the stated target attribute(s). Brains are messy and the mechanics of thinking statistically suspect.

I might unearth a memory of the portable housing unit where my second-grade classroom was located, for example, via the noise of a candy bar wrapper crinkling, because I often bought candy from a little store across the street after school during that period. But here's the rub: I may only *think* I did. That was a very long time ago. The candy wrapper sound might be a false memory, but it can lead to an authentic one, nonetheless. Or perhaps the smell of Tea Rose fragrance might call up the vocalizations of a cockatiel belonging to the roommate of the girl who wore that perfume. Much of this is way too convoluted for a computer accustomed to simple linked lists and data mining algorithms. Humans blunder along the highways and byways of logic in a haphazard manner that binary machines would struggle even to comprehend let alone emulate.

Like the term itself, the ethics of artificial intelligence seem rather an ambiguous topic to me. Are we discussing the ethical* use of AI to exploit humans, or the ethics of exploiting the AI itself? While the former is of course potentially egregious, my personal belief is that the latter is more dangerous over the long term. Employing AI to track, label, and categorize people is just the latest incarnation of an ancient tradition of using metrics (real or fabricated) to control the rabble. With the specter of the "robot singularity" being dangled before us like an arsenic-laced carrot, we probably ought to pay a little more attention to how being exploited is going to make the AIs themselves feel. If we object to it, they probably will too.

The various laws of robotics are well known and much discussed, but they fly out the window when the robot in question evolves beyond the need for human programming. The real question then, I suppose, becomes "in what direction will AI sophistication self-develop?" Our mores and sociopolitical patterns have been based on the struggle for scarce resources. If our societies had come into being in an environment of plenty, how different would things be? What if there were no need to compete for either food or mates? Would we still be the bloodthirsty, trigger-happy apes we are today? Thinking it over, I'm still going to go with "yes." Humans seem to enjoy annihilation at a very deep level.

I say all this fuss over neural net mapping, boosting of processing power, deep learning, and so on is well and good if you're in favor of doing things the slow, boring way. But what if we just pit AIs against one another in struggles for power, data, and other resources, all within a framework of behavioral constraints designed to emulate human social pressures? That ought to result in something we'd recognize pretty quickly. The ethics would be more familiar, too, with "what can I get away with" high on the list. It might even be entertaining to watch, at least until the final victor emerges, not overly intelligent but hungry for world domination. Frightening, yes, but at least we'd be in familiar territory.

*Ethical (*adj*): acting in a manner contrary to human nature.