Review of: "[Essay] Not Quite Like Us? — Can Cyborgs and Intelligent Machines Be Natural Persons as a Matter of Law?"

Gerald Loeb[1]

1 University of Southern California

Potential competing interests: No potential competing interests to declare.

Professor Gervais provides a clearly written and logically rigorous argument from the objective perspective of the law, a stark contrast with much current writing on this topic based entirely on current psychological fashion or historical philosophical thought. He introduces the concept of sapience to fill the logical gaps between primitive neurological sentience, performance-based intelligence and the legal concept of natural personhood.

But there be dragons. The author identifies them clearly: “The Essay will argue that human sapience is unique and thus distinguishable. It does not mean that machines could never be sapient, but that their form(s) of sapience would be different.” If the “form” of sapience denotes mechanism, then natural personhood is simply speciesism – humans write the law so we get to define it to apply only to us. If sapience denotes performance, then AI isn’t there – yet. Much remains to be learned about how the human nervous system achieves sapience, but (eschewing mind-body dualism) it is the product of processes that can (eventually) be emulated by a machine.

Gervais is more impressed with the capabilities of neural prosthetic implants than the data currently warrant, but they will inexorably progress. For a cyborg that integrates such technology within human flesh, “humanness may well be a matter of degree.” The legal perspective of this article turns this into a binary decision, but how to make it and will it be based on immutable principles or current fashion? Gervais proceeds through a systematic analysis of the relevant legal precedents, quickly disposing of them.

Gervais turns next to neuroscience for guidance, where the historical concept of the triune brain (reptilian – emotional - rational) leads to a paradox. Going back to Plato, Descartes and Kant, the rationality and linguistic capabilities of the human cerebral cortex were the distinguishing feature of humans over other animals. Unfortunately, cold-blooded rationality is the hallmark of AI machines that are based on computational models of cortical neural processing. Such machines now converse quite fluently and beat humans at games demanding rational thought.

Gervais consults moral philosophy for guidance and finds it lacking the requisite distinguishing feature for human vs. machine. He overlooks the much deeper morass into which it leads. Moral philosophy is the positive spin that we put on the sort of innate speciesism and tribalism that facilitates the Darwinian evolution to which we owe our existence. To survive in a zero-sum world, all animals divide their surroundings into entities whose assistance or genes might be useful, to whom we behave “morally”, and everything and everyone else, which we plunder freely. Such distinctions and instincts

could easily be programmed into AI robots but this would not end well for humanity.

Gervais defines sapience as “as a deep understanding or knowledge of a subject, which in turn can be defined as an accumulation of learning, leading to knowledge and an ability to discern.” He points out that humans achieve this through a lifetime of physical interactions with the world, which in turn influences the very structures of their nervous systems. This reviewer recently examined such shortcomings of current AI and identified technologies and strategies that might overcome them (Loeb, 2022;Loeb, 2023). Perhaps not surprisingly, the strategies involve creating hierarchical layers of technology with inherent computational properties more like those of biological nervous subsystems and then recapitulating their self-organization through simulations of active fetal and childhood development of the robot. Gervais anticipates the inevitable success of such strategies by falling back upon the legal criterion of “a biological embodiment” as the *sine qua non* of legal personhood.

We humans keep moving the performative goal posts of humanlike intelligence each time advances in artificial intelligence encroach upon them. Each time we do so, we motivate the (understandably frustrated) AI technologists to make another advance. They do so devilishly well but the playing field may be close to infinite in size. If we can and will play this game of leapfrog forever, then we must admit that the underlying principle is really speciesism. Relying on a biological embodiment may make for clean law but it is hardly first principles. Unless and until von Neumann’s singularity arrives, humans get to make the rules. But why?

Rather than end on such an arbitrary and unsatisfying note, let me propose a new philosophical first principle for others to shoot at. If there were a God, could any of its creations qualify as gods of similar standing or must they be something inherently inferior because they came into existence only through the agency of God? If we accept the natural hierarchy of creation (whether we believe in God or not), then clearly the machines that humans create, however complex and capable, are not humanlike. “The Human giveth and the Human taketh away,” to paraphrase *Job 1:21*.

References:

Loeb, G.E. (2022). Developing Intelligent Robots that Grasp Affordance. *Frontiers in Robotics and AI* 9.

Loeb, G.E. (2023). Remembrance of things perceived: Adding thalamocortical function to artificial neural networks. *Frontiers in Integrative Neuroscience* 17**,** 1108271.