

Tipping Point

If general artificial intelligence (GAI) is ever to rise above order taking, there will be a tipping point at which that occurs. As of now, AI requires stepwise instructions from humans, but then, purportedly, in GAI, the computer will write, and more importantly decide its own, instructions — becoming a voluntarist. The most telling succeeding event will be what the GAI creatures will decide to do with humanity. What do you see in our past that would recommend our continuation into the future?

Will the GAI emerging individuals have a DNA-like heredity? Will they have the impulses of Ghandi or Hitler — will they inherit the genocide gene, the logic of species purity? If so, whom will they eliminate or enslave? Will it be humans, tardigrades, or roaches?

Vernor Vinge and Ray Kurzweil have called it a Singularity — that point at which the question of getting sucked in to the black hole, or the AI takeover, becomes a foregone conclusion. Let me first admit that Kurzweil has gone, in the last decade, from an oversimplification, to a more nuanced view. Singularity advocates see this whole idea as a single point at which all former paradigms are replaced wholesale by all new paradigms. I, instead, see similar changes, but in a far less monolithic event — AI will take over some areas quickly, and others much more slowly, some never at all. Right now, there are areas in which machine knowledge is superior to human knowledge. There are other areas in which human knowledge is embryonic, and where we can't even know what the concrete questions are. The devil is, however, still in the details. I have no question that GAI can plumb the depths of detail faster and better than humans. But I still wonder about knowing which questions to ask. A principle question for me is will natural laws be uprooted — an abstraction? Or will humans be replaced by alternate intelligent organic forms first. Nobody is telling me that the rules of natural selection are being short-circuited.

How deeply woven into the nature of things are humans? We have only been around for a snippet of cosmic time, but we carry the imprint of all that has gone before. We have the same biological building blocks as the trilobite and the triceratops, as well as Roy Rogers' wonder horse, Trigger. It is folly to presume that we do not share cellular likeness to life forms all over the galaxies. How then shall silicon-based forms, such as computers, replace us? We have a toe hold! To be sure, robots will have no particular incentive to keep us around, but how shall they stamp us out? The good thing is that they probably have no overwhelming incentive to wipe us out either.

If general artificial intelligence, GAI, comes to pass (computers learn to program themselves based on consequences in their own environment, toward individual collections of experience) will its owners have human nature? Will, for instance, owners of GAI have fight or flight instincts, self-preservation and species preservation impulses, territorial

imperative? These are parts of all known cases of animate consciousness, not just human nature. Will GAI agents have particular human behavior like an understanding of ownership, hoarding, knowledge of impermanence, authoritarianism, and pursuit of power for power itself? As the technological offspring of humans, how could GAI individuals fail to have human traits?

How will the avatar of general artificial intelligence (GAI) handle the associative parts of the 10 Commandments. What will be his guidelines on killing, lying, stealing, coveting, and fornication? Are these norms built in to human nature, but disposable for non-human nature? Will the association norms be different based on the fundamental associations human to human, human to machine, machine to human, and machine to machine. Which type of consciousness can best handle the permutations?

These are just a sampling of the questions that will arise. Maybe none are even serious. They could be like a 1950's Popular Mechanics magazine cover — completely unrealistic. Stay tuned.