

BUILDING GODS OR BUILDING OUR POTENTIAL EXTERMINATORS?

Prof. Dr. Hugo de Garis

profhugodegaris@yahoo.com

Abstract

Hugo de Garis is concerned that massively intelligent machines (“artilects”) could become infinitely smarter than human beings, leading to warring factions over the question: should humanity risk building artilects? Result: gigadeath. (See the author’s “The Artilect War” book (amazon.com) for further details.)

Originally published February 26, 2001 on KurzweilAI.net.

Robot artificial intelligence is evolving a million times faster than human intelligence. This is a consequence of Moore’s law, which states that the electronic performance of chips is doubling every year or so, whereas it took a million years for our human brains to double their capacities.

It is therefore likely that it is only a matter of time before our machines become smarter than we are. It is also likely

that this development will occur this century if humanity chooses to allow it to happen.

My name is Prof. Hugo de Garis. My team and I are starting to design and build the world's first artificial brain at my lab, Starlab, in Brussels, Belgium, Europe, which should contain nearly 100 million artificial brain cells (neurons). In about 4 years, the next-generation artificial brain should contain a billion neurons.

Our human brains contain roughly 100 billion neurons, so it is not surprising that someone like me is preoccupied with the prospect of robot intelligence surpassing the human intelligence level. Admittedly, massive computational speed and size do not automatically equate to massive intelligence, but they are prerequisites. The potential is there. My brain-building team still faces the considerable challenge of architecting the artificial brain. We will need to become “BAs”–Brain Architects.

Despite this qualification, not only do I believe that artificial brains could become smarter than human beings, I believe that the potential intelligence of these massively intelligent machines (which I call “artilects” (artificial intellects) could be truly trillions of trillions of trillions of times greater.

If these astronomically large numbers sound like science fiction to you, consider the following. Moore's law is achieved by shrinking the size of electronic components such as transistors by a factor of two roughly every year. This halves the distance between components, and hence

doubles the speed at which electronic signals can move between them (at the speed of light, a constant of nature). This trend has been valid for 30 years, and is likely to continue until 2020, by which time the scale of electronic circuitry will have reached atomic levels.

In other words, within a single human generation it will very probably be possible to store a single bit of information on a single atom. There are a trillion trillion (a 1 with 24 zeros after it) atoms or molecules in objects of human scale, such as an orange. An object as large as an asteroid (to be found in the asteroid belt circling the sun between Mars and Jupiter) can be hundreds of kilometers across and contain a trillion trillion trillion atoms.

The bits stored on these atoms could switch (bit flip) from a 0 to a 1 and vice versa in a femtosecond (a thousandth of a trillionth of a second). That's an information-processing capacity of about ten million trillion trillion trillion (a 1 with 55 zeros) bit flips a second. When one compares the comparable information-handling capacity (in bit flips per second equivalent) of the human brain, the estimated answer is about ten thousand trillion bit flips a second (a 1 with 16 zeros), which is a thousand trillion trillion trillion times smaller. These artifacts could potentially be truly god-like, immortal, have virtually unlimited memory capacities, and vast humanly incomprehensible intelligence levels.

I foresee humanity splitting into two major ideological, bitterly opposed groups over the "species dominance" issue, i.e., should humanity build artifacts or not. These two

groups I label the “Cosmists” (in favor of building them) and the “Terrans” (who are opposed).

To the Cosmists (based on the word “cosmos”), building artifacts will be a religion (compatible with and based upon modern science), as the destiny of the human species and as the magnificent goal of creating the next rung up the ladder of dominant species.

To the Terrans (based on the word “terra,” the earth), building such artifacts means accepting the risk that one day, in an advanced state, these artifact gods might decide, for whatever reason, that the human species is so inferior and such a pest, that they should exterminate us. With their gargantuan intellects, such a task would not be difficult for them.

The Terrans, in the limit, will try to exterminate the Cosmists if the latter insist on building artifacts, for the sake of preserving the survival of the human species. Since the stake is so high (namely whether the human species survives or not) the passion levels will be high. The Cosmists will anticipate the murderous hatred of the Terrans and will defend themselves.

We have thus all the makings of a major war. About 200 million people died for political reasons in the 20th century (wars, purges, genocides, etc) using 20th century weapons. Extrapolating up the graph until late 21st century, with 21st century weapons, we arrive at billions of dead—”gigadeath.”

So which am I, a Cosmist or a Terran? I'm both. Ultimately, I think it would be a cosmic tragedy if humanity chooses to freeze evolution at the puny human level (with our pathetic little lives of 80 years in a universe billions of year old, that contains a trillion, trillion stars—the “big picture”). For me, the tragedy of seeing the human species wiped out is less significant than not seeing the birth of the artefacts. This sounds monstrous, and it is, in human terms, but to deny the creation of the first true artefact, which would be “worth” a trillion, trillion, trillion human beings, would be a far greater tragedy, a “cosmic” tragedy.

As the planet's pioneering brain builder, I feel a terrible burden of responsibility toward the survival of the human species and the creation of godlike artefacts, because I am part of the problem. I am quite schizophrenic on this point. I would love to be remembered after I'm gone as the “father of the artificial brain,” but I certainly don't want to be seen in future historical terms as the “father of gigadeaths.”

Hence I try to raise the alarm now while there is still time before the artefacts come into being. If I were a true Cosmist, I would keep quiet and just get on with my work, but instead I feel that humanity should be given the chance to nip the rise of the 21st century artefact in the bud if it so chooses.

So should work on artificial brains be stopped now? I think not. For the next 30 years or so, brain-based computers will be far too useful to be suppressed. For example, they will

become smart enough to clean the house, teach the children, provide sex, and help human experts in decision making, etc. They will do most of the work and thus create great wealth for the whole planet.

So, in the short to middle term, brain building technology will be seen as a great boon to humanity. It is the longer term that terrifies me and keeps me awake at night. I see no way out of a gigadeath artelect war, so relentless is the logic.

The rise of the artelect will probably be inevitable. The economic and military pressures to build them will be enormous—hundreds of trillions of dollars a year worldwide will be spent in the brain-based computing market within 10-20 years, I believe. The debate over whether artelects should be built or not is already starting to heat up, at least amongst the researchers concerned with brain building and AI (artificial intelligence).

This debate is starting to spill over to other specialties. For example, I'm trying to persuade Prof. Peter Singer (Princeton University), the planet's best-known "applied ethics" professor, to write a book about "Artelect Ethics." At the rate at which this issue is hitting the headlines lately, my bet is that within a few years the "artilect debate" will be on everyone's mind.