

Unfriendly AI and its Consequences

Prof. Dr. Hugo de Garis
profhugodegaris@yahoo.com

ABSTRACT

This essay argues that the notion of “friendly AI” is both deluded and dangerous. It presents a series of arguments that show that it is virtually impossible to design an intelligent machine that remains “human friendly” no matter how increasingly intelligent it becomes in the future by modifying its own structure. This impossibility will have profound political implications for humanity. As a consequence of this, the essay then argues that passionate differences over the issue of “species dominance” (i.e. whether humanity should build godlike massively intelligent machines this century or not) will lead to the most passionate war that humanity has ever known, using 21st century weaponry, that will kill billions of people. 21st century technologies will allow humanity, if it chooses, to build creatures (called “artilects” i.e. artificial intellects) with intellectual capacities trillions of trillions of times (and more) above the human level. Humanity will then have to decide whether it will become “number 2 species” and abandon control over its own fate. The essay argues that humanity will split largely into 3 political ideological groups, labeled respectively “Cosmists” (who will look upon artilect building as “god building”, a science based quasi religion), the “Terrans” (who will be opposed to artilect building, seeing artilects as an existential threat to humanity’s survival), and the “Cyborgists” (who will choose to become artilects themselves by adding artilectual components to their own brains). The essay argues that the Terrans will lump the Cyborgists into the same ideological camp as the Cosmists, since the computational capacities of 21st century nanotech matter is so high that merely adding a grain of nanotech sand to a human brain will outperform the latter by a factor of trillions. As the home robot industry grows in the next few decades, billions of people will see their own robots growing in intelligence every year, causing the temperature to rise in the world wide “species dominance debate”. The year 2011 will see this debate become main stream as several movies and major TV documentaries on the topic will appear. Surveys have already shown that this issue divides people almost 50-50. The essay argues that the dominant question of “Who should own capital?” that so divided humanity in the 19th and 20th centuries will be replaced by the question “Should humanity build artilects?” The rival answers to this question, i.e. “yes” from the Cosmists and Cyborgists, and “no” from the Terrans, follow from two very powerful ideologies that will become increasingly bitter as the decades pass. The Terrans know that their “time window of opportunity” to defeat the Cosmists is limited. If the Terrans do not act quickly enough, the artilects and cyborgs will have become so advanced that any military struggle between the Cosmists/Cyborgists/artilects on the one hand and the Terrans on the other, would be like (quoting Ray Kurzweil) “the US army fighting the Amish.” To avoid this, the Terrans know they must “first strike” and will prepare for this. The Cosmists in turn will anticipate this and be ready. Since this “global civil war” will take place in the second

half of the 21st century, using 21st century weapons, the scale of the killing in WW3 will not be in the millions, as in major 20th century wars, but in the billions (“gigadeath”).

Bio :

Prof. Dr. Hugo de Garis was director of the Artificial Brain Lab (ABL) at Xiamen University, China. He is now retired and spends his time studying pure math and mathematical physics in order to write books on “topological quantum computing” and “femtotechnology”, etc. He is the author of two books “The Artilect War” from which the ideas in the above essay are taken, and “Multis and Monos” which argues that a global state will probably be in place before mid century (both available at amazon.com). He was selected by History Channel to introduce the “species dominance debate” to a world public, in a program “Prophets of Doom” that first aired in January 2011.