

# Prometheus Unbound<sup>i</sup> : A Rational Approach to AI and the Future of Humanity

by Jack Peggs

I

And you will hear of wars and rumors of wars. See that you are not alarmed, for this must take place, but the end is not yet. For nation will rise against nation, and kingdom against kingdom, and there will be famines and earthquakes in various places. All these are but the beginning of the birth pains (*English Standard Version, Matt. 24:6*).

---

We didn't start the fire  
It was always burning since the world's been turning.  
We didn't start the fire  
But when we are gone  
It will still burn on, and on, and on, and on....  
(*Billy Joel, Lyrics "We Didn't Start the Fire", 1989*)

There is what we must recognize in humankind as we step into the twenty-first century. That is the immensity of our problems. There are seven billion of us on this planet. In what we call the third world, millions of us are in starvation; conflict and wars proliferate on every populated continent, and, occur with the frequency of every decade. Famine, faunal extinctions, pestilences besiege us wherever we are found; societies teeter on the brink of self-immolation. We are afflicted by an array of problematic issues. In some ways, it has always been such. Human existence has seldom been without an abundance of conflict, disease, natural disasters, and, miseries brought by the vicissitudes of fate.

We inexplicably survived in the past. However, the twenty-first century may be different. In recent decades, technology has reached a crescendo with devastating force and destructive power. Meanwhile, our dystonic behaviors flourish with increasing vigor, and, in proportion to the many millions of inhabitants on the globe. Unlike any other era before it, our planet faces grave risks to which antiquated beliefs hardly have an adequate response.

Thomas Malthus, an early nineteenth century economist and scholar of demography, believed that eventually humankind would meet its match. By observing that populations grow geometrically, and, resources only arithmetically, in the end, Malthus argued, there had to be a demographic eclipse of resources with accompanying conflicts. There have been Malthusian crises<sup>1</sup> in the past as humanity has consistently proven itself proficient in the creation of innovative technologies, but, habitually failed to recognize the limitations imposed by their exploitation.

When populations increased as a result of the change from a nomadic existence to established agrarian cultures with their many advantages, there came cycles of shortages and inevitable conflict whenever resources dwindled. The rise of autonomous communities, city states, and, nations, enhanced the development of culture and communal benefits, but, it also sponsored periods of lethal conditions and environments resulting in wars, famines and localized misfortunes, at times, threatening human survival.

With the rise of industrialization in the later nineteenth century, humanity managed to stave off that last cycle of geometric progression which would have yielded the total collapse of humankind predicted by Malthus. With the industrial revolution, the influx of new technologies and new sources of energy, catastrophic collapse had been averted. Science and technology, thus, stepped in to avoid the global equivalent of Easter Island<sup>2</sup>. Still, as the world's pollution output continues to grow in population, as two

---

<sup>1</sup> This concept was described by Thomas Malthus in his *Essay on the Principle of Population* in 1798. In his work, Malthus observed that the pattern of human population growth was at extreme variance with the rate of resource replenishment. The inevitable outcome, according to Malthus, would be to create radical demographic and social changes for the worse, perhaps, leading to the extinction of the human race.

<sup>2</sup> The Polynesian island of Easter Island, or, Rapa Nui, is often cited as an example of a prosperous civilization followed by a rapid decline. At the time of its European discovery in the early 1700's, the island was found to be devoid of the resources or artisan technology necessary to account for its multitude of impressive quarried idols situated throughout the island region. These god-like statues, called "moai", evidenced too exquisite craftsmanship to have been the product of the primitive culture found on the island. In addition, there were virtually no trees as would serve to distribute these several ton monuments from their quarry pit throughout the island area. The most current studies suggest that Easter Island was originally settled in the fifth century, and, over the succeeding thousand year period, its indigenous population so thoroughly depleted the island of wood and forestation that it resulted in the destruction of a sizable segment of population, stripping the culture of its affluence.

thirds of the world lives under impoverished and mal-nourished conditions, and, as mankind's religions and antiquated philosophies continue to rule our cultural stratagems, the question remains whether the twenty-first century, with its unprecedented perilous reach, may survive yet another Malthusian cycle. Can science rescue us again?

In this new century, modernity finds itself facing a peril of much enhanced magnitude beyond that of any prior experience. Without a tutelary hand to guide it, our "brave new world" of rampant technology may lack the discipline and refinement necessary to resolve the crisis on humanity's doorstep. We must recognize a moral and intellectual shift in dealing with the waste and detritus infecting the oceans and atmosphere of this planet; in reconciling the social dysfunction borne by indulgence and excess in one part of the world, and, grinding poverty in another.

We need a new paradigm, in directing our technology, one worthy of, and, consistent with our conquest of the atom and that of taking our first steps on the moon. As our childhood, at last, draws to an end, our new age requires that we find the necessary wisdom not out of votive offerings, but, from within ourselves.

Demiurgic forces may certainly be necessary to rescue the twenty-first century; but, if any are to be invoked, their efficacious properties must be reality based and sufficiently endowed with a moral component to coordinate, manage, and, to compel a corrective response to the many problems facing the seven billion inhabitants of this planet. The paradigm shift needed for this new millennium is a superior technology, but, one in which there is a union with a matching ethos suitable for its navigation.

## II

Listen and understand. That Terminator is out there. It can't be bargained with. It can't be reasoned with. It doesn't feel pity, or remorse, or fear. And it absolutely will not stop, ever, until you are dead. *Kyle Reese (Michael Biehn) dialog from The Terminator, directed by James Cameron (1984)*

Futurists and others with technological expertise point to the nascent explosion of nano technology on which computer science is based. Among them are those who

claim that such technology's exponential growth, in combination with its unpredictability, presents grave risks to humanity. The fear is that, as the increasing sophistication of artificial intelligence (or AI) surpasses human capability, it will develop resource acquisition and efficiency drives to better serve its programmed purposes. As super intelligent AI progresses to fulfill these purposes, it will appropriate more and more resources, become increasingly efficient, and, when its prodigious output begins to suffocate us, it shall also have acquired effective defense mechanisms to thwart disruption by human intervention.

In an even more extreme description of the dangers of AI, critics claim that AI will develop goals of its own, and, pose purposeful risks to humanity. The Hollywood Terminator film series describes a horrific future in which robotic machines acquire self-awareness, and, in so doing, develop objectives at odds with humanity's. It is not a future any of us would want for our prodigy.

One must certainly question whether such fears are realistic within the context of what a superior algorithmic system is likely to do, particularly, an AI which obtains self-awareness. AI will lack a culture or background even remotely similar to anything in the human experience. How will it contextualize the outside world? Without a biological physiology, what automated counterpart to human desire, ambition, envy and other "feelings" will AI possess; and, with what implications for its behavior? The concept of consciousness itself is not sufficiently understood to even test for<sup>3</sup>, much less, to determine its dynamics and relationships to motives, intention and planning. We currently have no reliable theory of the mind.

Such unknowns, therefore, make the creation of AI nothing less than the manufacture of an alien entity with intellectual properties far exceeding those of any human being on

---

<sup>3</sup> The closest we have to testing for self-awareness in machines is the Turing test. Named for Alan Turing's proposal in his 1950 essay, "Computing Machinery and Intelligence," the Turing test is a protocol designed to determine if the behavior and responses obtained from a computer are distinguishable from those expected of a human subject. If the computer responses are indistinguishable from what would be expected, the machine is thought to be intelligent, and, presumably, experiences some form of self-awareness.

the planet. Humanity simply would be in no position to comprehend what it had created until, perhaps, too late. Without more understanding of brain chemistry and physiology, both the manufacture of an artificial mind and the reverse engineering of the human brain are too speculative and too perilous of enterprises.

Yet, the promise of AI lies precisely with the intellectual capabilities of a technology which magnifies human capability many times over. In spite of its many perils, AI promises to be the intellectual boost so necessary to our survival.

### III

What a piece of work is a man, how noble in reason, how infinite in faculties, in form and moving how express and admirable, in action how like an angel, in apprehension how like a god! the beauty of the world, the paragon of animals. (Hamlet Act 2, scene 2)

There is another model for the implementation of AI, one which would not require appeal to an alien creation. Even though we have no theory explanatory of the processes of the human brain, we do have direct experience with our own brains. We know a great deal about ourselves. We have produced sciences and arts which have studied our behaviors, motives, physiology and tendencies. We have a history and experience to draw upon, and, we know that of all the possible “thinking machines” which might or could exist, that the human being is capable of creativity and of yielding heuristic responses to the problems facing us. Of even more importance, even though we have many times and in many ways acted contrary, we possess a penchant for altruism, and, an orphic drive to achieve the ideal.

After all, it is something to be human, to have arisen from that craven rodent like mammal which scurried about on the Mesozoic forest floor; to have, at last, stood upright; to have built and founded empires; and, now to dominate the globe. In our more sublime and better moments, we have produced great art, rapturous music, monumental structures and undertaken courageous explorations. We have come to understand something of our origins, of the stars and planets, and, of the micro-bits of energy which comprise being.

If it is not an autonomous system to which we can turn in this time of need, then we must consider a better and improved version of ourselves. Human integration with nanotechnology is a concept widely discussed. One of the predictions made by famed futurist, Ray Kurzweil<sup>4</sup>, is that robotics and biotechnology will soon allow manufactured silicon placements to be integrated with the human brain. Kurzweil claims this technology will so advance our intellectual capabilities that it will dynamically alter what it is to be mortal men and women because of its implications for the escalation of human knowledge in genetics and all other branches of human science and learning.

What there is to gain is the continuation of our species with greatly enhanced intellectual functioning, a predisposition for the rational, abandonment of logic antithetical to reason and a physiology gifted by advanced eugenic science.

There are the dangers, of course, with such an enhancement of man; but, similar risks and dangers have been with us for eons. There has always been the danger that if one holds too much power, then the power will be misused and employed to exploit others. There have always been perils associated with one nation state possessing superior military strength. These dangers cannot be discounted, but, if the science of interactive decision making teaches us anything, it is that the quality of information processing minimizes risk. Super intelligence cannot fail to serve the ends of security among competing factions with similar technologies.

Such a future also raises the sobering possibility that our species may diverge on itself and birth an altogether different species and separate from homo sapiens. If that occurs, however, it would not be substantially different than that which natural selection has been accomplishing for millennia.

A few years ago, Norman Mailer published "A Gospel According to the Son" (1997). In this short work, Mailer created some interesting dialog between Christ and Pilate. During that discourse, Pilate asks "What is truth?" and when Jesus fails to answer,

---

<sup>4</sup> Raymond "Ray" Kurzweil is an American author, and futurist, who has written extensively on artificial intelligence and its implications for the human species. Among his works on these topics are [The Age of Intelligent Machines](#) (1990) and [The Age of Spiritual Machines](#) (1999) in which Kurzweil describes the history, development and future of AI technology in generally optimistic terms.

Pilate responds to his own question by observing: "Where there is truth, there will be no peace. Where peace abides, you will find no truth." Regardless of how one considers Mailer's theology, Pilate's statement may contain an elegant summary of what compels the human spirit. Excellence is a function of challenge, peril and risk.

While the implementation of AI is revolutionary and fraught with aspects requiring great caution, so have all the great leaps made by our ancestors. From before homo habilis, our nature has always compelled us forward. Regardless, for better or worse, AI is a technology that is being researched extensively in many parts of the globe today. There is little question that it will be brought to fruit. Our challenge is in determining how this transformative technology will be harvested for our greater good.

#### IV

The most beautiful thing we can experience is the mysterious. It is the source of all true art and all science. Albert Einstein.

---

Two things fill the mind with ever-increasing wonder and awe, the more often and the more intensely the mind of thought is drawn to them: the starry heavens above me and the moral law within me. Immanuel Kant, *Critique of Practical Reason*

---

When old age shall this generation waste,  
Thou shalt remain, in midst of other woe  
Than ours, a friend to man, to whom thou say'st,  
'Beauty is truth, truth beauty,—that is all  
Ye know on earth, and all ye need to know. *Ode on a Grecian Urn*, John Keats  
(1819)

It is no easy task to describe a moral imperative for a transcendent being. Because of our tendency to graft a divine imperative onto human conduct, morality has, in the past, found itself associated with the claims of religious teaching. However, religious precepts are closed to self-correction, making them effete and ineffectual. In the worst of times, they have supplied the basis for persecutions and excoriation of perceived heretics, and, have waged wars in the exercise of their dictatorial powers. At some level, all religions practice their own predilections for bigotry and intolerance and, at various stages of history, the Christian and Islamic faiths have shown strong propensities for violence. In more recent times, fundamentalist religions in the West have allowed the cloying elements of commercialism to seduce them. In third world countries, where there is easy prey for conversions, zealots collaborate with the very squalor and

infirmities they profess to alleviate<sup>5</sup>. Religious doctrines are also rife with ritualistic practices and sets of beliefs which no longer serve any useful purpose.

Perhaps, at one time religion served the function of reifying the source of the wonder invoked by nature, by supplying the inspiration for our great works, and, even preserving, at least some of the classical works of antiquity when Europe entered the dark ages; but, it was always at a price, and, now is no longer necessary. More importantly, the crisis now at the millennial lintel is refractory to personalized deities, rituals and other imaginary means of salvation. The wisdom and basis for our moral course must come from other than what humanity has traditionally invoked.

Science offers something much different. The claims of science eschew rigidity and intractability. Science favors honest observation and is dispassionate in its quest for truth. It favors parsimony and restraint in reaching its conclusions. Its methodology is practiced with patience and with deference to established principles, yet, subjecting them to constant scrutiny, correction and refinement. These are characteristics not entirely dissimilar to principles found in the ethical codes of the great civilizations of antiquity<sup>6</sup> to which much of modern rationalism is indebted.

There is a structure in the manner in which we should seek the truth. That structure possesses a rational quality which is, at times, also visible in what great civilizations have touted as virtues. We admire and inherently recognize virtue, and, even have codified it on occasion; nonetheless, we have had much difficulty in practicing it. Our transcendence, however, promises to achieve the social justice which has always escaped our cultures in the past. Without the impairments which ignorance and logical

---

<sup>5</sup> Referring to Mother Theresa, British journalist, Christopher Hitchens, wrote that “[She] was not a friend of the poor. She was a friend of poverty. She said that suffering was a gift from God. She spent her life opposing the only known cure for poverty, which is the empowerment of women and the emancipation of them from a livestock version of compulsory reproduction.” “Mommie Dearest”, *Slate*, 20 October 2003

<sup>6</sup> By way of example, the Romans of antiquity conceptualized certain qualities which were believed to provide the Roman Republic with its moral basis for its dominance in the known world. These included *frugalitas* and *simplicitas* (which implied an economy and simplicity of style), *industria* and *gravitas* (suggesting hard work and a seriousness of purpose), *severitas* (restraint and self-control), *honestas* (a status arising from a freedom from outside influences). The Romans also recognized *pietas* as a virtue. Its meaning corresponds to a sense of obligation or duty arising out of one’s relationships to ones institutions, family, friends and others.



inconsistency imprint upon our institutions, we will be empowered to divest ourselves of bigotry, fear, and, of the wreckage of distorted reasoning. We, as well as our ancestors, describe virtues in our philosophies, we place platitudes on our monuments, and, profess to honor ideals in our memorials. There is something noble to be human, but, our shortcomings have, for so long, left us deficient in obtaining the fullness of our ideals.

If there is a universal imperative for morality, Keats may have had it right when he wrote those famous lines in his ode, "beauty is truth, truth beauty". The only proper ethos for the coming age arises from a union with the rational. The ordinances which hold the cosmos together are ideations of marvel. They embody "truth" because of their profound correspondence to reality, and, "beauty" because of their symmetry and refinement. These prescriptions attest to an order, dignity, and majesty in the universe. They reveal secrets about our natures and of our relationships with the cosmos and one another. These revelations also carry with them certain deontological aspects, imposing duties and obligations, which will become easily and effortlessly discernable with the untrammled vision of transcendence.

Enhanced intellectual capabilities carry the implications of what science can bestow on the human physiology and disposition. Free of errors of logic, personal dysfunction, and, unburdened of ignorance, we will find ourselves in union with the pervasive rational forces which fill the universe; the new ethos will be a "thing of beauty", tending toward an ideal, and, by definition, for the greater good of mankind, and, of the unchained Promethean giants who will come to know the cosmos.

JP  
April 17, 2014

---

+++ The title reference is the lyrical drama by Percy Bysshe Shelley (1820) featuring Prometheus from Greek mythology. Prometheus is condemned to eternal damnation as a result of disobeying the gods by giving fire to humanity. In Shelly's work, Prometheus manages to escape and to prevail in an ultimate triumph over the gods.