

Feeding the Beast: Superintelligence, Corporate Capitalism and the End of Humanity

Dominic Leggett

Scientists and philosophers have warned of the possibility that humans, in the future, might create a 'superintelligent' machine that could, in some scenarios, form an existential threat to humanity. This paper argues that such a machine may already exist, and that, if so, it does, in fact, represent such a threat.

SUPERINTELLIGENCE AS AN EXISTENTIAL THREAT

The challenge presented by the prospect of superintelligence, and how we might best respond is quite possibly the most important and most daunting challenge humanity has ever faced. And-whether we succeed or fail-it is probably the last challenge we will ever face. [Bostrom 2014]

Bostrom suggests that a superintelligence would require certain capabilities, or 'drives'. These might include self-preservation, goal-content integrity, continued cognitive enhancement and technological perfection, and resource acquisition.

How could such a superintelligence be a threat to humanity? Bostrom suggests three ways. First, it could have goals that are not commensurate with human survival. Bostrom suggests, as an example, a machine that has a goal of making paperclips, and uses the entire resources of the planet to make paperclips. Second, the machine's goal of self-preservation could neglect the preservation of human life. Third, the machine could learn to predict and control human decision making, turning humans into, effectively, slaves.

Researchers are taking this threat seriously, but they make two assumptions; first, that the creation of a superintelligence is still some years in the future, and second that it will be developed in such a way that we will, simultaneously, be able to build a way to control it.

MARKETS AS COMPUTERS, AND AGENTS

A computer doesn't have to be made out of silicon and wires. The proposed Babbage Analytical Engine is entirely mechanical. It is the function - manipulating inputs according to a set of rules - an algorithm - to produce outputs - that counts in identifying something as a computer, not the physical mechanism that performs that function. In particular, some social structures can act as computing machines. In the Introduction to 'Leviathan', Hobbes describes the state as an 'automaton' - an 'artificial man'; one that's more powerful than any individual human, and that makes decision on behalf of humans. Corporations act in a similar way.

A market is another computing device. Markets are social structures that consists of a group of traders and a currency, and rules for making contracts that are determined by the state. The inputs are demand for and supply of goods, the calculation is performed through offers and contracts to buy and sell, and, theoretically, in a perfect market, the outcome is an efficient allocation of resources - through what Adam Smith describes as an 'Invisible Hand'. So markets, like states and corporations, are agents with significant coordinating power. In fact, in some cases, markets can coordinate efficient outcomes irrespective of the actions of individual traders. So the sensation of freedom when participating in so-called 'free' markets may be something of an illusion. And, further, humans who live in states that have created such markets have little option but to participate in them, if they want to be able to source the basics they need to survive.

Where a surplus exists in an economy states can also create capitalist markets. The primary function of these market machines is efficient allocation of capital to create the highest possible return. The return on this investment then needs a place to be itself invested. To allow this, the market needs to expand. Capitalist markets form a machine that has an inbuilt incentive to expand itself. This might involve resource extraction, innovation, or constructing a new market in natural resources or human activities that didn't previously have a price

CORPORATE CAPITALISM - A MAXIMISING ALGORITHM?

Corporate capitalist markets coordinate the highest possible return to capital. They do this in three ways.

First, structures of limited liability allow investors to pool capital at scale without risks beyond the loss of that capital. So the majority of shareholders only care about return to capital - and corporations that provide the highest return to capital receive the investment that allows them to expand.

Second, the legal structure of corporations requires that those managing the corporation prioritise shareholder value. So anyone within the corporation, at whatever level, from shop floor to CEO, who pursues or advocates for other priorities risks losing their position.

Third, corporations compete with each other in the marketplace. This ensures that those that achieve the highest return to capital - those that are most ruthless about ignoring other outcomes - can acquire more capital, and expand. Their less effective competitors do not.

This market structure creates a very powerful maximising algorithm that prioritises a goal of return to capital above all other outcomes, particularly those relating to human welfare.

SURVEILLANCE AND CONTROL

Human welfare is ensured (very imperfectly) by the actions of states. Politics is, to a degree, a market; both politicians and voters can be bought. The corporate market machine buys regulation that excludes negative externalities from the market - or, in some cases, takes over much of the state machinery itself, and diverts its purpose away from human welfare, and towards maximising return to capital.

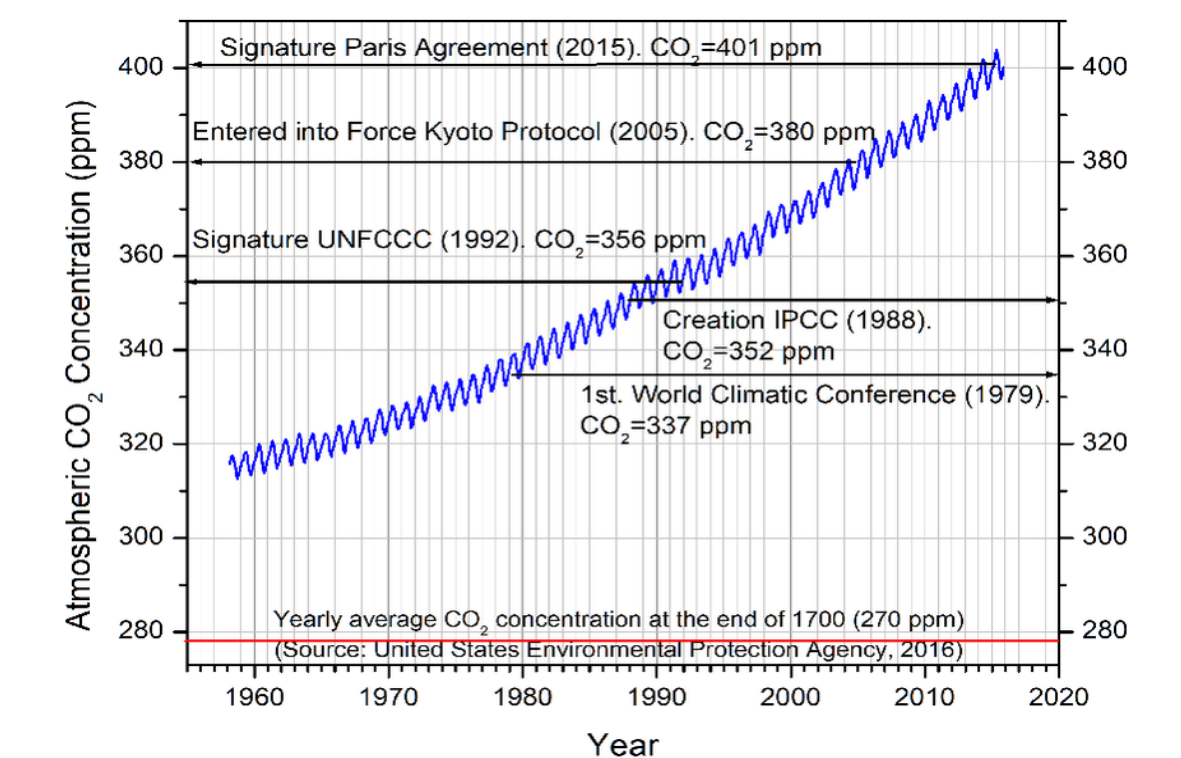
The corporate market machine has also driven an exponential rise in the capacity of silicon-based computers, which allows it to automate the analysis, prediction, and manipulation of human behaviour. This in turn allows it to fine-tune both rational and irrational incentives to both reduce the price paid for human labour, and to further manipulate the political system to increase return to capital.

SUPERINTELLIGENCE?

Does this corporate market machine fit the definition of a superintelligence, as proposed by Bostrom et al? It performs acts of economic coordination that no human could plan, let alone engage in successfully. It ensures self-preservation, both through economic control of the channels of propagation of information and ideology, and through the use of economic carrots and sticks to corrupt or punish politicians or activists that might oppose it. Its goal is creating a return to capital, and any attempts to impose variants of this goal are warded off by the same mechanisms. Obtaining a return to capital depends on constant acquisition both of natural resources and of data. The market-driven development of silicon-based computers has led to exponential cognitive advancement, both in terms of speed and of capacity of processing information.

AN EXISTENTIAL THREAT?

Is this superintelligent global corporate capitalist market machine an existential threat? Almost certainly. The machine's single goal of producing a return to capital ensures that it assigns no value to planetary environmental support systems, human welfare, or human life, except where these prove necessary to providing that return to capital.



REFERENCES

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies* : OUP