All Watched over by Machines of Silent Grace?

Mark Bishop

Published online: 11 August 2011
© Springer-Verlag 2011

August 1, 2011

When discussing the place of robotic system in civilian and military society, typically two opposing key themes will emerge: dreams of robot heaven and nightmares of robot hell. And it was in the 1960s, in this dual context—the robot as facilitator, the robot as destroyer—that the poet Richard Brautigan highlighted a vision of humans and robots working together to build a new society, a utopian cybernetic world ‘all watched over by machines of loving grace’ (Brautigan 1967).

In the intervening four decades since Brautigan's poem was first published, a number of core technologies required to underpin such cybernetic utopias have been successfully developed, even if the concomitant social changes he promised have yet to materialise. In the civilian world, current examples of the forecast army of 12 million plus¹ robot workers include: the ubiquitous yellow ‘Fanuc’ industrial robots; the ‘i-Robot’ vacuum cleaner, the ‘Robomow’ lawnmower; the ‘Cogniron’ home companion; ‘Forex’ bot financial traders, etc.²

Similar new robotic technologies have equally permeated the military world, perhaps most famously in the guise of the cruise missile. Typically, this is a jet-powered robot missile, used to remotely project superpower (with minimal risk to superpower personnel) as it autonomously carries a [conventional or nuclear] explosive payload towards a land or sea-based target. Once it has been targeted and launched by its support crew and systems, the cruise missile will autonomously plot a course enabling it to

---

¹The IFR Statistical Department forecast deployment of approximately 11 million service and leisure robots and 1 million industrial robots worldwide. Source: World Robotics Report 2010 (<http://www.worldrobotics.org/>).

²Background information on robots referenced in this article sourced from relevant manufactures' and Wikipedia web pages.

Founded in 1964, the AISB was the world's first society for the study of Artificial Intelligence and the Simulation of Behaviour.

M. Bishop (✉)
Goldsmiths, University of London, New Cross, London SE14 6NW, UK
e-mail: m.bishop@gold.ac.uk

appropriately engage the enemy target. In contrast to unmanned aerial vehicles (UAVs), cruise missiles are only used as weapons and are never deployed for reconnaissance.

Conversely, unmanned aerial [drone] aircraft are controlled from remote video consoles. These devices are increasingly being used by the US military in Afghanistan and by the CIA in Pakistan and other places outside of traditionally recognized war zones. The widespread use of such drones, alongside inevitable civilian deaths, raises both technical concerns about their effectiveness and philosophical questions about their ethics; for example, who is responsible for any civilian deaths? A government that is complicit in their deployment on its territory? The remote pilots who guide the UAVs? The aerospace firm that designed the control software?, etc. Worryingly, as Noel Sharkey highlighted in a 2008 'Science Perspective' article, "... despite [these] potential problems, no international or national legislation or policy guidelines exist except in terms of negligence" (Sharkey 2008).

More recent developments in state-of-the-art military robotics include: the now famous Boston Dynamics 'Big Dog' robotic mule (a rough terrain robot designed to carry heavy loads) whose 2008 video has currently racked up over 11 million YouTube hits; the Secom 'robot guard', a six-wheeled surveillance robot (likened to a cross between a dodgem car and R2D2) developed to scare off intruders, which releases a dense, billowing cloud of smoke as it films and chases its prey; and the Samsung SGR-A1. This is the sentry robot which, in 2007, due to the diminishing numbers of troops available to patrol its demilitarised zone (DMZ), South Korea announced it planned to install along its border with North Korea (Kumagai 2007). An ominous portent of things to come perhaps, the SGR-A1 has been specifically engineered to either wait for the order to fire from its human controllers or shoot 'at will'.

It's technological developments such as these that have led futurologists such as Hugo de Garis,³ Kevin Warwick (Warwick 1998) and Ray Kurzweil (Kurzweil 1999) to issue dark prophecies warning of fast-emerging robot hell; a future dystopian age of widespread human subjugation. The scenarios of techno-destruction outlined by these futurologists are usually based on the notion of the 'rogue' robot: the machine (or machines) which eventually turn rabidly against their human masters. However, recent history demonstrates that future techno-Armageddon does not have to be willful or intentional—as typically portrayed in films like 'The Terminator'—but can also occur naturally as a consequence of ill-considered pre-programmed 'mechanical' actions.

Thus, it was on September 6, 1983, during operation 'Able Archer', that an automatic military surveillance system—a 'monitoring agent' to borrow Haag's 2006 taxonomy of software agents (Haag 2006)—almost led to World War III (Lebedev 2004). During the height of what Russia perceived to be an intimidating US military exercise in central Europe, a malfunctioning Soviet alarm system alerted a Soviet colonel that the USSR was apparently under attack by multiple US ballistic missiles. Fortunately, the colonel had a hunch that his alarm system was malfunctioning, and reported it as such. Many believe that the colonel's quick and correct decision on how to respond may have averted east–west nuclear Armageddon.

³ Hugo de Garis, *The Artilect War: Cosmists vs. Terrans: A Bitter Controversy Concerning Whether Humanity Should Build Godlike Massively Intelligent Machines*, ETC Publications (2005).

Similarly, albeit a little less explosively, a future robot-instigated, global economic Armageddon is a possibility. Given the worldwide fear of economic melt-down that resulted from the 2008 Lehman Brothers collapse, it is perhaps not so farfetched to imagine potential malfunctioning auto-traders dragging the financial world into financial and hence political ruin; indeed an algorithmic agent-based system was implicated in the May 6, 2010 Flash Crash (Lauricella 2010)⁴ when the Dow Jones Industrial Average plunged about 600 points. At that time, it was the second largest point swing (1,010.14 points) and the biggest 1-day decline (998.5 points) in Dow Jones history.

Such examples demonstrate that technological route maps to drive robots a long way towards the key themes of robot heaven and hell already exist. However, contra Warwick, de Garis, Kurzweil and the prophets of dystopian robotic-singularity; or Brautigan and Moravec's post-human robot utopias (Moravec 2000)—it is my opinion that for both key themes, a huge conceptual and technological wall remains to be climbed before any future robot heaven or robot hell can ever usher forth. That wall is a giant edifice built on autonomy and teleology.

In a very strong sense, I suggest that systems lacking genuine autonomy will fundamentally always remain mere 'servants' of their human masters and similarly systems lacking teleology will fundamentally always be directed by the human gaze. These two notions not only delimit the possible space of human/robot interactions but also shape the field of the [artificial] intelligence that grounds it. All [Turing machine-powered] robotics can hope to achieve is, to paraphrase Dennett, a weak form of 'as-if' autonomy and 'as-if' teleology, which in reality, merely reflects their engineers' designs and their end-users' wishes. Such devices—although more than capable of serving a future James Bond-style villain's dreams of world domination—remain incapable of dreaming these dreams for themselves (Bishop and Nasuto 2005).

Furthermore, as has been remarked elsewhere, how can a robot seek to engage in the human world without experience of genuine phenomenal states? (Bishop 2009); how can a robot possibly ever be said to genuinely understand anything of the world without a genuine intentional stance in the world? (Searle 1980) It was such observations, together with his own insight into [the lack of] mechanised understanding, that led Roger Penrose to reassuringly, if controversially, conclude that "computers...would always remain subservient to us, no matter how far they advance with respect to speed, capacity and logical design" (Penrose 1994). And powered by mere 'silent' computers, how can the robot ever experience any kind of lived, transcendent grace-loving or otherwise?

References

- Bishop, M. (2009). Why computers can't feel pain. *Mind & Machine*, 19(4), 507–516. Kluwer.
 Bishop, J. M., & Nasuto, S. J. (2005). Second order cybernetics and enactive perception. *Kybernetes*, 34 (9/10), 1309–1320. Emerald.
 Brautigan, R. (1967). *All watched over by machines of loving grace*. communication company USA.

⁴ Graham Bowley, Lone \$4.1 Billion Sale Led to Flash Crash in May. The New York Times (October 1, 2010).

- Haag, S. (2006). *Management information systems for the information age*, 224–228. New York: McGraw-Hill Publishing Company.
- Kumagai, J. (2007). A robotic sentry for Korea's demilitarised zone. *IEEE Spectrum*, IEEE.
- Kurzweil, R. (1999). The age of spiritual machines. *Viking*.
- Lauricella, T. (2010). How a trading algorithm went awry. *The Wall Street Journal*.
- Lebedev, A. (2004). 'The man who saved the world finally recognized'. *MosNews*.
- Moravec, H. (2000). *Robot: Mere Machine to Transcendent Mind*. OUP.
- Penrose, R. (1994). *Shadows of the mind*, 35. OUP.
- Searle, J. (1980). Minds, brains and programs. *The Behavioral and Brain Sciences*, 3, 417–424. CUP.
- Sharkey, N. (2008). *Science*, 322(5909), 1800–1801. AAAS.
- Warwick, K. (1998). *In the mind of the machine*. London: Arrow Books.