Automating warfare is ethically dangerous

Noel Sharkey outlines the disturbing trends in military robotics.

War-fighting is currently undergoing a factory-style revolution. The use of robotics platforms for carrying weapons is coming on track at an increasing rate. US plans from all of the armed forces indicate a massive build up of military robots, and I have personally tracked and validated 51 countries that have either bought them or have development programmes. Currently all armed robots in war are remotely piloted by humans, so called 'man-in-the-loop' systems. Humans are responsible for both target selection and decisions about lethal force. But this is all set to change.

Since 2004, all of the roadmaps and plans of the US forces have been pushing the development and deployment of autonomous battlefield robots.[1] The UK Ministry of Defence Joint Doctrine Note,[2] published in 2011, follows suit. Fulfilment of these plans is well underway. There will be a staged progression towards autonomous operation; first for flight (take-off, navigation, obstacle avoidance, etc.) then for target selection. The end goal is robots that operate autonomously to locate their own targets and kill them without human intervention.[3]

It is said, perhaps too often, that for now a person will remain somewhere in the loop. But their role will shrink to vanishingly small: "humans will no longer be '*in* the loop' but rather '*on* the loop' – monitoring the execution of certain decisions. Simultaneously, advances in AI [artificial intelligence] will enable systems to make combat decisions and act within legal and policy constraints without necessarily requiring human input."[4] So, essentially a person will be on the loop to send in the autonomous swarm and possibly call it off again if radio or satellite contact is available.

Autonomous systems that can select targets and kill them are likely to face a number of ethical and legal problems.[5,6,7] In brief, no autonomous robots or artificial intelligence systems can discriminate between combatants and innocents.[8] International humanitarian law and the laws of war state clearly that belligerents may not attack civilians, wounded soldiers, the sick, the mentally ill, or captives. There are no visual or sensing systems for robots that are up to that challenge. Current sensing apparatus and processing can just about tell if something resembles a human but little else.

Armed drone on patrol: MQ-9 Reaper

A computer can understand any procedure that can be written in a programming language. We could, for example, give a robot a programmed instruction such as, "if civilian, do not shoot". But there is no precise definition of a civilian. We certainly cannot get one from the laws of war. The Geneva Convention requires soldiers to use common sense. But computers have no common sense.

Even if there was a precise computational definition of civilian, and robots were equipped with the appropriate sensing apparatus to discriminate, it is not appropriate to kill enemy combatants in all circumstances. Both discrimination and appropriateness require the kinds of real-world reasoning that AI systems are notoriously bad at.

Another problem is the 'principle of proportionality', which requires balancing the risks of civilian death with the military advantage to be gained. Again, there is no computational reasoning that would allow a robot such a determination, nor is there any known metric to measure objectively needless, superfluous or disproportionate suffering.[9] It requires human judgement. And yes, humans do make errors and can behave unethically, but they can be held accountable. Who is responsible for the lethal mishaps of a robot? Certainly not the machine. There is a long causal chain associated with robot mishaps: the manufacturer; the programmer; the designer; the department of defence; the generals or admirals in charge of the operation; the robot operator; or the enemy.

Before automating war, lessons learned from the current use of remotely piloted armed robots need to be considered.[10] There is an illusion of accuracy that is leading to the inappropriate expansion of the 'battle space' where many civilians are dying and there are frequent illegal targeted killings by the CIA that allow no chance for surrender or trial. Even worse is the adoption of the technology by so many countries without any kind of international discussion about rules of use, or how the many complex algorithms will interact.

Action

You can help advocate international control of robotic armaments by adding your voice of dissent at: www.icrac.co.uk

Noel Sharkey is Professor of artificial intelligence and robotics and Professor of public engagement at the University of Sheffield. Twitter: StopTheRobotWar

References

1. Headquarters of the United States Air Force (2009). United States Air Force Unmanned Aircraft Systems Flight Plan 2009-2047. Washington DC. 18 May.
2. Ministry of Defence (2011). The UK Approach to Unmanned Aircraft Systems. Joint Doctrine Note 2/11. 30 March.
3. Sharkey N (2008). Cassandra or the False Prophet of Doom: AI Robots and War. IEEE Intelligent Systems, 23(4), pp.14-17.
4. Ibid note 1, p.41
5. Sharkey N (2007). Automated killers and computer professionals. IEEE Computer, 40, pp.106-108.
6. Sharkey N (2008). The Ethical Frontiers of Robotics. Science, 322, pp. 1800-1801.
7. Sharkey N (2010). Saying "No!" to lethal autonomous targeting. Journal of Military Ethics, 9, pp.369-383.
8. Sharkey N (2008). Grounds for Discrimination: Autonomous Robot Weapons. RUSI Defence Systems, 11(2), pp.86-89.
9. Sharkey N (2009). Death Strikes from the Sky: The calculus of proportionality. IEEE Science and Society, 28, pp.16-19
10. Sharkey N (in press). Automating Warfare: lesson learned from the drones. Journal of Law and Information Science.

6