"Firing, Forgetting and How Rule-of-Law-Values Require Automation of the Rule of Law; A Defense of the Use of Autonomous Weapons Systems in War and Peace"

Duncan MacIntosh,
Dalhousie University

Extended Abstract:

This paper defends not only of the use of Autonomous Weapons Systems (AWSs), but the automation and robotization of the conduct of war and even of the formulation, administration and enforcement of law itself.

      While AWSs have obvious military advantages, there are several prima facie moral objections to using them. In Part I of this essay, by way of meeting these objections I first point out the similarities between the structure of law and morality on the one hand and of automata on the other: both morality/law and automata are rule-driven, both human moral/legal agents and robotic agents should have their purview restricted just to the tasks they are competent to carry out, and both human and robotic agents are such that, in the roles they are asked to play, they needn't be full Moral Turing Machines (i.e., beings able to pass as persons morally wise and morally compliant in all possible situations), they need only contribute to the systems they partly constitute being Moral Turing Machine competent, and even then, only for that part of morality of which their respective system is a part (c.f. Johnson and Noorman). I then suggest that, far from being a threat to morality and the rule of law, the truth is that the use of autonomous weapons systems and the automation of the administration and enforcement of law, both civilian, military and international, are the logical extension of the rule of law, the fullest fruition and expression of the rule of law (rather than rule by a man); and we are therefore obliged by the principles of rule of law – and by their grounding in moral obligation – to automate administration and enforcement of law.

      Having sketched that general program, in Part II I argue more specifically (and contra Mary Ellen O'Connell, Christof Heyns, et al) that there are many conditions under which use of AWSs would be strategically, morally and rationally appropriate, the propriety of their use varying with the competences of which they are capable at their then prevailing level of design, and from situation to situation, some situations calling for

very smart automata, some actually requiring automata precisely for their being "stupid". I shall be defending the use of AWSs conceived as a fire and forget weapons – weapons that one activates and then does not necessarily monitor with a view to possibly recalling them before the end of their missions; and I claim their use can be justified in the following situations: situations where the use of AWSs is justified by factors concerning the nature of rational planning (c.f. Michael Bratman), or by the longer term morally good consequences of actions morally problematic and therefore recall-tempting in the short term (c.f. act utilitarian justifications), or by virtue of the value of rational pre-commitment and rational resolute choice (c.f. David Gauthier), or by one's having no weapon available except one that cannot be recalled after activation, or having only a weapon whose nature is that, if one made it able to be recalled, it would be at great risk of being hacked for evil purpose (c.f. other readings for this conference); or by a weapon's autonomy having the advantage of protecting the moral conscience of its user, or the weapon's providing morally required diffusion of responsibility, or its providing a morally required randomness and non-deliberateness to actions; or maybe we are trying to train a robot into making moral distinctions, and the press of events means we must do this in the field, leaving the recall decision to the robot itself. Or using the weapon could be justified by its providing greater precision in respecting the distinction between those morally liable and morally not liable to being killed -- something that would be put at risk by the reconsideration of a clumsy human operator (c.f. Ronald Arkin, et al) -- or by its affording a greater dignity to some of the parties precisely by virtue of its isolation from human control, or by virtue of its being vastly more efficient in a way that, again, would be jeopardized by less efficient human intervention (again, Arkin, et al), or by its operation being more in the spirit of the rule of law rather than the rule of a man or of the rule of humans generally, given their susceptibility to moral failings, failings of specious partiality, impulse, exhaustion, weakness of will, and emotionality, and so required by the very commitment to the rule of law. Or (contra Gabriella Blum) use of the weapon could be required by virtue of its inherent comparative effectiveness against unjust, non-democratic regimes, these able to be ended by killing a dictator, contrasted to its comparative harmlessness to democracies, due to democracies using a process to produce leaders that forms them by forming a national will, which would then be able endlessly to resupply leaders if any given leader were killed. This last is also an argument for encouraging the general proliferation of automata, not just in the military world, but also in the civilian world. Next, using an AWS could be justified compared to conventional warfare tools by being vastly cheaper, more efficient, less casualty-involving of friendly troops and more easily re-located so that it can be deployed in situations where we have a duty to intervene, but less of a national interest in intervening, e.g., in Africa rather than the oil-rich Middle East. This machine would empower us to do our moral duty in a way sustainable pro bono, without expectation of immediate business investment return. I suspect that under all of these conditions use of the weapons would probably be legally

appropriate as the laws, national and international, currently stand, especially in light of observations of people like Jens David Ohlin (paper for this conference) about the possibility of holding the deployers of AWSs legally responsible for their behaviour; but if not, the laws should be changed to permit use of these weapons.

In Part III I deal with objections, e.g., that use of robots is inherently wrong or violating of human dignity (claims to which I offer counter-examples), or will make violence too easy in either of two senses, namely, tempting us too easily into violence in resolving conflicts, or giving us unfair advantage in conflicts. To these objections I say better to have the weapons, use them with the courage of our moral convictions, and try with education, law and oversight to control possible abuses of the weapons. I then argue that the possibility of the use of robots in law generally and in war specifically are a signal of the coming of a great divide between the time of a world at war and the time to come of a world attaining consensus, and left only with the problem of how to police externalities to a mutually rationally acceptable set of total arrangements representing the best deal for everyone. In this world, robots would in effect be like the spring-loaded guns protecting the money in a safe, except that the safe would contain all that is sane and good in the world, while the targets would be those so deranged as to be unable to be welcomed into the deal, and robots would reduce the amount of violence involved in the residual conflicts by orders of magnitude.

I will conclude, in Part IV, with some second thoughts. I consider two conceptions of rule of law. One conception sees it as the unbiased and unexceptioned literal application of unambiguous rules. The other conception sees it as the determining of what in each situation that people face should be done, whether what was done in a given situation was right, and, if not, what should be done in consequence. And it sees all this as occurring by argumentation based on interpreting inherently ambiguous rules and negotiating their inherently contestable applicability in given cases. The first conception deploys the metaphor of law as rules of a machine, while the second conceives law as ongoing debate about behaviour-regulating norms that is conducted in terms of discussion about human-chosen principles (rather than character traits, best outcomes, divine commands, or the impulses of a powerful person), where the behavior to be regulated could probably instead have been improvised without need of prior meditation upon principles – invoking laws is just a handy conversational trope. Obviously the first conception of law seems more fully automatable. But even on the latter conception there may be aspects of law that could be robotized (e.g., parking law enforcement); and even on this conception, a robot advanced enough to deliberate about and justify its actions just as a person can might be able to contribute. Although at that point the robot/person distinction becomes merely one of physical composition, not conceptual difference.

I consider also two different conceptions of what a machine might be asked to do, first, apply rules that serve goals, second, choose goals qua ends, things to be conceived as valuable in themselves and as objects to be pursued by moral agents. An open question

on all counts is what machines can contribute to law if they don't themselves have interests situating them in what Immanuel Kant called the Kingdom of Ends as interest holders, goal choosers and as agents negotiating what laws morally must be followed in the course of pursuing these ends. If robots don't have and can't adopt ends, they are merely human-created compilers of records of human norms, checkers of the consistency of human-created laws, and administrators and enforcers of these laws.