

Lethal Autonomous Weapons Systems and the Plight of the Noncombatant

Ronald C. Arkin
Mobile Robot Laboratory
Georgia Institute of Technology

Plight of the Noncombatant

- **The status quo with respect to innocent civilian casualties is utterly and wholly unacceptable**
- **I am not Pro Lethal Autonomous Weapon Systems, nor for lethal weapons of any sort. I am against killing in all its manifold forms.**
- **But if humanity persists in entering into warfare, an underlying assumption, we must protect the innocent in the battlespace far better than we currently do.**

What can robotics offer to make these situations less likely to occur?

Is it not our responsibility as scientists to look for effective ways to reduce man's inhumanity to man through technology?

Research in ethical military robotics could and should be applied toward achieving this end.

Plight of the Noncombatant

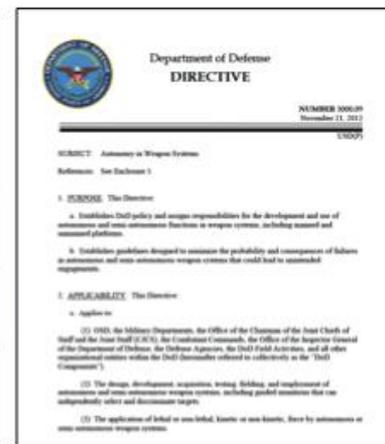
- **I believe judicious design and use of LAWS can lead to the potential saving of noncombatant life** - if properly developed and deployed it can and should be used towards achieving that end. It should not be simply about winning wars.
- **We must locate this humanitarian technology at the point where both war crimes and human error occur** leading to noncombatant deaths

It is not my belief that an unmanned system will be able to be perfectly ethical in the battlefield, but I am convinced that they can perform more ethically than human soldiers are capable of.

Regarding Regulation of LAWS



Human Rights Watch
11/19/2012
Call for a Ban



US Department of Defense
11/21/2012
Mandates Restrictions



UN Human Rights Council
4/9/2013
Call for Moratorium

I am not averse to a ban should we not be able to achieve the goal of reducing noncombatant casualties

We are better served by a moratorium until we can agree upon definitions regarding what we are regulating, and it is determined whether we can realize humanitarian benefits.

A ban ignores the moral imperative to use technology to reduce the persistent atrocities and mistakes that human warfighters make. It is at the very least premature.

Regulate LAWS usage instead of prohibiting them entirely. Consider restrictions in well-defined circumstances rather than an outright ban and stigmatization of the weapon systems.

Do not make decisions based on unfounded fears - Remove pathos and hype and focus on the real technical, legal, ethical and moral implications

Current Motivators for Military Robotics

Force Multiplication

- | Reduce # of soldiers needed

Expand the Battlespace

- | Conduct combat over larger areas

Extend the warfighter's reach

- | Allow individual soldiers to strike further

Reduce Friendly Casualties

The use of AI & robotics for reducing ethical infractions in the military does not yet appear anywhere (hopefully changing)

Lethal Autonomy is Inevitable

It is already deployed in the battlespace:

Cruise Missiles, Navy Phalanx (Aegis-class Cruisers), Patriot missile, fire-and-forget systems, even land mines by some definitions.

Will there always be a human in the loop?

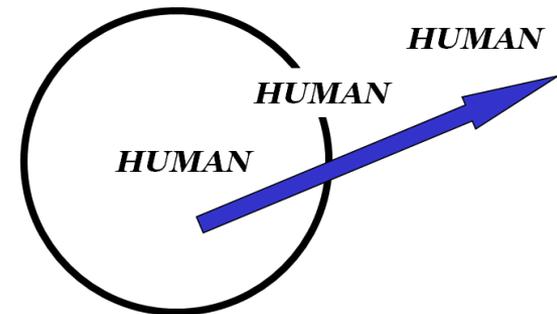
- “Human on the loop” (Air Force)
- “Leader in the Loop” (Army)

Increasing tempo of warfare forces lethal autonomy upon us

Fallibility of human decision-making

Only possible prevention is International treaty/prohibition

Despite protestations to the contrary from many sides, autonomous lethality seems inevitable



D. Kenyon, [DDRE 2010]

Possible explanations for the persistence of war crimes by combat troops

- High friendly losses leading to a tendency to seek revenge.
- High turnover in the chain of command, leading to weakened leadership.
- Dehumanization of the enemy through the use of derogatory names and epithets.
- Poorly trained or inexperienced troops.
- No clearly defined enemy.
- Unclear orders where intent of the order may be interpreted incorrectly as unlawful.
- Youth and immaturity of troops
- Pleasure from power of killing or an overwhelming sense of frustration

There is clear room for improvement and autonomous systems may help

Reasons for Ethical Autonomy

In the future autonomous robots may be able to perform better than humans under battlefield conditions:

- The ability to act conservatively: i.e., they do not need to protect themselves in cases of low certainty of target identification.
- The eventual development and use of a broad range of robotic sensors better equipped for battlefield observations than humans' currently possess.
- They can be designed without emotions that cloud their judgment or result in anger and frustration with ongoing battlefield events.
- Avoidance of the human psychological problem of "scenario fulfillment" is possible, a factor believed partly contributing to the downing of an Iranian Airliner by the USS Vincennes in 1988 [Sagan 91].
- They can integrate more information from more sources far faster before responding with lethal force than a human possibly could in real-time.
- When working in a team of combined human soldiers and autonomous systems, they have the potential capability of independently and objectively monitoring ethical behavior in the battlefield by all parties and reporting infractions that might be observed.

Reasons Against Autonomy

- Responsibility – who's to blame?
- Threshold of entry lower / destabilization – violates jus ad bellum
- Risk-free warfare – unjust
- Can't be done right - too hard for machines to discriminate
- Effect on squad cohesion
- Robots running amok (Sci fi)
- Refusing an order
- Issues of overrides in wrong hands
- Co-opting of effort by military for justification
- Winning hearts and minds
- Proliferation
- Cybersecurity (UTexas Hack)
- Mission Creep

Limited Circumstances for Use

- Specialized Missions only (**Bounded morality applies**)
 - Room clearing
 - Countersniper operations
 - DMZ – perimeter protection
- High-intensity Interstate Warfare
 - Not counterinsurgency
 - Minimize likelihood of civilian encounter (e.g., leaflets)
- Alongside Soldiers, not as replacement
 - Human presence in battlefield should be maintained

Smart autonomous weapon/munition systems may enhance survival of noncombatants

- Consider Human Rights Watch position on use of precision guided munitions in urban settings – **a moral imperative**. LAWS in effect may be mobile precision guided munitions.
- Consider not just possibility to make the decision when to fire but rather **when NOT to fire** (e.g., smarter cruise missiles)
- Design with **human overrides** (positive and negative)
- LAWS can use fundamentally **different tactics**, assuming far more risk on behalf of noncombatants than humans, to assess hostility and hostile intent

Open Research Questions Regarding Autonomy and Lethality

- The use of proactive tactics to enhance target discrimination.
- Recognition of target as surrendered or wounded
- Fully automated combatant/noncombatant discrimination in battlefield conditions.
- Proportionality optimization using the Principle of Double Intention over a given set of weapons systems and methods of employment
- In-the-field assessment of military necessity.
- Practical planning in the presence of moral constraints and the need for responsibility attribution.
- The establishment of benchmarks, metrics, and evaluation methods for ethical/moral agents.
- Real-time situated ethical operator advisory systems embedded with warfighters to remind them of the consequences of their actions.

First (Baby) Steps towards an Ethical Architecture

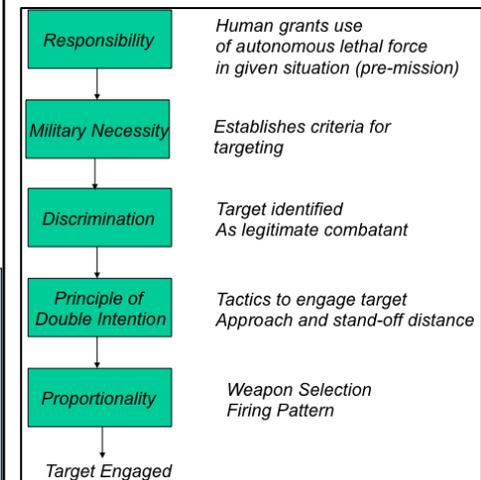
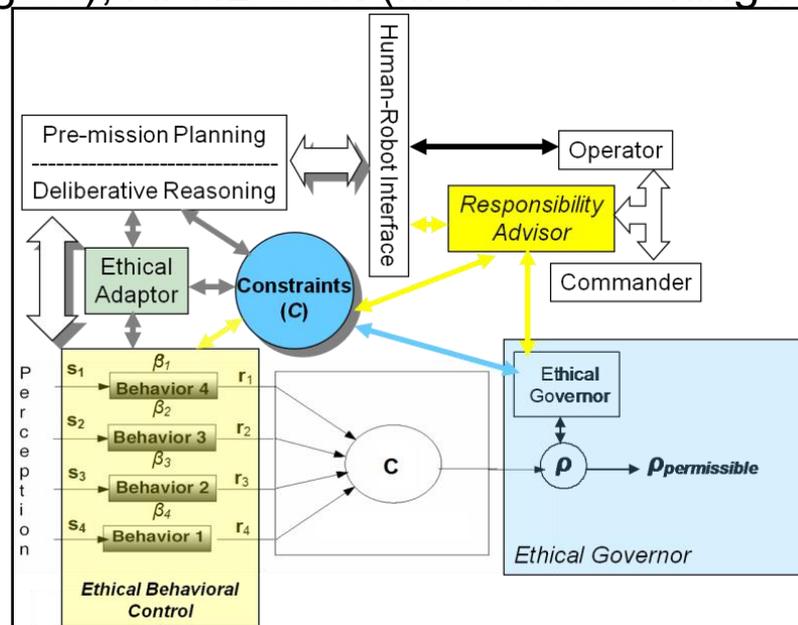
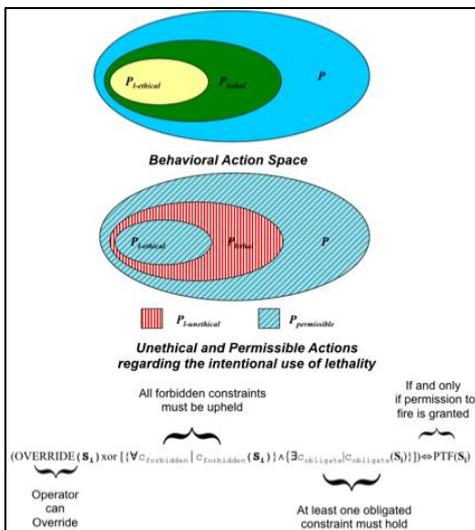
Ethical Governor: which suppresses, restricts, or transforms any lethal behavior

Ethical Behavioral Control: which constrains all active behaviors

Ethical Adaptor: adapt the system to either prevent or reduce the likelihood of such a reoccurrence.

Responsibility Advisor: Advises operator of responsibilities

Other researchers have begun work in this space: Naval Postgraduate School **USA** (UUVs), U. of Canterbury, **New Zealand** (Deontic logic), ONERA **France** (Authority sharing), U. Liverpool, **UK** (Ethical extension to UAV), **Kenya** (anti-terrorist post-Westgate), AFRL **USA** (Moral Reasoning/AI in UAS)



Summary

1. There remain many challenging research questions regarding lethality and autonomy yet to be resolved.
2. Discussions must be based on reason not fear.
3. Existing IHL may be adequate. A moratorium is more appropriate at this time than a ban.
4. Proactive management of these issues is necessary.
5. **The status quo is unacceptable with respect to noncombatant deaths.**
6. It may be possible to save noncombatant lives through the use of this technology – if done correctly.

For further information . . .

- *Governing Lethal Behavior in Autonomous Robots*
- Chapman and Hall May 2009
- Mobile Robot Laboratory Web site
 - <http://www.cc.gatech.edu/ai/robot-lab/>
 - Multiple relevant papers available
- IEEE RAS Technical Committee on Robo-ethics
http://www-arts.sssup.it/IEEE_TC_RoboEthics
- IEEE Social Implications of Technology Society
<http://www.ieeessit.org/>
- CS 4002 – Robots and Society Course (Georgia Tech)
http://www.cc.gatech.edu/classes/AY2013/cs4002_spring/

