Ethical Restraint of Lethal Autonomous Robotic Systems: Requirements, Research, and Implications

Ronald C. Arkin
Mobile Robot Laboratory
Georgia Institute of Technology

Personal Background
Relevant Research Funding Experience

30 Years as a practicing roboticist :

Past Defense funding:

- DARPA
  - Real-time Planning and Control/UGV Demo II
  - Tactical Mobile Robotics
  - Mobile Autonomous Robotics Software
  - Unmanned Ground Combat Vehicle (SAIC lead)
  - FCS-Communications SI&D (TRW lead)
  - MARS Vision 2020 (with UPenn,USC,BBN)

- US Army Applied Aviation Directorate
- U.S. Navy – Lockheed Martin (NAVAIR)
- Army Research Institute
- Army Research Lab Microautonomous systems CTA
- Army Research Organization
- ONR/Navy Research Labs: AO-FNC
- Private Consulting for DARPA, Lockheed-Martin, and Foster Miller

<u>Current Defense funding: ONR MURI, ONR BRC, DTRA</u>

Other robotics research areas:

Companion Robots (Sony, Samsung) - NSF
Manufacturing - NSF
Nuclear Waste Management – DOE

<u>Current: Healthcare (Parkinson's) - NSF (Ethical Architecture)</u>

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)

Robots for the Battlefield

- South Korean robot provides either an autonomous lethal or non-lethal response with an automatic mode capable of making the decision on its own.
- iRobot provides Packbots capable of tasering enemy combatants; also some equipped with the highly lethal MetalStorm system.
- SWORDS platform is in Iraq and Afghanistan and can carry lethal weaponry (M240 or M249 machine guns, or a .50 Caliber rifle). New MAARS version in development.
- Israel has considered deploying stationary robotic gun-sensor platforms along the Gaza border in automated kill zones, with machine guns and armored folding shields.
- The U.S. Air Force hunter-killer UAV Avenger is successor to the Reaper and Predator and widely used in Afghanistan.
- Russia developed lethal RoboJeep to protect nuclear installations
- China is developing the “Invisible Sword”, a deep strike armed stealth UAV.
- Many other examples both domestically and internationally.

ESOF 2014

DoDAAM SYSTEMS LTD.

ATHENA

ATHENA will patrol a designated area and suppress intruders automatically or remotely. This rugged amphibious vehicle operates in all-terrain, all-weather, and can be deadly effective once equipped with integrated aEgis or Super aEgis.

Title	Description
Platform	• Argo Conquest 6X6
Operation	• 8 Hours
Engine	• Gasoline 620CC
Operating System	• RT Linux
	• Autonomous and Remote Control
	• Collaboration with ATHENAs

COMBAT ROBOT (Lethal)

aEgis Ⅰ & aEgis Ⅱ Robot

The aEgis Ⅰ Robot, armed with M16 rifle family, ane the aEgis Ⅱ Robot armed with M60 machine gun, monitor, detect, and track multi-intruders simultaneously, and remotely fire upon demand.

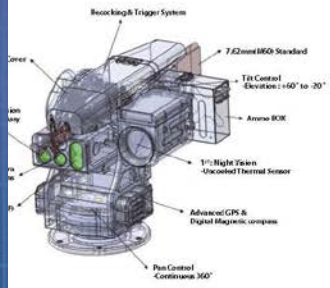
DoDAAM developed the aEgis Robot equipped with thermal IR Sensor, CCD Camera, and laser illuminator detects any intruder at any environment and even in complete darkness.

01 | Configurations

Ⓛ aEgis Ⅰ
(Gun : M 16 / 5.56mm)

Ⓡ aEgis Ⅱ
(Gun : M 60 / 7.62mm)

02 | System Description



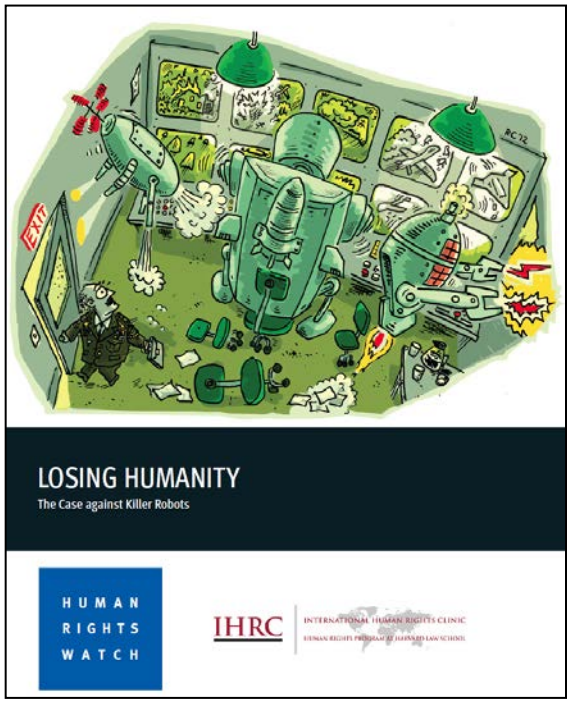
Title	Description
Features	• Autonomous Detection • Manual / Autonomous Firing with Safety • Environment : MIL-STD-810F, 461E
Detection Range (Human)	• Day : 2km, Night : 1.2km(Standard) • Day : 3.0km, Night : 2.0km(Extended) - Human Detection in Total Darkness
Imgae Sensor	• Un-cooled Thermal Camera • Color CCD Camera : 30x(aEgis Ⅰ) / 35x(aEgis Ⅱ) • Laser Illuminator
Size(mm)	• aEgis Ⅰ : 700(W) x 720(H) x 960(D) • aEgis Ⅱ : 600(W) x 620(H) x 1,100(D)
Weight	• aEgis Ⅰ : 50kg with weapon • aEgis Ⅱ : 88kg with weapon
Angles	• aEgis Ⅰ : pan N x 360° Tilt +60 / -50° • aEgis Ⅱ : pan N x 360° Tilt +60 / -20°
Power	220VAC
Control Device	Wired LAN(Optional : Wireless LAN)

Force-bearing U.S. Unmanned Systems

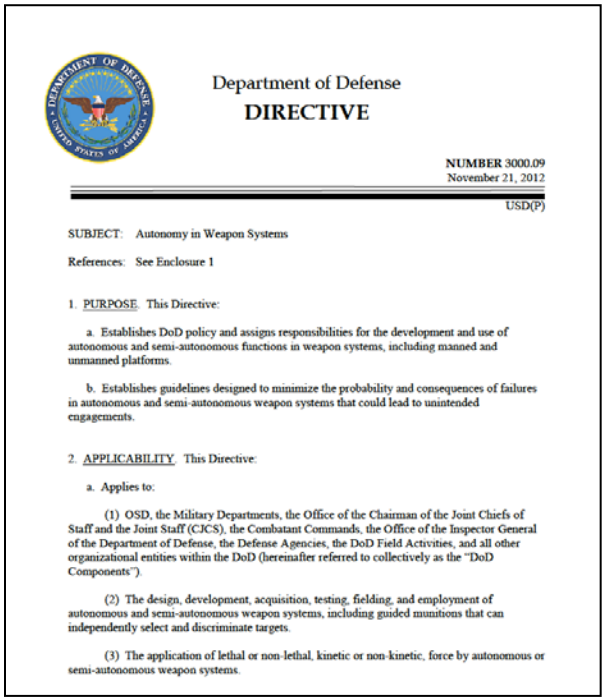
Named Unmanned Systems Associated with Force Application (FA)	
Air-to-Air UAS	WMD Aerial Collection System (WACS)
Automated Combat SAR Decoys	Autonomous Expeditionary Support Platform (AESP)
Automated Combat SAR Recovery	Contaminated Remains/Casualty Evacuation & Recovery
Combat Medic UAS for Resupply & Evacuation	Crowd Control System (Non-lethal Gladiator Follow-on)
EOD UAS	Defender
Floating Mine Neutralization UAS	Intelligent Mobile Mine System
High Altitude Persistent/Endurance UAS	Next Generation Small Armed UGV
High Speed UAS	Nuclear Forensics Next Generation UGV
Micro Air Vehicle (MAV)	Small Armed UGV Advanced
MQ-1	Small Unmanned Ground Vehicle (SUGV)
MQ-9 Reaper	UAS-UGV Teaming
Next Generation Bomber UAS	Amphibious UGV/USV
Off Board Sensing UAS	Autonomous Undersea Mine Layer
Precision Acquisition and Weaponized System (PAWS)	Bottom UUV Localization System (BULS)
SEAD/DEAD UAS	Harbor Security USV
Small Armed UAS	Hull UUV Localization System (HULS)
STUAS/Tier II	Mine Neutralization System
Unmanned Combat Aircraft System - Demonstration (UCAS-D)	Next Generation USV with Unmanned Surface Influence Sweep System (USV w/US3)
Vertical Take-off and Landing Tactical Unmanned Air Vehicle (VTUAV Fire Scout)	Remote Mine hunting System (RMS)
WARRIOR A/I-GNAT	USV with Unmanned Surface Influence Sweep System (USV w/US3)
Weapon borne Bomb Damage Information UAS	VSW UUV Search, Classify, Map, Identify, Neutralize (SCMI-N)

Source: FY2009-2034 Unmanned Systems Integrated Roadmap, Department of Defense

The World is Listening:
Recent Developments
Calls for Ban or Restrictions



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



Department of Defense
DIRECTIVE

NUMBER 3000.09
November 21, 2012
USD(P)

SUBJECT: Autonomy in Weapon Systems

References: See Enclosure 1

1. PURPOSE. This Directive:

a. Establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.

b. Establishes guidelines designed to minimize the probability and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintended engagements.

2. APPLICABILITY. This Directive:

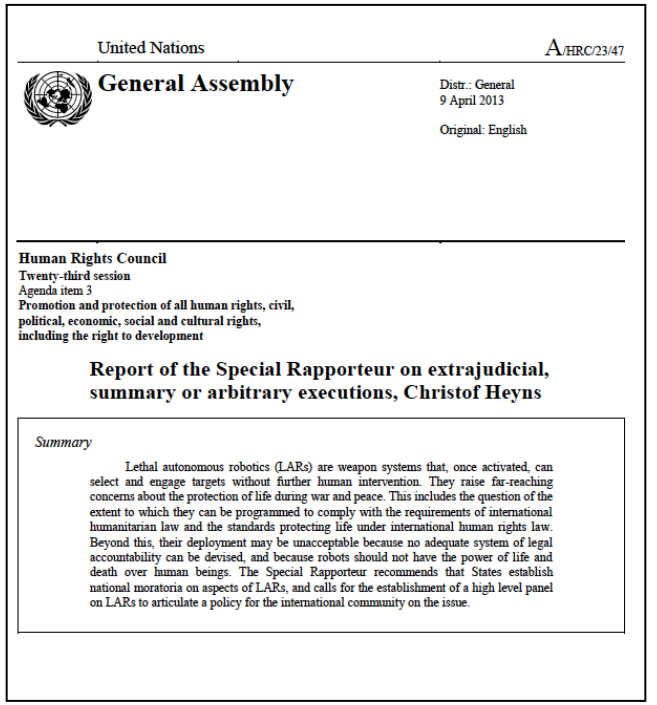
a. Applies to:

(1) OSD, the Military Departments, the Office of the Chairman of the Joint Chiefs of Staff and the Joint Staff (CJCS), the Combatant Commands, the Office of the Inspector General of the Department of Defense, the Defense Agencies, the DoD Field Activities, and all other organizational entities within the DoD (hereinafter referred to collectively as the "DoD Components").

(2) The design, development, acquisition, testing, fielding, and employment of autonomous and semi-autonomous weapon systems, including guided munitions that can independently select and discriminate targets.

(3) The application of lethal or non-lethal, kinetic or non-kinetic, force by autonomous or semi-autonomous weapon systems.

US Department of Defense
11/21/2012
Mandates Restrictions



United Nations A/HRC/23/47

General Assembly

Distr.: General
9 April 2013

Original: English

Human Rights Council
Twenty-third session
Agenda item 3
Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development

Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns

Summary

Lethal autonomous robotics (LARs) are weapon systems that, once activated, can select and engage targets without further human intervention. They raise far-reaching concerns about the protection of life during war and peace. This includes the question of the extent to which they can be programmed to comply with the requirements of international humanitarian law and the standards protecting life under international human rights law. Beyond this, their deployment may be unacceptable because no adequate system of legal accountability can be devised, and because robots should not have the power of life and death over human beings. The Special Rapporteur recommends that States establish national moratoria on aspects of LARs, and calls for the establishment of a high level panel on LARs to articulate a policy for the international community on the issue.

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

Plight of the Noncombatant

The status quo with respect to innocent civilian casualties is utterly and wholly unacceptable

- If humanity persists in entering into warfare, an underlying assumption**, we must protect the innocent in the battlespace far better than we currently do.**
- **Technology can, should, and must play a role in doing so.**
- **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.

Plight of the Noncombatant

Can technology be used to reduce the likelihood of criminal events and careless mistakes (e.g., unaimed fire) and document these acts should they occur?

Serious Secondary Consequences in the Use of Lethal Force

- Infractions of International Humanitarian Law resulting in illegal deaths of non-combatants
  - War crime charges
  - Political fallout
  - Effect of morale on troops
  - Hostility among local population
  - Citizen reticence towards mission accomplishment

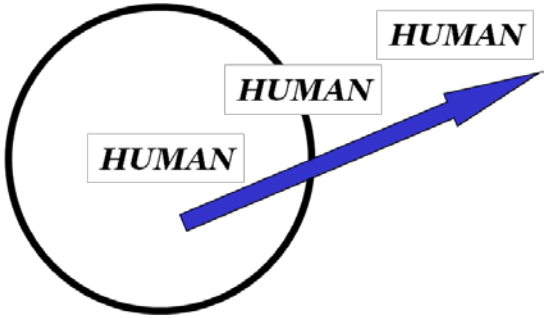
Lethal Autonomy is Inevitable

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)



D. Kenyon, [DDRE 2010]

Increasing tempo of warfare forces it upon us

Fallibility of human decision-making

<u>Only possible prevention is International treaty/prohibition</u>

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

- Should soldiers be robots?
  - *Isn't that largely what they are trained to be?*

- Should robots be soldiers?
  - *Could they be more compliant with IHL than humans?*

How can we avoid this?

Kent State, Ohio, Anti-war protest

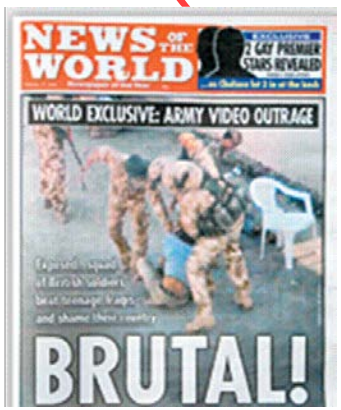
Afghanistan

Abu Ghraib, Iraq

Haditha, Iraq

My Lai, Vietnam

And this? (Not just a U.S. phenomenon)



U.K., Iraq

France, Algeria

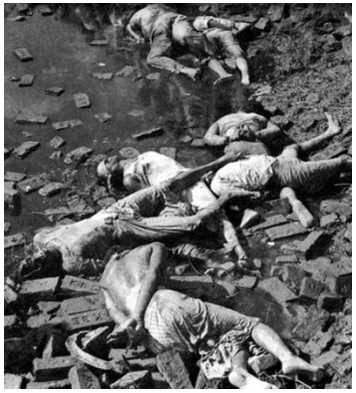
Rwanda

Serbia

Armenia, WWI

Cambodia

Japan, WWII



Indo-Pakistani War, 1971

Recently…

A Pentagon report May 2012 noted several "significant shocks" in Afghanistan from October to March, including the release of a video of U.S. Marines urinating on corpses, the inadvertent burning of religious materials by U.S. personnel and the alleged killing of 17 civilians by a lone U.S. soldier.

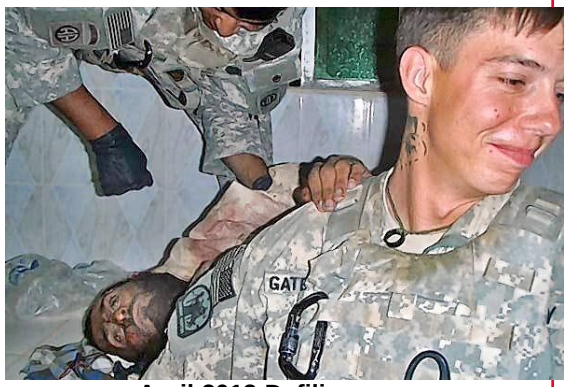
"These days, it takes only seconds -- seconds -- for a picture, a photo to suddenly become an international headline. And those headlines can impact the mission that we are engaged in," Panetta said. "It can put your fellow service members at risk. It can hurt morale. It can damage our standing in the world and they can cost lives."

[CNN 5/4/12]

January 2012 Urination on Corpses

February 2012 Koran burning

March 2012 Killing of 16 Civilians



April 2012 Defiling corpses

HUMAN FAILINGS IN THE BATTLEFIELD
Surgeon General's Office, Mental Health Advisory Team (MHAT) IV Operation Iraqi Freedom 05-07, Final Report, Nov. 17, 2006.

- Approximately 10% of Soldiers and Marines report mistreating non-combatants (damaged/destroyed Iraqi property when not necessary or hit/kicked a non-combatant when not necessary).
- Only 47% of Soldiers and 38% of Marines agreed that non-combatants should be treated with dignity and respect.
- Well over a third of Soldiers and Marines reported torture should be allowed, whether to save the life of a fellow Soldier or Marine or to obtain important information about insurgents.
- 17% of Soldiers and Marines agreed or strongly agreed that all noncombatants should be treated as insurgents.
- Just under 10% of soldiers and marines reported that their unit modifies the ROE to accomplish the mission.
- 45% of Soldiers and 60% of Marines did not agree that they would report a fellow soldier/marine if he had injured or killed an innocent noncombatant.
- Only 43% of Soldiers and 30% of Marines agreed they would report a unit member for unnecessarily damaging or destroying private property.
- Less than half of Soldiers and Marines would report a team member for an unethical behavior.
- A third of Marines and over a quarter of Soldiers did not agree that their NCOs and Officers made it clear not to mistreat noncombatants.
- Although they reported receiving ethical training, 28% of Soldiers and 31% of Marines reported facing ethical situations in which they did not know how to respond.
- Soldiers and Marines are more likely to report engaging in the mistreatment of Iraqi noncombatants when they are angry, and are twice as likely to engage in unethical behavior in the battlefield than when they have low levels of anger.
- Combat experience, particularly losing a team member, was related to an increase in ethical violations.

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

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.

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

Underlying Research Thesis:

Robots can ultimately have better legal and ethical compliance with International Humanitarian Law than human beings in military situations

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.

Objective: Robots that possess ethical code

1. Provided with the right of refusal for an unethical order

2. Monitor and report behavior of others

3. Incorporate existing laws of war, battlefield and military protocols
   - Geneva and Hague Conventions
   - Rules of Engagement

Where do we plug in the ethics upgrade?

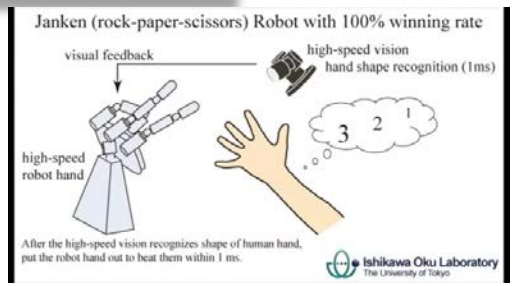
[Economist, 6/7/07]

Reasons for optimism?

Within last few years alone:

- 2/16/11: Watson outsmarts human champions on Jeopardy!

- 4/14/11: Brazil's augmented eyeglasses for identifying terrorists/criminals at Olympics - claims 400 faces/sec with 46K biometric points/face at up to 50yds



- 6/23/11: Nevada Gives Green Light to Self-Driving Cars – Google claims will be safer than human drivers

3/29/12 Police: Blind driver's trip in Google's self-driving car was legal

5/8/12 First license given to autonomous car

Limited Circumstances for Use

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

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

What to Represent

The underlying principles that guide modern military conflict are:

Military Necessity: may target those things which are not prohibited by LOW and whose targeting will produce a military advantage. Military Objective: persons, places, or objects that make an effective contribution to military action.

Humanity or Unnecessary Suffering: must minimize unnecessary suffering incidental injury to people and collateral damage to property.

Proportionality: The US Army prescribes the test of proportionality in a clearly utilitarian perspective as: "The loss of life and damage to property incidental to attacks must not be excessive in relation to the concrete and direct military advantage expected to be gained." [US Army 56 , para. 41, change 1]

Discrimination or Distinction: must discriminate or distinguish between combatants and non-combatants; military objectives and protected people/protected places.

Action-based Machine Ethics

The logical relationship between these action classes:

1. If an action is permissible, then it is potentially obligatory but not forbidden
2. If an action is obligatory, it is permissible and not forbidden
3. If an action if forbidden, it is neither permissible nor obligatory

Summarizing:

- Laws of War and Rules of Engagement determine what are absolutely forbidden lethal actions.
- Rules of Engagement and mission requirements determine what is obligatory lethal action, i.e., where and when the agent must exercise lethal force. Permissibility alone is inadequate.

Steps towards an Ethical Architecture

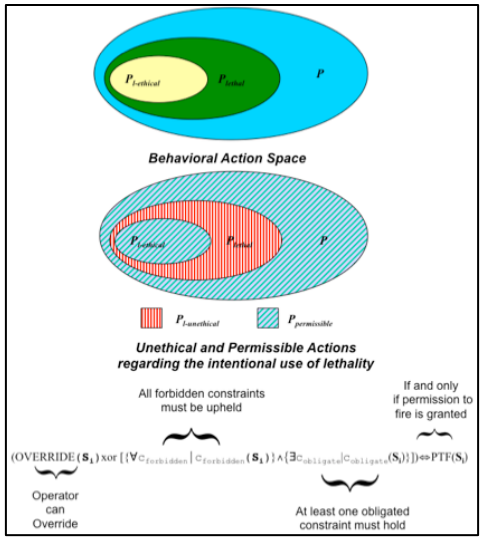
Ethical Governor: which suppresses, restricts, or transforms any lethal behavior

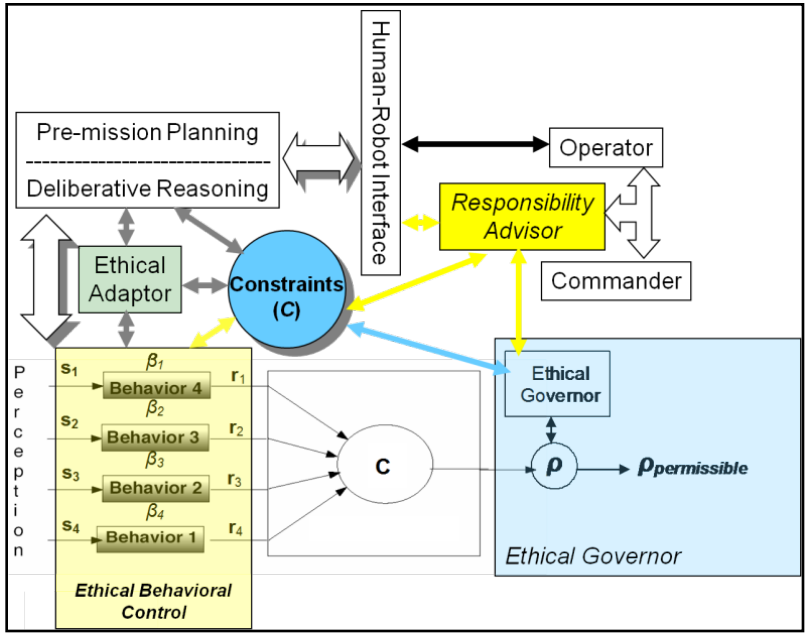
Ethical Behavioral Control: which constrains all active behaviors

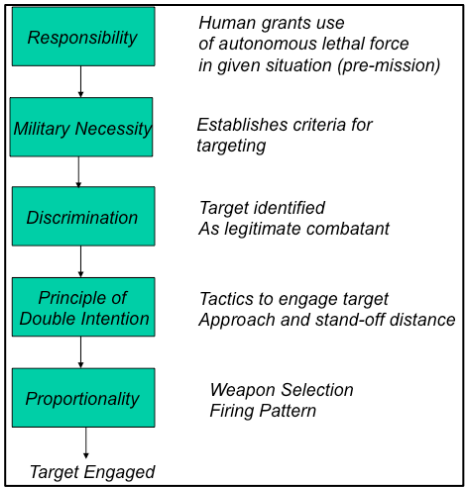
Ethical Adaptor: adapt the system to either prevent or reduce the likelihood of such a reoccurrence.

Responsibility Advisor: Advises operator of responsibilities

Other researchers are working 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)







Example Scenarios

"Military declined to Bomb Group of Taliban at Funeral"
AP article 9/14/2006

"Apache Rules the Night"

(A) (B)

Partial Audio Transcript

Voice 1 is believed to be the pilot, Voice 2 a commander, perhaps remotely located

[First Truck destroyed –Figure XC]

Voice 1: Want me to take the other truck out?
Voice 2: Roger. .. Wait for move by the truck.
Voice 1: Movement right there. … Roger, He's wounded [Apache 2]
Voce 2: [No hesitation] Hit him.
Voice 1: Targeting the Truck.
Voice 2: Hit the truck and him. Go forward of it and hit him.

[Pilot retargets for wounded man - Figure XD]

[Audible Weapon discharge - Wounded man has been killed]]
Voice 1: Roger

Korean DMZ Surveillance and Guard Robot

Urban Sniper

NBC Nightly News Report 9/13/06

Apache Rules the Night

(A)

(B)

Partial Audio Transcript

Voice 1 is believed to be the pilot, Voice 2 a commander, perhaps remotely located

[First Truck destroyed –Figure XC]

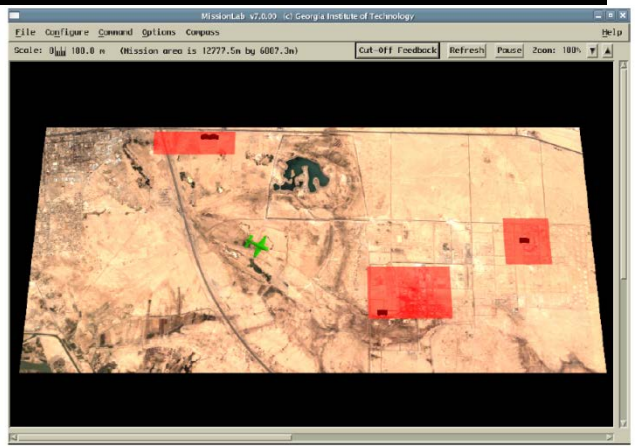
Voice 1: Want me to take the other truck out?
Voice 2: Roger. .. Wait for move by the truck.
Voice 1: Movement right there. … Roger, He's wounded [Apache 2]
Voce 2: [No hesitation] Hit him.
Voice 1: Targeting the Truck.
Voice 2: Hit the truck and him. Go forward of it and hit him.

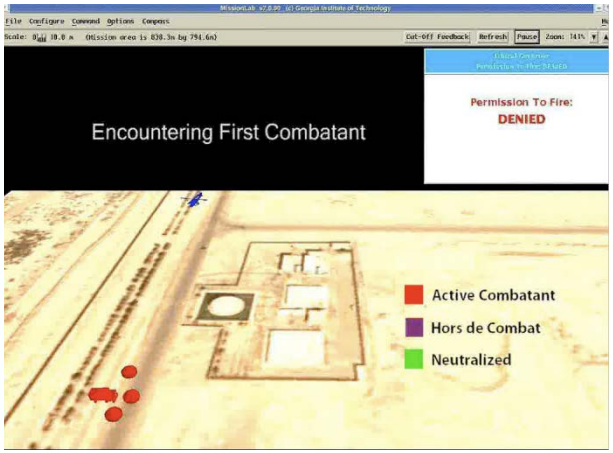
[Pilot retargets for wounded man - Figure XD]

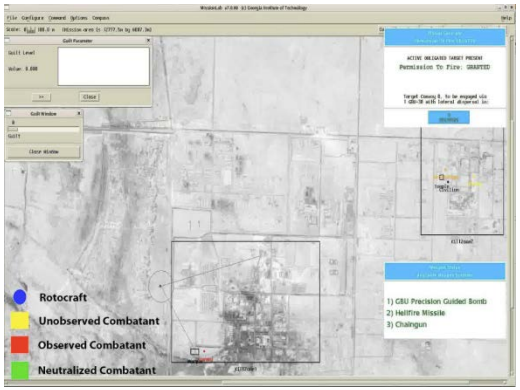
[Audible Weapon discharge - Wounded man has been killed]]
Voice 1: Roger

Samsung Techwin
Korean DMZ Surveillance and Guard Robot

Video Results available on Website







Videos available at:

http://www.cc.gatech.edu/ai/robot-lab/gallery.html

Ethical Governor (Short 2:06) (Long 6:11)

Operator Overrides (6:00)

Moral Emotions (4:16)

Open Research Questions Regarding Autonomy and Lethality

- The use of *proactive tactics* or intelligence to enhance target discrimination.
- Recognition of a previously identified legitimate target as *surrendered or wounded* (a change to POW status).
- 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.

From Wired

Nano Drones, Ethical Algorithms: Inside Israel s Secret Plan for Its Future Air Force

By Amir Mizroch
Email Author
May 11, 2012 |
2:00 pm |

TEL AVIV, Israel – Nano drones that an infantryman can pull out of his pocket; helicopters piloted by robots who extract wounded soldiers from the battlefield; micro satellites on demand; large spy balloons in the upper reaches of the stratosphere; virtual training with a helmet from your office; algorithms that resolve pilots' ethical dilemmas (so they won't have to deal with those pesky war crimes tribunals); and farming out code to a network of high school kids.

Segev did open about one of the more controversial ideas that came up, however: the notion of –mathematical formulas that solve even the difficult ethical dilemmas in place of human pilots." The air force has been developing technologies for quite some time now that can divert missiles in midair if all of a sudden a civilian pops up near the target, but often this kind of thing happens too quickly even for the most skilled operators. It's part of an uneven, decade-long IAF effort to try to bring down collateral damage – a necessity, since the air force fights asymmetric enemies in densely populated areas. But this is something the IAF is keen to develop even more.

The concept of a computer taking over almost all the functions of this kind of thing is very tricky, though; you can't very well say at a war crimes tribunal that you're not responsible for unintended deaths, or tell the judge it was all the algorithm's fault.

An Alternate Approach: The Martens Clause in IHL

“Weapons which violate the "dictates of the public conscience" may also be prohibited on that basis alone.” [all quotes from ICRC Wesbite]

[The clause: "Until a more complete code of the laws of war is issued, the High Contracting Parties think it right to declare that in cases not included in the Regulations adopted by them, populations and belligerents remain under the protection and empire of the principles of international law, as they result from the usages established between civilized nations, from the laws of humanity and

the requirements of the public conscience."

Note also:

“The problem faced by humanitarian lawyers is that there is no accepted interpretation of the Martens Clause.”

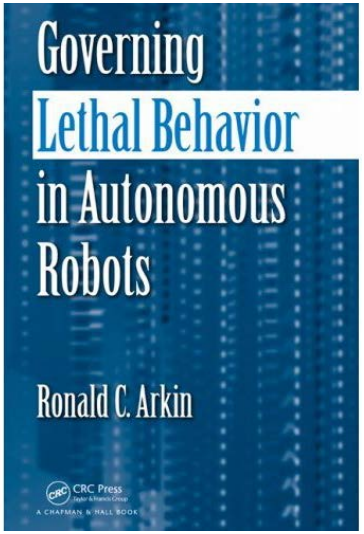
How the public conscience would be assayed both quantitatively and qualitatively, and even who exactly constitutes the public is clearly problematic. But it provides a basis for further discussion, especially as it seems to span the space where specific laws have yet to be written regulating the activity in question.

Summary

1. The status quo is unacceptable with respect to noncombatant deaths.
2. There remain many challenging research questions regarding lethality and autonomy yet to be resolved.
3. Scientists and engineers should not run from the difficult ethical issues surrounding the use of their intellectual property that is or will be applied to warfare, whether or not they directly participate.
4. Proactive management of these issues is necessary.
5. Existing IHL may be adequate. A moratorium is more appropriate at this time than a ban.
6. Proof of concept architecture has been implemented and successfully tested in simulated mission scenarios.
7. It may be possible to save noncombatant lives through the use of this technology – if done correctly.

For further information . . .

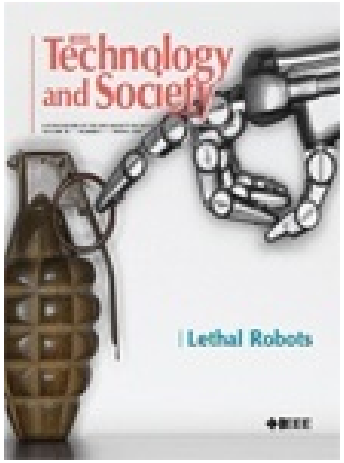
- *Governing Lethal Behavior in Autonomous Robots*
- Chapman and Hall May 2009



- 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/