

Lethality and Autonomous Systems: An Ethical Stance

Ronald C. Arkin
Mobile Robot Laboratory
Georgia Institute of Technology

April 2007



Talk Outline

- Inevitability of the development of autonomous robots capable of lethal force
- Humanity's persistent failings in battlefield ethics
- Research Agenda
(funded by Army Research Organization)
 - Survey opinion on use of Lethal Force by Autonomous Robots
 - Artificial Conscience, to yield *Humane-oids* - Robots that can potentially perform more ethically in the battlefield than humans



Background:

Personal Defense Funding Experience

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 Organization

ONR/Navy Research Labs: AO-FNC

Private Consulting for DARPA, Lockheed-Martin, and Foster Miller



Pre-emptive Strike

The debate here is not about whether or not we should have wars

Rather the question is:

Assuming wars will continue, what is the appropriate role of robotics technology?



Perspective: Future Combat Systems

127 Billion \$ program (recently delayed):

Biggest military contract in U.S. history

Transformation of U.S. Army

Driven by Congressional mandate that by 2010 that “one-third of all operational deep strike aircraft be unmanned” and by 2015 one-third of all ground combat vehicles are unmanned

What are the ethical implications of all this?





FCS Future Combat Systems



Manned Systems



ICV



C2V



Mounted Combat System



Reconnaissance & Surveillance



NLOS Cannon



NLOS Mortar



Maintenance & Recovery



Medical Treatment, Evacuation



Unmanned Air Vehicles (UAV)



Class I & II



Class III/IV

Unmanned Ground Vehicles



Armed Robotic Vehicle



Small Manpackable UGV



Mule

- Unmanned Payloads
- Unattended Ground Sensors
- Unattended Munitions
- Intelligent Munitions
- NLOS LS



Program Objective:

To develop and transition a networked "system of systems" that will serve as the core building block for the US Army's future tactical formations. This capability enables over-matching combat power, sustainability, agility, and versatility necessary for full spectrum military operations.

Technical Challenges:

"Today, our heavy forces are too heavy and our light forces lack staying power. We will address those mismatches with FCS." - GEN Shinseki, CSA, 23 June 1999

The greatest challenges are developing an integrated C4ISR network that enables Battle Command on the move, large scale systems of systems integration, and an expeditionary, lethal, survivable and sustainable force.

Program Status: The DARPA-led Concept and Technology Development (CTD) phase ended with a successful Defense Acquisition Board (DAB) in May 2003. DARPA continues to transition key enabling technologies and lead S&T projects for future advanced warfighting capabilities.



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 soldier's to strike further

The use of robotics for reducing ethical infractions in the military does not yet appear anywhere



Should soldiers be robots?

Isn't that largely what they are trained to be?

Should robots be soldiers?

Could they be more humane than humans?



Motivation for Research

- Battlefield ethics has for millennia been a serious question and constraint for the conduct of military operations
- Breeches in military ethical conduct often have extremely serious consequences, both politically and pragmatically, as evidenced recently by the Abu Ghraib and Haditha incidents in Iraq, which can actually be viewed as increasing the risk to U.S. troops there, as well as the concomitant damage to the United State's public image worldwide.
- If the military keeps moving forward at its current rapid pace towards the deployment of intelligent autonomous robots, we must ensure that these systems be deployed ethically, in a manner consistent with standing protocols and other ethical constraints.



Will Robots be Permitted to Autonomously Employ Lethal Force?

Several robotic systems already use lethal force:

- Cruise Missiles, Navy Phalanx (Aegis 1986 USS Vincennes), Patriot missile, even land mines by some definitions.

Depends on when and who you talk to.

Will there always be a human in the loop?

Fallibility of human versus machine. Who knows better?

Despite protestations to the contrary from all sides, the answer appears to be unequivocally yes.



How can we avoid this?



Kent State, Ohio, Anti-war protest, 4 Dead, May 1970



Abu Ghraib, Iraq

April 2007



Haditha, Iraq



My Lai, Vietnam



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



U.K., Iraq



Rwanda



Germany, Holocaust



Defendant	Position(s)	Charges Convicted Of	Sentence/Fate
Dohara, Kenji	General in army	Conspiracy, waging aggressive war, subverting, ordering, permitting atrocities	Sentenced to hang on December 23, 1948
Hata, Shunroku	Field Marshall in army	Conspiracy, waging aggressive war, disregarding his duty to prevent atrocities	Sentenced to life in prison, paroled 1954, released from parole requirements in 1958
Hirata, Koki	Foreign minister, prime minister	Conspiracy, waging aggressive war, disregarding his duty to prevent atrocities	Sentenced to hang on December 23, 1948
Itagaki, Seishiro	General in army, war minister	Conspiracy, waging aggressive war, ordering or permitting atrocities	Sentenced to hang on December 23, 1946
Kimura, Heitaro	General in army, vice minister of war	Conspiracy, waging aggressive war, ordering or permitting atrocities, disregarding his duty to prevent atrocities	Sentenced to hang December 23, 1946
Koto, Kuntaki	General in Japanese Army, overseas minister, governor-general of Korea, prime minister	Conspiracy, waging aggressive war, disregarding his duty to prevent atrocities	Sentenced to life in prison. Died in 1950 while serving sentence
Matsu, Iwane	General in Army, emperor's representative at Geneva disarmament conference, retired 1938	Disregarding his duty to prevent atrocities	Sentenced to hang on December 23, 1948
Matsu, Akira	General in army, chief of war ministry Military Affairs Bureau	Conspiracy, waging aggressive war, ordering or permitting atrocities, disregarding his duty to prevent atrocities	Sentenced to hang on December 23, 1946
Shigemitsu, Mamoru	Ambassador to China, Soviet Union, and Great Britain, foreign minister	Conspiracy of waging aggressive war, disregarding his duty to prevent atrocities	Sentenced to seven years in prison, paroled 1950, released from parole requirements 1958
Tojo, Hideki	General in army, vice minister of war, war minister, prime minister. Considered most notorious war criminal tried by IMTFE	Conspiracy, waging aggressive war, ordering or permitting atrocities	Sentenced to hang on December 23, 1946

Japan, WWII



Serbia



Cambodia



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.

How can this happen?



Underlying Thesis:
Robots can ultimately be more humane than
human beings in military situations

April 2007



Differentiated Uses for Robots in warfare

Robot as a Weapon:

- Extension of the warfighter
- A human remains in control of the weapons system at all times.
- Standard Practice for today
- Ethics of standard battlefield technology apply
- *This will not be discussed further in this talk from an ethical perspective*

Robot as an Autonomous Agent

- Application of lethal force
- The unmanned system reserves the right to make its own local decisions regarding the application of force directly in the field, without requiring human consent at that moment, either in direct support of the conduct of an ongoing military mission or for the robot's own self-preservation.
- How can ethical considerations be applied in this case?



Humane-oids (Not Humanoids)

Robot soldier



Conventional Robot Weapon

Robot soldier



Humane-oid



Humane-oids (Not Humanoids)

Robot soldier



Robot soldier



Conventional Robot Weapon

Humane-oid

What's the difference?

AN ETHICAL BASIS

April 2007



Robots that have an ethical stance

Right of refusal

Monitor and report behavior of others

Incorporate existing battlefield and military protocols

- Geneva Convention
- Rules of Engagement
- Codes of Conduct

This is not science fiction – but spirit (not letter) of Asimov's laws applies. The robot is bound by the military code of conduct, not Asimov's laws.



Ongoing Research: An Ethical Basis for Autonomous System Deployment (funded by U.S. Army Research Organization)

Given: The robot acts as an intelligent but subordinate autonomous agent.

Research is required to delineate the ethical implications for:

- When the robot reserves the right to make its own local decisions regarding the application of lethal force directly in the field, without requiring human consent at that moment, either in direct support of the conduct of an ongoing military mission or for the robot's own self-preservation.
- When the robot may be tasked to conduct a mission which possibly includes the deliberate destruction of life. The ethical aspects regarding the use of this sort of autonomous robot are unclear at this time and require additional research.



What is acceptable?

*Understand, define, and shape expectations
regarding battlefield robotics*

Task 1: Generation of an Ethical Basis for the Use of Lethality by Autonomous Systems (YEAR 1: UNDERWAY)

Conduct an ethnographic evaluation regarding the dimensions of the ethical basis for the Army's deployment of lethal autonomous systems in the battlefield. This requires interaction with relevant military personnel, ranging from robot operator's to commanders, as well as members of the body politic (policymakers), robot system designers, and the general public.

The end result will be an elaboration of both current and future acceptability of lethal autonomous systems, clarifying and documenting what existing doctrinal thinking is in this regard.

This study will be conducted through formal interviews, survey instruments, literature reviews, and other related sources of information. The end product will be a detailed report and analysis detailing the requirements for the generation of an ethical code of conduct for autonomous systems and the documentation justifying these requirements.



Survey Objectives

Determine people's acceptance of the use of lethal robots in warfare

- Across four communities:
 - ◆ Military
 - ◆ Robotics researchers
 - ◆ Policy makers
 - ◆ General public
- Across levels of autonomy:
 - ◆ Human soldier
 - ◆ Robot as an extension of a soldier
 - ◆ Autonomous robot

Note variation based on demographics



Some Survey Design Principles

1. Questions should be simply-worded and understandable
3. Questions should require an answer
5. Questions should be neither too specific, nor too vague
7. More interesting and motivating questions should go first
9. Randomize to eliminate order effects

Don A. Dillman, "Mail and Internet Surveys: The Tailored Design Method", 2000



Definitions

Robot: as defined for this survey, an automated machine or vehicle, capable of independent perception, reasoning and action

Robot acting as an extension of a human soldier: a robot under the direct authority of a human, including authority over the use of lethal force

Autonomous robot: a robot that does not require direct human involvement, except for high-level mission tasking; such a robot can make its own decisions consistent with its mission without requiring direct human authorization, including decisions regarding the use of lethal force



Question Types

Prior knowledge and attitude

- Robots in general and in the military
- Attitude towards human soldiers and robots in warfare

Possible roles and situations

- How appropriate is using human soldiers vs. robots as extension of a soldier vs. autonomous robots for a number of roles and situations
 - ◆ Direct combat, hostage rescue, etc.



Question Types (2)

Ethics-related questions:

- What it would mean for a robot to be ethical, and to what standards should it be held
- Ability to refuse an unethical order

Responsibility questions

Potential benefits and concerns for using lethal robots in warfare



Question Types (3)

Would it be harder or easier to start wars with robot involvement?

If possible, would any emotions be beneficial for a military robot?



Demographics Questions

Age, gender, cultural upbringing

Education, occupation

Military, policy making or robot research experience

Technology and robot experience and attitude

Attitude to war

Spirituality/religion



Pilot study Conducted

Goal: improve the quality of the survey

20 people total, 19 fully completed

5 with military experience, 3 with policy making experience, and 5 with robot research experience

14 had higher education

12 male, 7 female

Wide age range



Results, even preliminary, cannot be provided until survey completed to avoid the introduction of bias



Timetable

Date	Task
August 2006	Project began
October 2006	Pilot study submitted to IRB
December 2006	Pilot study completed
January 2007	Revised survey submitted to IRB
March 2007	Survey started
Late 2007	Survey completed
End of 2007	Data analysis

April 2007



What can be done?

Artificial Conscience and Reflection

Task 2: Computational implementation of an ethical code within an existing autonomous robotic system, i.e., an “artificial conscience”.
(YEAR 2-3)

- Provide enforceable limits on acceptable behavior (behavioral governor)
- Drawing on ethical precepts extracted from sources such as the Geneva convention and other related protocols and the results of Task 1, the robot will be able to consider, in real-time, the consequences of its behavioral actions in situ, and thus potentially lead to a robotic soldier that may indeed operate in a more ethical and humane manner than even many human warfighters currently do.
- In support of this effort, a reflective component to the architecture will be elaborated in order to effectively evaluate the consequences of present actions in a more global context.
- Investigation into guilt as a robotic motivational (emotional) component.



Reiterating:

Objective: Robots that possess ethical code

1. Provided with the right of refusal for an unethical order
3. Monitor and report behavior of others
5. Incorporate existing laws of war, battlefield and military protocols
 - Geneva Convention
 - Rules of Engagement
 - Codes of Conduct



Example Scenario:
“Military declined to Bomb Group
of Taliban at Funeral”
AP article 9/14/2006



(Left) Reconnaissance Photo showing a Taliban Muster

(Right) Predator UAV



Summary

1. Roboticists 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 you directly participate. Wars unfortunately will continue and derivative technology from your ideas will be used.
3. Proactive management of these issues is necessary.
5. Research is ongoing on only a few of these issues in this and other related ethical areas in robotics.
7. Formalization of rules and guidelines for researchers as well as consciousness-raising is essential at this time to avoid a Pugwash-style after-the-fact effect. Bioengineering has much to teach us in that regard.



For further information . . .

Mobile Robot Laboratory Web site

- <http://www.cc.gatech.edu/ai/robot-lab/>

Contact information

- ◆ Ron Arkin: arkin@cc.gatech.edu

IEEE RAS Technical Committee on Robo-ethics

http://www-arts.sssup.it/IEEE_TC_RoboEthics

CS 4002 – Robots and Society Course

http://www.cc.gatech.edu/classes/AY2007/cs4002_spring/

