### *Humane-oids*: Towards An Ethical Basis for Autonomous System Deployment

Ronald C. Arkin CNRS-LAAS/ Toulouse and Mobile Robot Laboratory Georgia Tech Atlanta, GA, U.S.A.

### Talk Outline

 Inevitability of the development of autonomous robots capable of lethal force

Humanity's persistent failings in battlefield ethics

#### • An alternate view:

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 (NAVAIR)
- Army Research Institute
- ONR/Navy Research Labs: AO-FNC
- Corporate: SAIC
- Private Consulting for DARPA 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?

#### **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

# Differentiated Uses for Robots in warfare

Robot as a Weapon:

- Extension of the warfighter
- 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
   How can ethical considerations be applied

# Will Robots be Permitted to Autonomously Employ Lethal Force?

- Several robotic systems already use lethal force:
  - Cruise Missiles, Navy Phalanx, 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.

That is not the end of the discussion.

#### **Perspective:** Future Combat Systems

 127 Billion \$ program (recently delayed): Biggest military contract in US history

Transformation of 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?



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.

#### **Representative US Military Robotic Programs**

#### • Note:

- All video material that follows is classified for public release, distribution unlimited or downloaded from internet.
- All credit for the videos shown vests with the system developer.

#### Boeing Unmanned Combat Aerial Vehicle (UCAV)



#### GDRS Experimental Unmanned Vehicle (XUV)

#### Ft. Indiantown Gap November 2001

#### USMC Tactical Unmanned Ground Vehicle (TUGV)



#### UAV Launch from MDARS (SPAWAR)



#### **UUVAutonomous Reconnaissance, Surveillance & Docking** Woods Hole Oceanographic Institution



### **Current Deployments**

in | Change edition

#### B B C NEWS WORLD EDITION

Last Updated: Sunday, 23 January, 2005, 16:45 GMT
E-mail this to a friend
Frintable version

US plans 'robot troops' for Iraq

The US military is planning to deploy robots armed with machine-guns to wage war against insurgents in Iraq.

Eighteen of the 1m-high robots, equipped with cameras and operated by remote control, are going to Iraq this spring, the Associated Press reports.

The machine is based on a robot already used by the military to disable bombs.



Soldering on: A private firm helped the military make the robo-soldier

Officials say the robot warrior is fast, accurate and will track and attack the enemy with relatively little risk to the lives of US soldiers.

Unlike its human counterparts, the armed robot does not require food, clothing, training, motivation or a pension.

When not needed in war, it can be mothballed in a warehouse.

#### Roboethics Atelier

However, the robot will rely on its human operator, remotely studying footage from its cameras, for the order to open fire.

#### So where does ethics fit?

#### One possible view

### Humane-oids (Not Humanoids)



#### Conventional Robot Weapon

### Humane-oids (Not Humanoids)



Robot soldier



#### Conventional Robot Weapon Humane-oid What's the difference?

### Humane-oids (Not Humanoids)



Robot soldier



#### **Conventional Robot Weapon** Humane-oid What's the difference? AN ETHICAL BASIS **Roboethics** Atelier

20

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

#### 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 of Asimov's laws applies. The robot is bound by the military code of conduct, not Asimov's laws.

### Why is this needed?

# Can robots outperform humans on an ethical basis?

WARNING: Many of the following slides are extremely graphic – illustrating man's inhumanity to man in warfare

### U.S. - Abu Ghraib



#### Roboethics Atelier





24

### British - Iraq





## Germany - Holocaust



# Japan – World War II





Defendant	Position(s)	Charges Convicted Of	Sentence/Fate
Doihara, Kenji	General in army	Conspiracy; waging aggressive war; authorizing, ordering, or 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
Hirota, 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
Koiso, Kuniaki	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
Matsui, 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
Muto, 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	Convicted 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

### Cambodia



## Rwanda



## U.S. – My Lai, Vietnam







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?

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?

### Baby steps forward?

#### A few ideas gleaned from two proposals I generated in this area

- 1. Non-lethal force for mob control
- 2. Ethical Battlefield Autonomous Systems It may sound oxymoronic but here I refer to robotic systems that are potentially more ethical than human warfighters

### Case 1: How can we avoid this?



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

#### Non-Lethal Force in Crowd/Riot Control



Goal: Use robotics to avoid fatalities as seen in:

- U.S. Kent State Anti-war protest
- Afghanistan Pakistan Cartoon Riots
- Numerous others

#### **Proposal: Cognitive Actuation: Agonistic Behavior for Robot-Human Interaction**

 Can models of agonistic behavior, suitably embedded in a hybrid deliberative/reactive robotic architecture be used to defuse and manage human conflict?

 Can cognitive models of human affective state, both individual and collective (i.e., a mob mentality) be used to control action-selection to produce desirable outcomes in human-robot conflict without resorting to lethality?

### Natural Agonistic Systems



#### **Group behavior in flocking**



# Deimatic behavior in the praying mantis

## Proposed Research Agenda

- 1. Develop cognitive models of human affective state that pertain to both individuals and mobs and use them to influence behavior selection consistent with producing desired changes in the surrounding human behavioral environment.
- 2. Create an agonistic subsystem for conflict resolution drawn from biological models, to deflect attacks before they occur or reflect them when they occur but without lethal force.
- 3. Create agonistic robotic behaviors to manage ("actuate") humans in dangerous situations, permitting robots to induce changes in human behavior causing people to move out-of-harm's way, consistent with high-level goals.
- 4. **Develop robot behaviors from existing crowd control protocols** ensuring that rules-of-engagement and other doctrine are adhered to.
- 5. Unpredictablity is a hallmark characteristic for action selection throwing people off-guard to defuse the situation. Confrontational and appeasement behaviors will be created.
- 6. Incorporate cognitive models of human individuals and crowds to monitor the situation and evoke suitable behaviors as needed, drawn on emotional models of both mobs and individuals.

Case 2: An Ethical Basis for Autonomous System Deployment

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

Research is required to delineate the ethical implications for:

- 1. 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.
- 2. 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.

### Motivation

- Battlefield ethics has for millennia been a serious question and constraint for the conduct of military operations by commanders, soldiers, and politicians, as evidenced for example by the creation of the Geneva conventions, the production of field manuals to guide appropriate activity for the warfighter in the battlefield, and the specific rules of engagement for a given military context.
- Breeches in military ethical conduct often have extremely serious consequences, both politically and pragmatically, as evidenced recently by the Abu Ghraib incident 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 that draw from cultural relativism (our own society's or the world's ethical perspectives), deontology (rightbased approaches), or within other related ethical frameworks.

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

- Conduct an ethnographic evaluation of the ethical basis for lethal autonomous systems in the battlefield. This requires interaction with military personnel, from robot operator's to commanders, as well as policymakers, robot designers, and the public.
- The 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.
- Interviews, survey instruments, literature reviews, and other sources will be used, resulting in a report and analysis of the requirements for the generation of an ethical code of conduct for autonomous systems and documentation justifying these requirements.

### What can be done? Artificial Conscience and Reflection

- Task 2: Computational implementation of an ethical code within an existing autonomous robotic system, i.e., "artificial conscience".
  - Provide enforceable Limits on acceptable behavior (behavioral governor)
  - Drawing on ethical precepts from sources such as the Geneva convention and other related protocols and Task 1 results, the robot will consider in real-time the consequences of its actions in situ, and potentially lead to a robotic soldier that may operate in a more ethical and humane manner than even many human warfighters currently do.
  - A reflective component will be elaborated to effectively evaluate the consequences of present actions in a more global context.
  - Investigation into guilt as a robotic motivational (emotional) component.

#### What should we do be next?

Follow Bioethics Community Lead

- Hold Asilomar-style Conference
- Delineate All Classes of Robotics Research (not just military) on basis of ethical considerations
- Generate recommendations for each class

 Produce a roadmap and use existing societal and political bodies (e.g., IEEE) to further an ethical agenda

### For further information . . .

Mobile Robot Laboratory Web site - http://www.cc.gatech.edu/ai/robot-lab/ Contact information Ron Arkin: <u>arkin@cc.gatech.edu</u> IEEE RAS Technical Committee on Roboethics http://www-arts.ssup.it/IEEE\_TC\_RoboEthics CS 4002 – Robots and Society http://www.cc.gatech.edu/classes/AY2005/cs4002\_spring/