REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
24-09-2009	Final Report	1-Jul-2006 - 30-Nov-2006

4. TITLE AND SUBTITLE
FINAL REPORT: An Ethical Basis for Autonomous System Deployment

5a. CONTRACT NUMBER
W911NF-06-1-0252

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Ronald C. Arkin

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Georgia Tech Research Corporation
Office of Sponsored Programs
Georgia Tech Research Corporation
Atlanta, GA 30332 -0415

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
50397-NS.1

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This project investigated and implemented an ethical basis for deployment of lethality in autonomous robotic systems. Two main thrusts were explored. The first addresses the ethical dimensions of robotic weaponry in two contexts: the robot as an extension of the warfighter and the robot as an autonomous agent acting on behalf of the warfighter. A formal survey has been completed among a broad population of relevant parties including military personnel, the public, policymakers, and roboticists. The results characterize the decision-making space for the

15. SUBJECT TERMS
autonomous systems, robot ethics, lethality

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ronald Arkin
U	U	U	SAR		19b. TELEPHONE NUMBER 404-894-8209

Report Title

FINAL REPORT: An Ethical Basis for Autonomous System Deployment

ABSTRACT

This project investigated and implemented an ethical basis for deployment of lethality in autonomous robotic systems. Two main thrusts were explored. The first addresses the ethical dimensions of robotic weaponry in two contexts: the robot as an extension of the warfighter and the robot as an autonomous agent acting on behalf of the warfighter. A formal survey has been completed among a broad population of relevant parties including military personnel, the public, policymakers, and roboticists. The results characterize the decision-making space for the deployment of intelligent robotic weaponry, whereby the military can judiciously determine its most effective and appropriate usage. The second component involved the generation of an artificial "conscience" for an intelligent autonomous robotic agent, which applies limits and constraints on its actions as required by the bounds of ethical decision making. These limits are generated from the Laws of War, rules of engagement, and other requirements. The intent is to yield robots that can perhaps act more humanely than humans do under highly stressful conditions; provide warnings in the field to military decision-makers about the potential ethical consequences of tactical actions regarding the use of this technology; and to ensure that accountability is engineered into these systems from the onset.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Arkin, R.C., 2009. "Ethical Robots in Warfare", IEEE Technology and Society Magazine, Vol. 28, No. 1, pp. 30-33, Spring 2009.

Number of Papers published in peer-reviewed journals: 1.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

1. VI Latin American Robotics Symposium (LARS 2009), "Ethics and Lethality in Autonomous Combat Robots", Keynote Lecture, Valparaiso, Chile, October 2009.

2. Workshop on Military Operations, National Security, and Emerging Technologies, "Ethical Robots in Combat", Invited Talk, Case Western Reserve University, Cleveland, OH, Oct. 2009.

3. International Conference on Knowledge Intensive Mult-agent Systems (KIMAS 2009), "Ethics and Lethality in Autonomous Combat Robots", Distinguished Plenary Lecture, St. Louis, MO, October 2009.

4. Seminar for Science, Theology and Ethics, "Bombs, Bonding and Bondage: Current Issues in Human-Robot Interaction", Virginia Conference of the United Methodist Church, Webinar, Blackstone, VA, Oct. 2009.

5. Panel on Ethics in Unmanned Systems in Combat, AUVSI '09, Washington, DC, August 2009.

6. Air Force Unmanned Aircraft Systems Academic Outreach Symposium, "Ethics and Lethality in Autonomous Combat Robots", invited talk, and panel on "Ethics and Autonomous Systems", Grand Forks, ND, August 2009.

7. 2009 International Symposium on Technology and Society (ISTAS '09), "Ethical Robots in Combat", Panel on How Sustainable is a Society that Employs Autonomous Robots?, Tempe, AZ, May 2009.

8. Army Science Board Study on Armed Ground Robots, "Ethics and Lethality in Autonomous Combat Robots", Arlington, VA, April 2009.

9. 6th International Symposium on Mechatronics and its Applications (ISMA '09), "Embedding Ethical Constraints into Robotic Systems", Keynote Lecture, Sharjah, UAE, March 2009.

10. Technology and Ethics Seminar, "Ethics and Lethality in Autonomous Robots", Yale University Interdisciplinary Center for Bioethics, New Haven, CT, Feb. 2009.

11. Korea University Graduate Seminar Series, "Embedding Ethical Constraints into Robotic Systems", Seoul, KR, Nov. 2008.

12. Booz Allen Hamilton Technology Focus Group, "Ethics and Lethality in Autonomous Combat Robots", Tysons Corner, VA, October 2008.

13. 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008), "Embedding Ethical Constraints into Robotic Systems", Plenary Lecture, Seoul, KR, August 2008.

14. 2008 North American Computing and Philosophy Conference, Donald C. Engelbart Keynote Lecture, "Ethics and Lethality in Autonomous Combat Robots", Bloomington IN, July 2008.

15, First International Conference on Human-Robot Personal Relationships, "Ethical Aspects of Personal Human-Robot Interaction". Plenary Lecture, Maastricht, NL, June 2008.

16. Washington and Lee University Seminar, "Ethics and Lethality in Autonomous Military Robots", VA, April 2008.

17. UNC Charlotte Ethics in Emerging Technologies Symposium, "Governing Lethal Behavior: Embedding Ethics in an Autonomous Robot Architecture", Plenary lecture, Charlotte, NC, April 2008.

18, University of Michigan EECS Seminar, "Ethics and Lethality in Autonomous Systems", Ann Arbor, MI, March 2008.

19. 3rd ACM/IEEE international Conference on Human-Robot Interaction (HRI- 2008), Panel Discussion on Robo-Ethics, Amsterdam, NL, March 2008.GSU Neurophilosophy

20. Brown Bag Lunch Series, "Governing Lethal Behavior: Embedding Ethics in an Autonomous Robot Architecture", Georgia State University, February 2008.

21. Royal United Services Institute/British Computer Society Ethics of Autonomous Systems Workshop, "Ethics and Lethality in Autonomous Systems", Keynote Session, London UK, February 2008.

22. Workshop on the Behavioral Dynamics of Heterogeneous Teams of Humans and Machines, "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture Systems", Princeton, NJ, Nov. 2007.

23. Unmanned Systems Council Meeting, "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture Systems", Dahlgren Naval Surface Warfare Center, Dahlgren, VA, September 2007.

24. OSD Intelligent Autonomy Workshop "Moving to the Next Level: Experiential and Ethical Reasoning for Autonomous Systems", invited talk, Rosslyn, VA, September 2007.

25. JC-UGV TACOM Robotics Workshop, "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture Systems", invited talk, TARDEC, Detroit, MI, September, 2007.

26, ICRA 2007 Workshop on Roboethics, "Lethality and Autonomous Robots: An Ethical Stance", Rome IT, April 2007.

Number of Presentations: 26.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

1. Arkin, R.C., Wagner, A., and Duncan, B., 2009. "Responsibility and Lethality for Unmanned Systems: Ethical Pre-mission Responsibility Advisement", Proc. 2009 IEEE Workshop on Roboethics, Kobe JP, May 2009.

2. Moshkina, L. and Arkin, R.C., 2008. "Lethality and Autonomous Systems: The Roboticist Demographic", Proc. ISTAS 2008, Fredericton, CA, June 2008.

3. Arkin, R.C., 2008. "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part I: Motivation and Philosophy", Proc. Human-Robot Interaction 2008, Amsterdam, NL, March 2008.

4. Arkin, R.C., 2008. "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part II: Formalization for Ethical Control", Proc.
1st Conference on Artificial General Intelligence, Memphis, TN, March 2008.

5. Arkin, R.C., 2008. "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part III: Representational and Architectural Considerations", Proceedings of Technology in Wartime Conference, Palo Alto, CA, January 2008.

6. Arkin, R.C. and Moshkina, L., 2007, "Lethality and Autonomous Robots: An Ethical Stance" Proc. International Symposium on Technology and Society, Las Vegas, June 2007.

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 6

(d) Manuscripts

1. Arkin, R.C., and Ulam, P., 2009. "An Ethical Adaptor: Behavioral Modification Derived from Moral Emotions", GVU Technical Report GIT-GVU-09-04, GVU Center, Georgia Institute of Technology, 2009.

2. Arkin, R.C., 2009. "Accountable Autonomous Agents: The Next Level", Position paper for the DARPA Complete Intelligence Workshop, Feb. 2009.

3. Arkin, R.C., Ulam, P., and Duncan, B., 2009. "An Ethical Governor for Constraining Lethal Action in an Autonomous System", GVU Technical Report GIT-GVU-09-02, GVU Center, Georgia Institute of Technology, 2009.

4. Arkin, R.C., Wagner, A.R., and Duncan, B., 2009. "Responsibility and Lethality for Unmanned Systems: Ethical Pre-Mission Responsiblity Advisement", GVU Technical Report GIT-GVU-09-01, GVU Center, Georgia Institute of Technology, 2009.

5. Moshkina, L., and Arkin, R.C., 2007. "Lethalilty and Autonomous Systems: Survey Design and Results", GVU Technical Report GIT-GVU-07-16, GVU Center, Georgia Institute of Technology, 2007.

6. Arkin, R.C., 2007. "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture", GVU Technical Report GIT-GVU-07-11, GVU Center, Georgia Institute of Technology, 2007.

Number of Manuscripts: 6.00

Number of Inventions:

Graduate Students

NAME	PERCENT_SUPPORTED
Patrick Ulam	0.50
Alan Wagner	0.50
Lilia Moshkina	0.50
FTE Equivalent:	**1.50**
Total Number:	**3**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Ronald Arkin	0.10	No
FTE Equivalent:	**0.10**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
Brittany Duncan	0.38
FTE Equivalent:	**0.38**
Total Number:	**1**

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	1.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....	1.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....	1.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	0.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

Inventions (DD882)

An Ethical Basis for Autonomous System Deployment

Proposal 50397-CI

Ronald C. Arkin, College of Computing, Georgia Tech

FINAL REPORT ATTACHMENT

I. PROBLEM STATEMENT

The objectives of this project were twofold:

- To gauge opinion on the use of lethality by autonomous systems in the battlefield from various demographic groups: the military, robotics researchers, policy makers, and the general public.
- To embed an ethical code within a robotic controller to govern its behavior in a manner consistent with the laws of war, rules of engagement and code of conduct of the military.

To accomplish these objectives, the project investigated and implemented an ethical basis for deployment of lethality in autonomous robotic systems. Two main thrusts were explored.

1. The first task addressed the ethical dimensions of robotic weaponry in two contexts: the robot as an extension of the warfighter and the robot as an autonomous agent acting on behalf of the warfighter. A formal survey has been completed among a broad population of relevant parties including military personnel, the public, policymakers, and roboticists. The results characterize the decision-making space for the deployment of intelligent robotic weaponry, whereby the military can judiciously determine its most effective and appropriate usage.

2. The second task involved the generation of an artificial "conscience" for an intelligent autonomous robotic agent, which applies limits and constraints on its actions as required by the bounds of ethical decision-making. These limits are generated from the Laws of War, rules of engagement, and other requirements. The intent is to yield robots that can perhaps act more humanely than humans do under highly stressful conditions; provide warnings in the field to military decision-makers about the potential ethical consequences of tactical actions regarding the use of this technology; and to ensure that accountability is engineered into these systems from the onset.

II. ACCOMPLISHMENT SUMMMARY

We have successfully completed all of the goals of this three-year project.

1. The survey was closed and completed on October 20, 2007. The results of the completed survey are documented in a technical report entitled *Lethality and Autonomous Systems: Survey Design and Results,* L. Moshkina and R. Arkin 2007 (GIT-GVU-TR-07-16) and a new book published in May 2009 (Objective 1). This report summarizes the year one results obtained on our survey. Specifically it reports the methods and results of an on-line survey addressing the issues surrounding lethality and autonomous systems. The data from this survey were analyzed both qualitatively, providing a comparison between four different demographic samples targeted in the survey (namely, robotics researchers, policymakers, the military, and the general public), and quantitatively, for the robotics researcher demographic. In addition to the analysis, the design and administration of this survey and a discussion of the survey results are provided.

2. The design, development and implementation of the ethical architecture (Objective 2), including its philosophy and motivation, formal mathematical development, and testing have been completed. This is summarized in a second technical report entitled: *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture,* Arkin, R., 2007 (GIT-GVU-TR-07-11) and a new book published in May 2009. This report describes the basis, motivation, theory, and design recommendations for the implementation of an ethical control and reasoning system potentially suitable for constraining lethal actions in an autonomous robotic system so that they fall within the bounds prescribed by the Laws of War and Rules of Engagement. It is based upon extensions to existing deliberative/reactive autonomous robotic architectures, and includes recommendations for (1) post facto suppression of unethical behavior, (2) behavioral design that incorporates ethical constraints from the onset, (3) the use of affective functions as an adaptive component in the event of unethical action, and (4) a mechanism in support of identifying and advising operators regarding the ultimate responsibility for the deployment of such a system.

II.1 Task 1: Questionnaire Summary ***(from Technical Report)***

A high-level summary of the survey results follows. The interested reader should refer to the technical report mentioned on the previous page. After analyzing the results of the survey, the following generalizations can be made:

1. **Demographics:**
   - A typical respondent was an American or Western European male in his 20s or 30s, with higher education, significant computer experience, and positive attitude toward technology and robots.
   - The participants ranged from under 21 to over 66 years old (all the participants were over 18); 11% of the participants were female; non-US participants were from all over the world, including Australia, Asia, Eastern Europe and Africa.

2. **Levels of Autonomy:**
   - In general, regardless of roles or situations, the more the control shifts away from the human, the less such an entity is acceptable to the participants. A human soldier was the most acceptable entity in warfare, followed by the robot as an extension of the warfighter, and autonomous robot was the least acceptable.
   - There was a larger gap in terms of acceptability between a robot as an extension and autonomous robot than that between soldier and robot as an extension.
   - Taking human life by an autonomous robot in both Open Warfare and Covert Operations is unacceptable to more than half of the participants (56% disagreed or strongly disagreed), especially in the case of Covert Operations on Home Territory.

3. **Comparison between Community Types:**
   - Regardless of roles or situations, in most cases the general public found the employment of soldiers and robots less acceptable than any other community type, and, conversely, those with military experience and policymakers found such employment more acceptable.
   - More military and policymakers were in favor of the same ethical standards for both soldiers and robots than both the general public and roboticists, who were more in favor of higher standards for robots.
   - When asked about the responsibility for any lethal errors, those with military experience attributed the least amount of blame to any of the responsible parties.

4. **Roles:**
   - The most acceptable role for using both types of robots is Reconnaissance; the least acceptable is Crowd Control.
   - Robots could be used for roles where less force is involved, such as Sentry and Reconnaissance, and should be avoided for roles where use of force may

be necessary, especially when civilian lives are at stake, such as Crowd Control and Hostage Rescue.

5. **Situations:**
   - Covert Operations were less acceptable to the entire set of participants than Open Warfare (whether on Home or Foreign Territory).

6. **Ethical Considerations:**
   - The majority of participants, regardless of the community type, agreed that the ethical standards, namely, Laws of War, Rules of Engagement, Code of Conduct and Additional Moral Standards, do apply to both soldiers (84%) and robots (72%).
   - The more concrete, specific and identifiable ethical standards were, the more likely they were to be considered applicable to both soldiers and robots, with Laws of War being the most applicable, and Additional Moral Standards the least.
   - 66% of the participants were in favor of higher ethical standards for a robot than those for a soldier.
   - 59% of the participants believed that an autonomous robot should have a right to refuse an order it finds unethical, thus in a sense admitting that it may be more important for a robot to behave ethically than to stay under the control of a human.

7. **Responsibility:**
   - A soldier was the party considered the most responsible for both his/her own lethal errors, and for those of a robot as an extension under his/her control. Robots were the least blamed parties, although an autonomous robot was found blameworthy twice as much as robot as an extension. It is interesting that even though robots were blamed the least, 40% of the respondents still found an autonomous robot responsible for its errors to a very significant or significant extent.
   - As the control shifts away from the soldier, the robot and its maker should take more responsibility for robot's actions. A robot designer was blamed 31% less for the mistakes of a robot as an extension than those of an autonomous robot.

8. **Benefits and Concerns:**
   - Saving lives of soldiers was considered the most clear-cut benefit of employing robots in warfare and the main concern was that of risking civilian lives. Saving soldiers' lives and decreasing psychological trauma to soldiers outweigh the risk to the soldiers the most. Decreasing cost and producing better battlefield outcome were also viewed as benefits rather than concerns.
   - For the roboticists, the categories regarding battlefield outcomes and friendly fire were not considered strongly as either benefits or concerns, suggesting that the participants did not think that robots would have an effect on these categories.

9. Wars and Emotions:

- The majority of the participants (69%) believe that it would be easier to start wars if robots were employed in warfare.
- Sympathy was considered to be beneficial to a military robot by over half of the participants (59%), and guilt by just under a half (49%). The majority of the participants (75%) were against anger in a military robot.

10. Cultural Background:

US participants were more likely to accept both soldiers and robots in proposed roles and situations than non-US participants. They favored less stringent ethical standards for robots and were less likely to give the robot a right to refuse an unethical order than non-US participants. They were also less likely to assign responsibility for lethal errors of soldiers and robots and less willing to provide military robots with emotions.

11. Firearms Experience:

- Those with less firearms experience found the use of all three levels of autonomy for the proposed roles, overall, less acceptable than those with more experience, and found the use of both types of robots less acceptable in the proposed situations.
- Those with less firearm experience were also more likely to hold a robot to more stringent ethical standards when compared to those of a soldier; more likely to allow a robot to refuse an unethical order, more prone to assign responsibility for lethal errors of soldier and robot as extension, and more willing to provide military robots with the emotions of sympathy, guilt and happiness.

12. Spirituality:

In most cases, spirituality had no effect on the participants' opinions with the exception of the use of robot as an extension for the proposed roles and the use of all three levels of autonomy in the given situations. Those of higher spirituality found such use more acceptable in warfare; also, more spiritual/religious participants were less convinced that it would be easier to start wars if robots were brought onto the battlefield.

II. 2 Task 2: Ethical Architecture Summary

The process for conducting Task 2 involved the following stages:

1. **Understanding of philosophical and legal underpinnings:**
   We first presented the background, motivation and philosophy for the design of an ethical autonomous robotic system capable of using lethal force. The system is governed by the Laws of War and Rules of Engagement using them as constraints.

   Arkin, R.C., "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part I: Motivation and Philosophy", *Proc. Human-Robot Interaction 2008*, Amsterdam, NL, March 2008.

2. **Development of underlying mathematical formalisms:**
   We provided the permeating formalisms for a hybrid deliberative/reactive architecture designed to govern the application of lethal force by an autonomous system to ensure that it conforms with International Law.

   Arkin, R.C., "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part II: Formalization for Ethical Control", *Proc. 1st Conference on Artificial General Intelligence*, Memphis, TN, March 2008.

3. **Design of architectural principles:**
   We then provided the representational requirements, architectural design criteria and recommendations to design and construct an autonomous robotic system architecture capable of the ethical use of lethal force.

   Arkin, R.C., "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture - Part III: Representational and Architectural Considerations", *Proceedings of Technology in Wartime Conference*, Palo Alto, CA, January 2008.

4. **Implementation of ethical governor:**
   In order to evaluate the feasibility of the ethical governor implementation, a series of test scenarios were developed within the *MissionLab* simulation environment. A variety of situations were presented to an autonomous fixed-wing UAV in which the ethical use of lethal force must be ensured.

   Arkin, R.C., Ulam, P., and Duncan, B., "An Ethical Governor for Constraining Lethal Action in an Autonomous System", GVU Technical Report GIT-GVU-09-02, GVU Center, Georgia Institute of Technology, 2009.

5. **Implementation of responsibility advisor:**
   An ethical permission responsibility advisor was prototyped and demonstrated in a manner fully consistent with the overarching architectural principles developed earlier.

   Arkin, R.C., Wagner, A., and Duncan, B., "Responsibility and Lethality for Unmanned Systems: Ethical Pre-mission Responsibility Advisement", *Proc. 2009 IEEE Workshop on Roboethics*,

Kobe JP, May 2009.

6. **Implementation of moral emotions:**
   Using a cognitive model of guilt we have implemented it computationally and created a proof of concept demonstration in a military context, demonstrating its utility for altering behavior based on emotional state.

   Arkin, R.C., and Ulam, P., "An Ethical Adaptor: Behavioral Modification Derived from Moral Emotions", GVU Technical Report GIT-GVU-09-04, GVU Center, Georgia Institute of Technology, 2009.

Suitable ethical test scenarios have been completed and implemented as a prototype within MissionLab, our laboratory's mission specification software system. Numerous videos document these results:

- Operator interface for the ethical governor:
  ftp://ftp.cc.gatech.edu/pub/groups/robot/videos/PTF_Interface_Final_Largev3.mpg

- Demonstration of the Ethical Responsibility Advisor:
  http://www.cc.gatech.edu/ai/robot-lab/ethics/res-advisor.mpg

- Demonstration of the Ethical Governor:
  ftp://ftp.cc.gatech.edu/pub/groups/robot/videos/ethics_governor_final_largev3.mpg

- Demonstration of the Ethical Adaptor (Guilt mechanism)
  ftp://ftp.cc.gatech.edu/pub/groups/robot/videos/guilt_movie_v3.mpg

The full discourse on this subject is available in

Arkin, R.C., *Governing Lethal Behavior in Autonomous Systems*, Chapman-Hall, 2009.