1. REPORT DATE (DD-MM-YYYY)  21-04-2010
3. DATES COVERED (From - To) September 2009 - April 2010
4. TITLE AND SUBTITLE Robot Wars: An Ethical Way-Ahead
5a. CONTRACT NUMBER N/A
5b. GRANT NUMBER N/A
5c. PROGRAM ELEMENT NUMBER N/A
5d. PROJECT NUMBER N/A
5e. TASK NUMBER N/A
5f. WORK UNIT NUMBER N/A
6. AUTHOR(S) LCDR Gabriel Cavazos, USN
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) USMC Command and Staff College Marine Corps University 2076 South Street Quantico, VA 22134-5068
8. PERFORMING ORGANIZATION REPORT NUMBER N/A
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A
10. SPONSOR/MONITOR'S ACRONYM(S) N/A
11. SPONSORING/MONITORING AGENCY REPORT NUMBER N/A
12. DISTRIBUTION AVAILABILITY STATEMENT Unlimited
13. SUPPLEMENTARY NOTES N/A
14. ABSTRACT Technology alters the nature of warfare and has consistently challenged the accepted jus in bello principles of warfare. Advancements in robotic technology will soon provide governments with fully autonomous systems that their militaries can employ in combat. This research paper reviews the foundations of today's Laws of Armed Conflict (LOAC) and highlights the development of artificial intelligence and the challenges associated with developing ethical decision-making models for robots. This paper also examines the technological and ethical benefits associated with employing fully autonomous systems in combat. After exploring these topics, this paper examines a potential concept of operations to permit the employment of fully autonomous combat systems while adhering to the LOAC and applicable rules of engagement. After careful analysis, this paper maintains that the U.S. Navy's proposal is correct by claiming that fully autonomous robots can be used to target enemy objects, but they cannot ethically be used to target enemy combatants.
15. SUBJECT TERMS Autonomous combat systems, Laws of Armed Conflict
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER OF PAGES
a. REPORT Unclass UU 20
b. ABSTRACT Unclass

c. THIS PAGE Unclass 19. a. NAME OF RESPONSIBLE PERSON Marine Corps University / Command and Staff College
19b. TELEPHONE NUMBER (Include area code) (703) 784-3330 (Admin Office)

Standard Form 298 (Rev. 8-98) Prescribed by ANSI-Std Z39-18
Robot Wars: An Ethical Way-Ahead

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OR THE DEGREE OF
MASTER OF MILITARY STUDIES

LCDR Gabriel B. Cavazos, USN

AY 09-10

Mentor and Oral Defense Committee Member: Dr. Rebecca Johnson
Approved: 
Date: 04/14/10

Oral Defense Committee Member: Dr. Francis H. Marlo
Approved: 
Date: 14 April 2010
Executive Summary

Title: Robot Wars: An Ethical Way-Ahead

Author: LCDR Gabriel B. Cavazos, United States Navy

Thesis: Militaries can employ fully autonomous robots in combat situations if these systems engage enemy objects only. Governments could morally, and legally, justify any unfortunate civilian casualties incurred while engaging a valid enemy object since these casualties would be completely unintended.

Discussion: Technology alters the nature of warfare and has consistently challenged the accepted jus in bello principles of warfare. Advancements in robot technology will soon provide governments with fully autonomous systems that their militaries could employ in combat. This research paper reviews the foundations of today’s Laws of Armed Conflict (LOAC) and highlights the development of artificial intelligence and the challenges associated with developing ethical decision-making models for robots. This paper also examines the technological and ethical benefits associated with employing fully autonomous systems in combat. After exploring these topics, this paper examines a potential concept of operations to permit the employment fully autonomous combat systems while adhering to the LOAC.

Conclusion: It is imperative that any newly developed weapons systems adhere to the jus in bello principles of discrimination and proportionality. After careful analysis, this paper maintains that the U.S. Navy’s proposal is correct by claiming that militaries can use fully autonomous robots target enemy objects, but militaries cannot ethically use these systems to target enemy combatants.
DISCLAIMER

THE OPINIONS AND CONCLUSIONS EXPRESSED HERIN ARE THOSE OF THE INDIVIDUAL STUDENT AUTHOR AND DONOT NECESSARILY REPRESENT THE VIEWS OF EITHER THE MARINE CORPS COMMAND AND STAFF COLLEGE OR ANY OTHER GOVERNMENTAL AGENCY. REFERENCES TO THIS STUDY SHOULD INCLUDE THE FOREGOING STATEMENT.

QUOTATION FROM, ABSTRACTION FROM, OR REPRODUCTION OF ALL OR ANY PART OF THIS DOCUMENT IS PERMITTED PROVIDED PROPER ACKNOWLEDGEMENT IS MADE.
Preface

I became interested in the rapid development and integration of autonomous combat systems given their increased involvement in recent conflicts. This research paper allowed me the opportunity to learn about the history of the Laws of Armed conflict and advancements in robotic technology, especially artificial intelligence and the complex programming required to allow robots to operate with full autonomy. Given the effectiveness of autonomous systems in combat, it is likely that militaries will employ these systems more commonly in future conflicts. I wanted to learn if and how man could ensure that the employment of fully autonomous systems in combat would be morally acceptable.

I would like to thank my research advisor, Dr. Rebecca Johnson, for her wonderful perspectives and insights that I integrated into my paper. I would also like to thank her for her time and patience while assisting me with this project.

More importantly, I would like to thank my loving wife, Tasha, for her understanding and support during the many hours I spent at my computer, and my son, Gabriel, for bringing a smile to my face every time he came downstairs to see how I was progressing with my paper.
Abstract

Technology alters the nature of warfare and has consistently challenged the accepted *jus in bello* principles of warfare. Advancements in robotic technology will soon provide governments with fully autonomous systems that their militaries can employ in combat. This research paper reviews the foundations of today’s Laws of Armed Conflict (LOAC) and highlights the development of artificial intelligence and the challenges associated with developing ethical decision-making models for robots. This paper also examines the technological and ethical benefits associated with employing fully autonomous systems in combat. After exploring these topics, this paper examines a potential concept of operations to permit the employment of fully autonomous combat systems while adhering to the LOAC and applicable rules of engagement. After careful analysis, this paper maintains that the U.S. Navy’s proposal is correct by claiming that fully autonomous robots can be used to target enemy objects, but they cannot ethically be used to target enemy combatants.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCLAIMER</td>
<td>i</td>
</tr>
<tr>
<td>PREFACE</td>
<td>ii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>JUST WAR THEORY</td>
<td>3</td>
</tr>
<tr>
<td>THE ROBOTIC REVOLUTION</td>
<td>6</td>
</tr>
<tr>
<td>MAKING THE CASE FOR FULLY AUTONOMOUS COMBAT SYSTEMS</td>
<td></td>
</tr>
<tr>
<td>TECHNOLOGICAL ADVANTAGES</td>
<td>9</td>
</tr>
<tr>
<td>ETHICAL ADVANTAGES</td>
<td>10</td>
</tr>
<tr>
<td>HERE TO STAY</td>
<td>13</td>
</tr>
<tr>
<td>AN ETHICAL WAY-AHEAD</td>
<td>15</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>18</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>23</td>
</tr>
</tbody>
</table>
Introduction

"The future ain’t what is used to be.” – Yogi Berra

The rules of warfare, known today as the Laws of Armed Conflict (LOAC), have evolved over time and are largely dependent on what is considered to be moral based on the socio-economic and political factors of the era. Technological improvements in weapons consistently challenge these accepted rules of war and alter the ways states fight wars. For example, the employment of a new, mechanized, more lethal weapon, the crossbow, prompted Pope Innocent II to forbid its use against Christians. These technological advances also progress at an exponential rate. A person living during the medieval times would not recognize this progression due to the few technological changes that occurred during their lifetime; however, the number of new weapons developed during the 20th Century greatly changed the character of warfare. During World War I and II, unrestricted submarine warfare, in which a submarine attacked a merchant vessel without warning, appalled the international community because this type of warfare was considered immoral since it did not adhere to the accepted LOAC. The development and application of chemical weapons in WWI and nuclear weapons in WWII each redefined the LOAC.

This adaptation continues today. The character of contemporary war today has made warfare more and less personal at the same time; it is more personal due to the interactive nature of counterinsurgency operations and less so due to the technology’s ability to project and direct firepower from greater distances, which increases the physical distance among combatants. For example, today’s “point and click” methods of warfare allow operators in Creech AFB, Nevada, to engage hostile targets, both humans and objects, in Afghanistan. Taken to the extreme, employing fully autonomous combat systems will completely dissociate the soldier physically
from the combat experience and greatly reduce the risk incurred with military action, which
could greatly influence a country’s decision to employ the military arm of politics.

Today, governments invest more to develop autonomous unmanned systems that their
militaries can employ in combat. Increased autonomy reduces manning requirements, which
provides a cost savings benefit, and decreases the risk of human casualties to the state employing
the autonomous systems. It is imperative, however, to understand if the employment of such
systems adheres to today’s LOAC. Also, it is critical to understand the implications employment
of these systems will have on the nature of warfare. If autonomous combat robots cannot adhere
to the LOAC, then governments should either:

1) Cease to pursue autonomous combat systems, which would save possibly billions of
dollars,
2) Design autonomous systems to operate and target enemy forces independently while
   maintaining a human operator in control of weapon release authority, or
3) Program autonomous systems to engage only valid enemy objects to fulfill the
   requirements of the LOAC; militaries could not use autonomous systems to attack
   combatants.

The international community must consider carefully these options prior to employing fully
autonomous systems in combat.

Militaries can employ fully autonomous robots in combat situations if these systems
target enemy objects only. Governments could morally and legally justify any unfortunate
civilian casualties incurred while engaging a valid enemy object since these casualties would be
completely unintended. Militaries cannot ethically allow these systems to engage enemy
combatants to avoid the possibility of having a robot committing a war crime and not being able
to hold someone accountable for the robot’s actions.

To assess the feasibility of employing fully autonomous robots in combat situations, this
paper will first review the foundations of today’s Just War Theory (JWT), and will then review
robotic development and current military applications of robots. Finally, this paper will examine a concept of operations (CONOPS) for employing autonomous robots in combat proposed by the U.S. Navy and then explore the concerns associated with allowing these systems to engage enemy combatants.

**Just War Theory**

"All sects are different, because they come from men; morality is everywhere the same, because it comes from God." — Voltaire

Just War Theory (JWT) evolved after Christianity became the official religion of the Roman Empire. St. Augustine sought to ensure warfare was constrained by Christian values, and as Walzer notes, "the rules actually observed or violated in this or that time and place are necessarily a complex product, mediated by cultural and religious norms, social structures, formal and informal bargaining between belligerent powers, and so on." Consequently, JWT evolved over time due to changes in technology and perceptions of acceptable behavior regarding the use of force to achieve political and humanitarian goals.

The study of ethics examines the principles of right and wrong used to govern human behavior. These principles provide the basis for making moral decisions: JWT seeks to provide the moral foundations to justify the use of force between states and provide limits to armed combat; therefore, establishing the medium between extreme passivism, which condemns war on all grounds, and realism, which views any action as potentially acceptable during times of war. Just war thinking provides the foundation for the LOAC and informs the constraints warfighters accept when following rules of engagement (ROE).

There are two fundamental components to JWT: *jus ad bellum*, the justice of war, and *jus in bello*, justice in war. *Jus ad bellum* provides tenants for countries to justify going to war, and *jus in bello* outlines appropriate conduct for fighting once war has begun. These two
components of the JWT are distinct and must be analyzed separately; a country that is justified in
going to war may fight unjustly, and conversely, a country that lacks a just cause for waging war
may fight justly.  

Two essential principles establish the foundation for *jus in bello*: discrimination and
proportionality. The principle of discrimination addresses those individuals who combatants
may rightfully target and develops from a commitment to preserving non-combatant immunity.
Discrimination requires militaries to segregate populations into two categories: combatants,
those who militaries may target justly, and non-combatants, those who do not actively engage in
combat and who militaries consider immune from intentional attack. More than a matter of
tradition, international law incorporates the concept of discrimination, as outlined in the Hague
Conventions of 1907, the Geneva Conventions of 1949, and in the 1977 Protocols of the Geneva
Conventions. For example, Article 48 of Protocol I, Part IV, Section I, Chapter I of the 1977
Protocols of the Geneva Conventions states: “In order to ensure respect for and protection of the
civilian population and civilian objects, the Parties to the conflict shall at all times distinguish
between the civilian population and combatants and between civilian objects and military
objectives and accordingly shall direct their operations only against military objectives.”

The international community deems weapons and tactics that cannot fulfill this principle, such as
anti-personnel mines and weapons of mass destruction, unacceptable.

The principle of proportionality seeks to minimize the brutality of warfare by placing
moral limitations on the use of force in combat; that is, the means used in battle must not cause
more harm than is warranted by the military gains the act was intended to achieve. This
principle seeks to prevent escalation of force by encouraging the minimum use of force to
achieve the desired end state. Decision-makers must estimate the amount of force and types of
weapons necessary to subdue the enemy. Johnson elaborates this point further stating, “one may not maim the unjust opponent if it is possible to disarm him without doing so; one may not kill the opponent if it is possible to secure the desired end by only injuring him.” This principle is also fundamental to the laws of war as outlined by the Hague and Geneva Conventions and its Protocols. Article 35, Basic Rules, of Protocol I, Part IV, Section I, Chapter I of the 1977 Protocols of the Geneva Conventions deserves quotation in full to give the sense of the degree to which armed forces are expected to limit the means by which they wage war:

1. In any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited.
2. It is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering.
3. It is prohibited to employ methods or means of warfare which are intended, or may be expected, to cause widespread, long-term and severe damage to the natural environment.

Simply put, armed forces cannot do anything they deem necessary to win the war nor attempt to justify their actions by claiming an act was a military necessity. The debate continues regarding whether Allied Forces were justified in conducting the fire-bombings of Dresden and Tokyo and the nuclear attacks on Hiroshima and Nagasaki.

Using the two principles of *jus in bello*, it is evident that what is morally acceptable in war is the appropriate use of force against legitimate military targets. Unfortunately, it is unlikely that militaries can completely exclude from the tragedies of war unless adversaries wage war in remote locations, such as on the high seas or in the desert. The concept of “double effect” provides a moral means of “reconciling the absolute prohibition against attacking non-combatants with the legitimate conduct of military activity.” This concept maintains that acts that cause harm to non-combatants are considered acceptable provided the act is a legitimate act of war against a valid military target and the effects to non-combatants are completely unintentional.
The LOAC have solid foundations in JWT and provide guidelines to make relative moral decisions in order to justify and limit the brutality of war. Violations of the laws of war constitute war crimes. Service members make great efforts to ensure they honor the tenants of proportionality and discrimination. These efforts, however, can be challenging, especially in conflicts where the enemy blends in to the population. Any fully autonomous combat system must have the capability to differentiate between combatants and non-combatants and then it must possess the ability to choose an appropriate weapon with which to engage the target. The introduction of fully autonomous systems in combat, while beneficial to improving force protection and reducing manpower requirements in the combat zone, raises ethical question of whether the system can satisfy the *jus in bello* principles. The following section will examine advances in robotics and their current applications in the military.

**The Robotic Evolution**

“Never send a man to do a machine’s job.” — Agent Smith, *The Matrix*

Karel Capek first mentioned the word “robot” in his 1921 play, *R. U. R.* (Rossum’s Universal Robots). In his story, robots capable of thinking for themselves initially serve humans. The robots rebel against the humans, and by the end of the story robots kill off all humans and the robot “race” goes on to rule the world.\(^{14}\) Since then, robots have intrigued humans and robot development has progressed at an exponential rate. Robots serve in many capacities, from searching for injured persons following the 9/11 attacks to engaging hostile targets in the mountains of Afghanistan. Human operators direct many actions of unmanned systems; however, today’s technological advances, especially relating to computing power, allow engineers to develop more complex intelligence systems, which improves the capabilities of and enables more autonomous operation of unmanned systems. This section will first analyze the
difference between robots and machines and then review Moore’s Law, which has contributed
directly to the dramatic improvements in robotics capabilities. This paper will then examine the
issue of autonomy, which refers to the level of independent decision-making capability granted
to a robot and the ability to act on those decisions, to understand why autonomous systems
present a challenge to the LOAC.

A common misconception is that the terms ‘robot’ and ‘machine’ are interchangeable. Merriam-Webster Dictionary defines a machine as “(1): an assemblage of parts that transmit
forces, motion, and energy one to another in a predetermined manner (2): an instrument (as a
lever) designed to transmit or modify the application of power, force, or motion; a mechanically,
electrically, or electronically operated device for performing a task.” Although robots meet these
requirements, a machine must possess more capabilities to be classified a robot. Robots are:
man-made devices built on the “sense-think-act” paradigm and consist of three essential
components: “sensors” used to monitor and detect changes in the surroundings, “processors” or
“artificial intelligence” (AI) that determine how to respond in situations, and “effectors” that
allow the robot to cause changes in its environment based on its decisions. If a machine lacks
any one of these components, then it is not a robot; therefore, robots are machines, but machines
are not necessarily a robot. 15 AI stores the programming that allows a robot to operate
autonomously, and there is a direct correlation between the level of autonomy and required
computing capabilities.

It requires tremendous computing power to enable autonomous operations of unmanned
systems. In 1965, Intel co-founder Gordon Moore, noted that manufacturers could double the
number of transistors placed on a microchip approximately every two years. His prediction
regarding the exponential improvements in computing power, commonly referred to as Moore’s
Law, has been upheld since and has improved to doubling at intervals of nearly eighteen months. If Moore’s Law holds, “it seems only a matter of time until more sophisticated cognitive states, perhaps wildly different from our own, will evolve.” Having the necessary computing power contained within a reasonable size allows engineers to develop more sophisticated, capable artificial intelligence (AI) for robots. This AI leads to the concept of autonomy.

Many vaguely understand the concept of autonomy associated with unmanned systems. With respect to the science of robotics in military applications, autonomy pertains to the relative independence given to a robot and its ability to make decisions and act upon those decisions without human intervention to support the given mission. As Georges notes, “applied to robots, the term [autonomy] merely obscures the real issue, which is about the levels of decision making and goal setting we want to reserve for ourselves versus how much power and control we want to give to them [robots]. At best, autonomy is a matter of degree.” These degrees, or levels, of autonomy vary from human directed operation at the low end to “adaptive” at the high end of the spectrum. Humans interact less with the robot as its level of autonomy increases. If given “adaptive” capabilities, a robot can learn from its experiences.

Current military applications employ robotic systems with varying levels of autonomy. For example, the PackBot, which is a remote controlled robot that militaries can reconfigure for various operations ranging from explosive ordinance disposal to identifying hazardous materials, operates with low levels of autonomy; a human operator directs many of the PackBot’s actions. Conversely, the Global Hawk reconnaissance aircraft operates at the high-end level of autonomy. This aircraft can independently take off once given permission, surveil an area and decide what to take images of, and then return to base and land; a human operator merely directs the aircraft to and from the hangar and the runway.
Making the Case for Fully Autonomous Combat Systems

“*They don’t get hungry. They’re not afraid. They don’t forget their orders. They don’t care if the guy next to them has just been shot. Will they do a better job than humans? Yes.*” – Gordon Johnson of the U.S. Joint Forces Command at the Pentagon Referring to Robot Soldiers

**Technological Advantages**

Fully autonomous combat systems offer many advantages, both tactically and ethically. Tactically, these systems reduce manpower requirements and do not strain lines of communication. Ethically, these systems can improve discrimination via improved target identification and precise, accurate weapons delivery while reducing combat trauma.

Technology has not only improved the lethality of weapons, but it has also granted warfighters greater stand-off range from the enemy. As a result,

in antiquity, when you divided the number of people fighting by the area they would typically cover, on average it would take a Greek hoplite and five hundred of his buddies to cover an area the size of a football field...By the time of the American Civil War, weapons had gained such power, distance, and lethality that roughly twenty soldiers would fight in that same space of a football field. By World War I, it was just two soldiers in that football field. By World War II, a single soldier occupied roughly five football fields to himself. In Iraq in 2008, the ratio of personnel to territory was roughly 780 football fields per one U.S. soldier.

Today, operators can simply “point and click” to engage targets via unmanned platforms from half a world away and some weapons such as the Tomahawk Land Attack Missile (TLAM) engage pre-assigned targets. The operator launches the TLAM without knowledge of what target the missile will engage.

The United States and other nations have developed and employ many systems with varying autonomy. The Aegis weapons systems carried onboard U.S. warships has an “auto” mode, which enables the system to detect-track-classify-engage without human intervention. Global Hawk, scheduled to replace the U-2, operates autonomously at high altitudes for more
than thirty hours. The system navigates to preplanned waypoints via Global Positioning System (GPS) and executes its assigned mission independently. The human operator merely clicks the mouse to have the unit take off and land. Perhaps the most revealing demonstration of autonomous systems' capabilities occurred in February 2005 when two Boeing Joint Unmanned Combat Air Systems (X-45A) successfully completed a simulated combat mission. When presented with multiple threats, the two platforms decided which would attack based on position with respect to the target, weapons load, and fuel remaining. These examples illustrate the usefulness of autonomous systems in combat and prove these systems are worth developing further.

Unmanned systems are becoming more commonplace within the U.S. military. For example, by 2008, the number of unmanned planes outnumbered manned planes by a factor of nearly two-to-one. A majority of unmanned systems currently employed still have a human operator “in the loop” who selects and engages targets; this distancing illustrates man’s increasing aversion to physically engaging the enemy. Fully autonomous systems are the next logical step.

**Ethical Advantages**

Despite man’s instinctual advantage in combat, humans have many faults and physical limitations, which combat magnifies, that robots lack. Man’s shortcomings include, but are not limited to, the need for food, water, and sleep. Additionally, extreme environmental conditions negatively affect human performance in general. Conversely, robots can operate in remote, harsh environments for extended periods without suffering negative effects on performance and, perhaps more importantly, without straining lines of communication. With respect to aviation, unmanned systems can fly faster and turn harder than manned systems since man’s ability to
overcome G-forces limits the maneuverability of manned platforms. Quite simply, robots can execute tasks faster and with more accuracy than humans. One official at Defense Advanced Research Projects Agency (DARPA) noted that humans are quickly becoming the weakest link in defense systems. Unmanned systems provide a solution to overcoming man’s limitations.\(^{24}\)

This section will explore deficiencies in human behavior that can lead to immoral behavior in combat and illustrate how robots can mitigate this human shortfall. Next, this section will review why the employment of robots in combat is a merely a matter of time. Lastly, this section will examine the advantages of employing robots, in lieu of humans, in combat environments.

With respect to adhering to the LOAC, robots are not susceptible to human short-comings such as giving in to unlawful orders and acting immorally in the battlefield under the stress of combat. Prior to a conference in February 2008, Dr. Arkin, a leader in the field of the developing ethical AI and the ethics of employing autonomous robots, expressed his confidence that militaries can justly employ fully autonomous robots in the battle space. The trends in human behavior on the battlefield regarding adherence to legal and ethical requirements are suspect as evident in the following report.\(^{25}\)

According to a recent report from the Surgeon General’s Office that assessed the mental health and battlefield ethics of Soldiers and Marines participating in Operation Iraqi Freedom, “only 47% of Soldiers and only 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 (41% and 44%, respectively) or to obtain important information about insurgents (36% and 39%, respectively).”\(^{26}\) Regarding battlefield ethics violations, the report cited that:

the most likely battlefield ethics violation that Soldiers and Marines would report included a unit member injuring or killing an innocent non-combatant, with 55% of
Soldiers agreeing that they would report a unit member and 40% of Marines agreeing that they would report a fellow Marine. Soldiers and Marines were least likely to report a unit member for unnecessarily destroying or damaging private property, with 43% of Soldiers indicating that they would report a unit member and 30% of Marines indicating that they would do so. 27

Additionally, “Soldiers and Marines who had high levels of anger were twice as likely to engage in unethical behaviors on the battlefield” compared to those who had low levels. 28 Similarly, “Soldiers who screened positive for a mental health problem (anxiety, depression or acute stress) were twice as likely to engage in unethical behavior compared to those Soldiers who did not screen positive.” 29

Employing autonomous systems can help alleviate these troubling findings. Robots would not be susceptible to emotions, such as frustration, anger, and resentment, which could lead to actions resulting in possible war crimes. Integrating a robot that is programmed to adhere, at all times, to the LOAC and applicable ROE with human military units may also promote better adherence to these moral and legal foundations, which outline acceptable behavior during war. If militaries equip robots video cameras that continuously record events, then Soldiers would also be less likely to commit LOAC or ROE violations knowing that the robot is recording their actions.

While one motivation for adopting fully autonomous systems is to reduce LOAC and ROE violations due to combat stress, another is to improve force protection. America’s refusal to sacrifice American lives in combat has contributed to the U.S. Government’s insistence on developing and employing unmanned systems. Major General Robert Scales maintains that the new age of warfare is one in which “dead soldiers are America’s most vulnerable center of gravity.” 30 Senator Warner shares this view:

“When you look at the history of casualties, beginning with almost half a million killed in WWII, over 35,000 killed in Korea, and more than 50,000 killed in Vietnam, and zero
combat deaths in Kosovo, in my judgment this country will never again permit the armed forces to be engaged in conflicts which inflict the level of casualties we have seen historically...The driving force is the culture in our country today, which says, 'Hey! If our soldiers want to go to war, so be it. But don’t let any of them get hurt.'" 31

Given these sentiments, it is easy to understand why the U.S. Government is investing so much in developing autonomous unmanned systems.

A desire to reduce manpower requirements is a third driver toward fully autonomous systems. As noted by VADM McCullough, former deputy chief of Naval Operations for Integration of Capabilities and Resources (N8), high end autonomy is the "key to getting dollar savings out of this [unmanned systems]." 32 Another compelling reason to employ high-end autonomous systems is because it would be difficult for one operator to process the inordinate amount of data from and then direct the actions of one, much less a number of robots, such as an unmanned air systems squadron or a squad of ground combat systems. Granting robots with full autonomy allows the systems to maintain their own situational awareness and coordinate movement over the entire battle space.

Here to Stay

Robotics and war is not a matter of "if," but rather, "when." The 2001 National Defense Authorization Act, Public Law 106-398, Section 220 mandates that by 2010, one-third of the aircraft in the operational deep strike zone should be unmanned, and by 2015, one-third of the Army’s Future Combat System operational ground combat vehicles should be unmanned. 34 Additionally, Congress ordered the Pentagon to show a "preference for joint unmanned systems in acquisition programs for new systems, including a requirement under any such program for the development of a manned system for a certification that an unmanned system is incapable of meeting program requirements." 35 The Department of Defense (DoD) Unmanned Systems Integrated Roadmap extends to 2034 and asserts that "unmanned systems will continue to have a
central role in meeting our country’s diverse security needs, especially in the War on 
Terrorism.36 These three directives illustrate the military’s commitment to developing 
unmanned technology and intent to use it in the battle space. Given this intent, government and 
military officials must demand and engineers must develop robots that have the capacity to 
adhere to ethical considerations defined by the LOAC. The Department of Defense policy 
regarding the weapons systems approval process directs that weapons procurement and 
acquisition “shall be consistent with all applicable domestic law and treaties and international 
agreements.”37 The acquisition process, therefore, must consider the following three factors, 
which address the *jus in bello* tenants of proportionality and discrimination:

1. The weapon must not cause suffering that is needless, superfluous, or disproportionate 
to the military objective,
2. The weapon must be able to be controlled so that it can be directed toward a lawful 
target, and
3. The weapon must not violate an existing treaty or domestic law prohibiting the 
weapon’s use.38

The U.S. Army addresses this need in its 2007 Small Business Innovation Research (SBIR) 
Proposal, which states:

> Armed UMS are beginning to be fielded in the current battlespace, and will be extremely 
> common in the Future Force Battlespace...This will lead directly to the need for the 
systems to be able to operate autonomously for extended periods, and also to be able to 
collaboratively engage hostile targets within specified rules of engagement...Fully 
autonomous engagement without human intervention should also be considered, under 
user defined conditions, as should both lethal and non-lethal engagement and effects 
delivery means.39

It is important to note that this vision incorporates the *jus in bello* principles of discrimination 
and proportionality as evident by the verbiage “engage hostile targets,” and “both lethal and non-
lethal engagement,” respectively. One proposal presented by the U.S. Navy fulfills these *jus in 
bello* tenants and will be examined in the following section.
An Ethical Way Ahead

“If we continue to develop our technology without wisdom or prudence, our servant may prove to be our executioner.” – General Omar Bradley, USA

Autonomous systems, like human soldiers, must adhere to the LOAC; that is, they must possess the ability to discriminate between combatants and non-combatants and then they must use appropriate force. It is relatively easy to provide a robot with a program that delineates what type of weapon to use against various targets to satisfy the principle of proportionality. The *jus in bello* principle of discrimination, however, is considerably more difficult to translate into 1’s and 0’s. While a robot may differentiate between a cruise liner and a warship relatively easily based on acoustic and electronic signatures, it is much more difficult to provide a robot with the AI to differentiate between a group of individuals who are angry and protesting and a group of individuals who are intent on causing harm. In irregular warfare, combatants frequently dress like the civilian population; therefore, it is often difficult to differentiate between a civilian and a combatant when not engaged in combat. These insurgents do not adhere to the convention that directs belligerents to “have a fixed distinctive emblem recognizable at a distance.” It is also necessary to prevent human casualties that the enemy can publicize and exploit to tarnish America’s reputation and thwart America’s national strategic and operational objectives. This section will review a U.S. Navy proposal regarding the development and employment of autonomous combat systems with respect to the laws of war, which will likely provide the template for employing autonomous systems in combat. Next, this section will explore the challenges associated with developing AI capable of emulating morally acceptable behavior. Finally, this section will address the greatest obstacle to employing fully autonomous systems in combat.
The U.S. Navy proposal provides a practical means to allow militaries to use fully autonomous systems in combat immediately. NAVSEA director, Mr. John Canning, and his team developed an acceptable targeting subset for autonomous systems. Given the inherent difficulty associated with differentiating between belligerents in today’s counterinsurgency conflicts, NAVSEA proposes to allow robots to engage only military objects and exclude robots from targeting people altogether; thus, robots could target “either the bow or the arrow, but not the archer.” This CONOPS removes much of the ambiguity associated with the LOAC and ROE that pertains to engaging combatants, which requires soldiers to assess the belligerent’s intent. Governments would consider any human casualties morally acceptable under the principle of “Double Effect” since they occurred while targeting a legitimate military object and were completely unintentional. This CONOPS also addresses the *jus in bello* tenant of proportionality by recommending arming autonomous combat robots with non-lethal weapons that could be used against belligerents to convince them to abandon their weapons. A human operator would assume weapons release authority if the autonomous system needed to engage a human target. This CONOPS adheres to today’s JWT, is executable, and subscribes to the legal precedence of weapons systems currently in use that target the “bow or the arrow, but not the archer.” For example, the Aegis weapons and Patriot missile systems both have “auto” modes that allow for independent detection, tracking, classification, and target engagement, without operator action. This CONOPS allows for employment of autonomous combat robots given today’s technology; however, technology will continue to improve, thus, providing more capable sensors and AI that will eventually allow robots to discriminate between combatants and non-combatants.
Various AI programs, encoded with ethical guidelines, exist to generate morally acceptable behavior in autonomous systems. As McLaren notes, however, the realm of ethical reasoning is difficult for scientists to model and computers to calculate because of the following:

1. The rules that must be adhered to (i.e. laws, codes, principles) are provided at a highly conceptual, abstract level; conditions, premises, or clauses that are not precise or that cover a wide range of specific facts, or are highly subject to interpretation and may even have different meanings in different contexts.

2. The actions prescribed by the given rules, i.e., the rules’ conclusions, may also be abstract. Thus, even if one is able to determine that a particular rule applies to a given fact situation, the rule’s consequent recommendation may be difficult to execute because it is highly conceptual or vague.

3. Abstract rules often conflict with one another in particular situations with no deductive or formal means of arbitrating such conflicts. That is, more than one rule may appear to apply to a given fact situation, but neither the abstract rules nor the general knowledge of the domain provide clear resolution.\textsuperscript{44}

It is important to note, however, that the LOAC and applicable ROE remove much ambiguity by bounding acceptable behavior in combat “when compared to ungoverned solutions for autonomous robots.”\textsuperscript{45} It would be easier to develop an AI program based on the LOAC and applicable ROE, which would allow autonomous robots employed in combat situations to perform in a manner considered more ethical than humans. Although incapable of experiencing or exhibiting emotions such as guilt, compassion, and mercy, Arkin notes that “humanity is legislated into the Laws of War, and as such if they are followed, the robot will exercise restraint consistent with societal norms.”\textsuperscript{46} In his book \textit{Governing Lethal Behavior in Autonomous Robots}, Arkin goes in to great detail regarding the ethical models used to develop the prototype for his autonomous robotic system architecture, which is beyond the scope of this paper; however, Arkin maintains he can develop morally encoded AI that will allow robots to perform more ethically in battle compared to humans.\textsuperscript{49}
The greatest hindrance to employing fully autonomous robots in combat that are allowed to engage enemy combatants is overcoming the tenant of the LOAC that someone is able to be held accountable for the killing of non-combatants, which raises the question of accountability. Arkin introduces the concept of a "responsibility advisor," who is responsible for informing the deployer of a fully autonomous system of the possible consequences of deploying the system. This concept, however, does not overcome the dilemma of holding a human operator responsible for an autonomous robot's actions, especially if the robot is given "adaptive" AI, which would allow the robot to learn and perhaps modify its ethical algorithms. Arkin acknowledges this as a "daunting problem remaining" to be overcome. Until this issue is resolved, society will consider it unethical to allow autonomous robots to engage human targets.

**Conclusion**

"The use of long distance, remote control weapons, or weapons connected to sensors positioned in the field, leads to the automation of the battlefield in which the soldier plays an increasingly less important role. The counter-measures developed as a result of this evolution, in particular electronic jamming (or interference), exacerbates the indiscriminate character of combat. In short, all predictions agree that if man does not master technology, but allows it to master him, he will be destroyed by technology."

— International Committee of the Red Cross

Man has engaged in battle for various reasons since the beginning of time. History illustrates that man has also sought to justify aggression yet limit the destruction and brutality associated with war. Morality, which is dependent on variables such as social, political, and cultural norms associated with a particular time period, has provided the foundation for these limits. These limitations evolved into today's codified LOAC. *Jus ad bellum* provides guidance
for states to justify the use of force against other states, and *jus in bello* seeks to limit the
destruction caused in war. It is imperative that any newly developed weapons systems adhere to
the *jus in bello* principles of discrimination and proportionality.

Robots will quickly assume additional roles on the battlefield. They are not constrained
by food, water, sleep, and robots can perform tasks more quickly and with better accuracy than
humans. Perhaps more importantly, robots can behave more consistently with the LOAC and
would not be subject to man’s mental and emotional shortcomings that result in unethical
behavior on the battlefield. Feelings such as anger, rage, and fear that can cause humans to
behave unethically in combat and violate the LOAC cannot influence a robot’s moral-emulating
programming; therefore, the robot is unlikely deviate from the LOAC and applicable ROE due to
emotional stresses. Employing fully autonomous combat systems reduces manpower
requirements, which is essential to fully reaping the cost-benefits from unmanned systems.
America’s aversion to suffering casualties in combat has also contributed to Congress’ mandate
that the Pentagon explore autonomous systems for many future projects.

Autonomous unmanned systems employed in the battle space must adhere to the LOAC.
Engineers can easily program robots with the AI necessary to adhere to the *jus in bello* principle
of proportionality; however, engineers continue their efforts to develop a moral decision-making
program that can satisfy the tenant of discrimination. This is partly due to the need for improved
sensors, but primarily because belligerents engaged in today’s insurgencies engage in guerrilla
tactics and dress like the civilian population. The NAVSEA CONOPS provides targeting subsets
for autonomous combat systems adheres to the LOAC and alleviates much of the gray area
associated with ROE by intentionally limiting autonomous systems to engaging only legitimate
military objects. This targeting model allows militaries to employ autonomous systems
immediately in combat situations. Technological improvements will continue to progress and eventually enable robots to differentiate between combatants and non-combatants. It would be unethical, however, to allow these systems to autonomously engage human targets since no one would be able to be held accountable if the autonomous agent inadvertently violated the LOAC and ROE. A human operator must assume weapons release authority if the military uses the autonomous platform to engage a human target.

Advancements in weapons, from the crossbow to the Tomahawk missile, have produced more lethal weapons that have provided opponents with greater stand-off ranges. This distancing makes war less personal and, thus, soldiers become immune to the realities of war; therefore, technology, especially employment of autonomous robots, significantly limits the risk of human casualties to the nation possessing it and its allies, which is a significant factor when making the calculations to go to war.

While it is possible for robots to engage in lethal force in combat in a manner consistent with *jus in bello* restraints, their use may lower the threshold for *jus ad bellum* limitations on the use of force. Questions that require further study prior to the wide scale use of autonomous systems include: Will the use of autonomous combat systems make the countries employing them more apt to use force to achieve political objectives? What happens when robots evolve to the point where they are sentient? Do they deserve rights? Are we morally obligated to treat them with certain ethical standards? It is imperative that man consider these implications associated with employing autonomous systems prior to committing to such a course of action that will surely change the character of warfare.

---

2 Michael Walzer, 43.


“Protocols Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977, Part IV: Civilian Population, Section I: General protection against effects of hostilities, Chapter I: Basic rule and field of application, Article 48: Basic Rule” (accessed December 30, 2009). Although the United States has not signed the Protocols to the Geneva Conventions, the United States does accept the Protocols as customary law.


Johnson, 198.


Walzer, 131.

Johnson, 130.

Walzer, 153.

Johnson, 130. Michael Walzer, in Just and Unjust Wars, maintains that the principle of “double effect” is insufficient because it does not provide sufficient protection to non-combatants who may be considered as merely acceptable collateral damage. He, therefore, proposes the principle of “Double Intention,” which adds a fourth caveat to the principle of “double effect.” His fourth caveat maintains that a commander should make every effort to minimize collateral damage despite the increased risk to the combatants conducting the operation. Walzer’s principle of “double intention” is not recognized in the Judge Advocates General’s Law of War Desktop Workbook, which establishes the legal constraints for U.S. armed forces. Walzer’s principle, therefore, will not be discussed further in this paper.


Singer, 67.


Georges, 53.


Singer, 100.


Singer, 37.

Singer, 64.


Surgeon General’s Office, Mental Health Advisory Team (MHAT) IV Operation Iraqi Freedom 05-07, Final Report, November 17, 2006, 35.

Surgeon General’s Office, 36.

Surgeon General’s Office, 38.
29 Surgeon General’s Office, 38.
30 Singer, 59.
31 Singer, 60.
38 Canning, 8.
40 The Avalon Project at Yale Law School: Laws of War: Laws and Customs of War on Land (Hague IV); October 18, 1907,” (accessed December 28, 2009).
42 Canning, 1-32.
43 Canning, 1-32.
45 Arkin, 38.
46 Arkin, 44.
50 Arkin, 144.
51 Arkin, 211.
Bibliography


Canning, John S. "A Concept of Operations for Armed Autonomous Systems." Chief Engineer, G80 Division, Naval Surface Warefare Center, Dahlgren Division.


