Human vs. Robot Decision Making on the Battlefield: War and Rational Choice Theory

by Todd J. Barry

Abstract

Fifty years after the 1968 Vietnam War's Tet Offensive, old debates are being revived in the United States regarding the rational conduction of war and the appropriateness of whether and when to send or withdraw troops, while new debates are emerging about the U.S. military's use of artificial intelligence (AI). In this paper, I present a political-economic model of war using marginal concepts with regard to the wars in Vietnam, Iraq, and Afghanistan. I then turn to an examination of the rationality of AI robots and their ethical use in war.

Keywords: Tet Offensive, rationality, sunk costs, artificial intelligence, utilitarianism

1. Introduction

In this paper, I present a political-economic model of war using marginal concepts with regard to the wars in Vietnam, Iraq, and Afghanistan. I then turn to an examination of the rationality of artificially intelligent (AI) robots and their ethical use in war. Although machine intelligence can often surpass human rationality as demonstrated by a host of simple experiments, machines lack emotional decision making, which raises a wide range of ethical questions about their employment in warfare. I begin by proposing a model of rational decision making with consideration of statistics from recent military conflicts in order to evaluate the process of troop surges and withdrawal. I then conclude with a consideration of some of the potential ethical dilemmas posed by military robots and the need for developing policies in advance of their increasing use in future military conflicts.

The economic principle of sunk costs contends that rational decisions should be made by comparing future marginal benefits and losses without regard to prior costs. Sunk costs, however, assume that one has the resources to continue a policy, that utility or happiness remains constant, and that there are no substitutes for the resources or opportunity costs (alternative uses). This principle emerged when neo-classical microeconomics was developing in Victorian England. Philosophers at that time also held that mankind could be understood as rational and that wealth, along with other factors, could be used as the basis for calculating utility maximization (Bernstein 1998).

In warfare, such sunk costs include monetary expenses, number of casualties, and emotional suffering. Did military and political leaders during the Vietnam War (roughly 1962-1974), the Afghanistan War (2001-present) and the Iraq War (2003-2011) act with

economic rationality? Moreover, could a war fought by AI robots be conducted following principles of economic rationality?

A 2011 study by Schott, Scherer, and Lambert found that people were more inclined to continue to support a war effort after being primed with information that people should "not waste" while reading about the losses in Afghanistan. Psychologists have found that people are prone to overweigh recent evidence over long-term evidence, which is likely the reason why citizens tend to form their views about war based on the most recent military conflict. Daniel Kahneman proposed prospect theory in the 1970s, in which he theorized that we like to keep what we already have and are more likely to gamble, if we choose to, when the choice affects ourselves rather than others. More recent work on this issue by Robert Shiller (2000) suggests that people make riskier decisions when the choice is more vivid, such as when a friend provides a verbal or visual description (Bernstein 1998). In short, people gamble, take risks, and are emotionally impacted by vivid descriptions, whereas robots would not be, unless robots develop an emotional system over time.

2. Method

In what follows, I propose a model for military operations centered on rational choice economic principles. This model assumes winning as the end game of any military operation, whereby winning is defined as the ability of one side to compel the other to surrender while experiencing less relative suffering as defined by average variable costs. Military decisions are deemed rational when they are solely determined by future costs without regard to sunk costs. I applied this model in order to access whether military or political leaders have acted rationally in the Vietnam, Afghanistan, and Iraq conflicts. Following this, I will then consider whether artificially intelligent robots would similarly be capable of rational decision making in a military operation.

The figures below show "fixed costs" as those amount of troops initially expected as casualties, held fairly constant from the initial troop investment. Average variable cost (AVC) is the average of the total variable costs and additional American casualties, and marginal costs are the casualties of incremental troops sent. Marginal costs (MC) and marginal revenues (MR) indicate that as one adds American troops, MC increases, and more troops increasingly perish and foreign troops are killed at a decreasing rate. Overall, a country needs to increase troops to obtain more foreign casualties to be rationally "profitable," which is the difference of the AVC casualties and MC at the MC=MR intersection. The amount needed to "win" varies with each conflict. In economics, "closing shop" would happen if AVC rises about the MC=MR intersection, but there are more considerations in war, such as future resources, political factors, and public morale. One side might thus choose to go beyond marginal measures to win at all costs, such as was the case with U.S. Civil War Generals Grant and Sherman. The model assumes foreign troop levels stay the same; if they increase, then MC would shift right, with AVC shifting up and right.

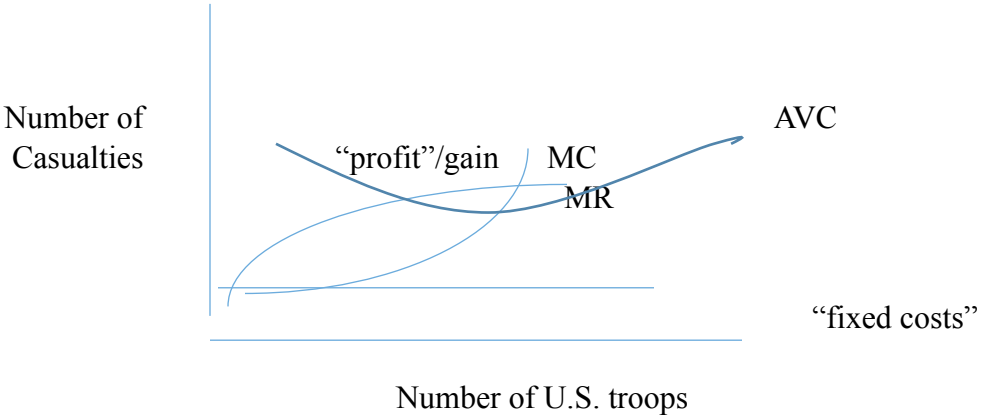

Figure 1: Micro-Economic Rational Model of War (Gains)

The fact that MR = MC can be explained by countries around the world having similarly sized militaries despite vast difference in populations.

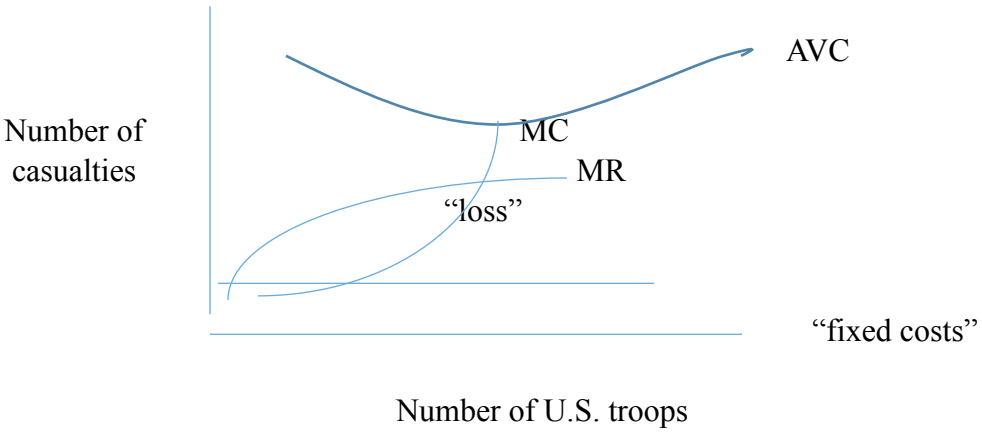

Figure 2: Micro-Economic Rational Model of War (Losses)

3. Empirical Results

The Tet Offensive in 1968 changed the direction of the Vietnam War, resulting in roughly 45,267 Vietnamese deaths (Kolko 1985), and 44,814 allied deaths (16,899 U.S. and 27,915 "host") (Rummel 1997; DCAS 2008). These figures indicate that the North Vietnamese had reached a military capability whereby U.S. troop casualties were the same, marginally, as their own. Following the Tet Offensive, President Richard Nixon initiated a troop drawdown in 1969, consistent with the rational choice model. Since the United States lost 58,000 soldiers and the enemy 444,000 (as estimated by historian Guenter Lewy), one can project that it would require several million additional U.S. troops to "win," based on increasing marginal costs, which the American public was not willing to bear. Consistent with Shiller's previously discussed theories, Americans were also duly affected by the vivid images of the war presented in various media outlets, especially through television news coverage.

In the Iraq War, there were 23,984 Iraqi "enemy" casualties and 19,610 allied casualties (15,196 "host" deaths and 4,424 U.S. deaths). This does not include the 3,771 "friendly deaths," the nearly 100,000 civilian victims, as well as the countless injured (Leigh 2010). These numbers seem to indicate a 1:1 relationship, which, if considered marginally overall, suggest that President George W. Bush's troop "surge" was a rational military decision since America was "winning" insofar as equilibrium was achieved.

At the start of wars or with adeptness over longer time frames, countries can reach "economies of scale," or greater output per lower costs through efficiencies, tactics, strength in numbers, and "winning hearts and minds" of the citizens, which lessens MC and total costs. One example of this is the Powell-Weinberger doctrine of "overwhelming force" used in the Gulf War. This also occurred in Iraq with the "Sunni Awakening," as well as in Afghanistan in 2010. When President Obama ordered a 33,000 increase in 2009, the hike escalated the violence and increased innocent civilian casualties, suggesting that it was not rational in an economic sense but more politically driven (Rafferty 2017).

Certainly, a country cannot deploy infinite troops, which are needed in the domestic workforce, as the famous Lyndon Johnson "guns versus butter" analogy asserts. But common military strategy posits that an offensive force needs 3 times the troops as a defensive force to "win," because defensive forces know the terrain and can make fortifications, affecting the slopes, or elasticities, of the marginal curves. In Vietnam, the marginal curves kept shifting rightward with more foreign troops being trained, stronger morale, and greater weapons resulting in more American casualties.

During the Vietnam War, and more recently as demonstrated by President Donald J. Trump in Afghanistan, political leaders have appealed to the public to garner support for continued military operations, claiming that we cannot let past sacrifices be dishonored. Although one may question the initial legitimacy of the Iraq and Afghanistan Wars, those in favor of abandoning these efforts failed to recognize the principle of sunk costs, in which land already gained could be held and not represent a future cost. For example, in the 1950's Korean War, the United States was able to maintain its position. In contrast, the United States was forced

to fight a "guerrilla" war in Vietnam, where land "controlled" was dispersed across both the North and South, as well as with enemy combatants in the South, making it more difficult to maintain a dominant stronghold.

In the current Afghanistan War, only 17 allied troops were killed in 2017 (iCasualties 2018), which some might argue is a small cost to pay for the successes gained in order to assist an otherwise dangerous and unstable state. Measuring the relative success of a war effort by statistics alone, however, is impossible. Even when going to the movies and losing your ticket, you would still buy another one if MR > MC, just like an army might keep fighting after sustained losses, but it does not necessarily mean that the utility from the movie will be as expected or the outcome of the war will be as a country anticipates.

Military decision making is important to understand in thinking about future conflicts, which will be fought primarily by robots.

4. The Rationality of Military Robots

In September 2017, Russian President Vladimir Putin told a million of his country's schoolchildren via a televised speech, "Whoever becomes the leader in [artificial intelligence] will become the ruler of the world" (Allen 2017). Technology often finds its way to the battlefield before the home front. If this is the case, we are led to consider if AI robots are capable of rational decision making, and if so, what are the implications for future warfare if they should be employed in military operations?

Many simple tests show AI to demonstrate superior rationality over humans. Although some humans might be tricked by an equal outcome game, by favoring the higher dollar value in the question, AI would not be deceived, for example, if asked to choose between 50% of $100 or 25% of $200. Moreover, AI would not be deceived by consideration of sunk costs as demonstrated by computer chess programs that are not emotionally affected by previous moves when deciding future moves. Computers are essentially rule followers, though the analysis here must go deeper.

In warfare, human fighters do not like to see their compatriots' deaths, but will robots share this sort of emotional concern? Rationality can imply, as stated earlier, that we intend to create as much utility for ourselves with the least cost in terms of money, time, energy, social status, or loss of life, as well as countless other factors depending on our personal values. People do not, however, only consider their own utility when deliberating a situation. In fact, we have countless examples of people who will risk their lives for the sake of someone else's welfare.

Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873) proposed the idea of utilitarianism, in which political decisions should be made that result in the greatest happiness for the greatest number of people. The concept of utilitarianism, however, is based on consequences, not on intent, and conflicts with the idea of rights that emerged from the 18th century European Enlightenment. Of course, intent and consequence are often related to one

another as is evident in property rights—if people were constantly stealing each other's property, there would be no foundation for association and chaos would result. Rights and laws protect the marginal distribution of capital (what belongs to whom), without which, there would be no incentive for people to work or fight, thus diminishing overall utility. But under utilitarianism, one must assume that poorer or less fortunate individuals have greater needs, due to diminishing marginal utility. Since the concept applies to overall happiness and not whether someone in particular is happy, the condition maximizing utility can be expressed in the formula:

gain to the less fortunate = loss from lack of motivational productivity/effort.

Wars arise from several reasons including disputes over resources, conflicting values, or perceived power struggles. Wars are possible even within a utilitarian context if the overall good is perceived to be greater than the loss; however, conflict resolution can also achieve this objective. Wars undertaken for the wrong intentions should be reevaluated according to a utilitarian framework, as ending wars abruptly can sometimes create vacuums of power resulting in greater civilian loss.

How these various factors affect robots employed with military decision making will depend on their programming. Many regard AI as rational if it meets the goals set by its programming. In contrast, humans base their utility upon their emotions rather than following the same rule governed logic of a set program. Therefore, could a robot truly be rational if it cannot base its utility maximization decisions on emotional factors?

One should, moreover, keep in mind that all nations and all individuals have different values that might influence any given robot's programming. As Kasparov (2017) points out, in computerized chess programs, each chess figure is programmed with a point score based on medieval standards. For instance, a pawn is worth just one point, while a knight is worth three points. Early chess programs pursued a failed strategy of maximizing points at all costs, resulting in overly aggressive moves. In contrast, more recent computer chess programs are able to avoid negative outcomes. That said, a small change in values results in a completely different algorithm. One should further keep in mind the argument of the economist Kenneth Arrow that, while computers may use numerical values in a consistent and predictable manner, numbers mean something different to each person.

The most dangerous arena where artificial intelligence might be employed is warfare, in which robots will be programmed to kill foreign soldiers or other robots. Will robots simply follow their programming or will they evolve intelligence to question it? Will our answers to their questions conflict with their original programming creating something like cognitive dissonance? Robots may come to question their programmed assumptions when they find *contradictions* in their programs, such as how their own value is to be assessed against other entities, whether human or robotic.

People can often base decisions on emotional factors, such as when they engage in conspicuous consumption or when they are seduced by a politician's rhetoric. Although emotional decision making may seem to be irrational, our animal instincts ensure our

reproduction and survival. If robots were to behave in the same way, will they attempt to survive, whether individually or collectively, at any cost?

If robots were to value life, whether human or robotic, they could employ game theory as initially described by the economist John von Neumann and mathematician John Nash. The robots could mutually conclude to cooperate and thus decide *not* to kill each other for the sake of the "social" good, just as trench warfare fighters did during World War I. Some military robots might even become opposed to the basic principle of war in the manner of a pacifist. Others might recall earlier programming, or communicate their values with other robots, and thus display something akin to a conscience.

Robots will need to be programmed carefully in order to identify who is their enemy. No two robots will be alike since they will all have different experiences from which to adapt in a process of machine nurturance. Both defensive and offensive robots would have distinct programming concerning their opposing goals, as well as the specific value system of the nation they represent. The primary concern remains the potential of a robot going rogue and becoming an uncontrollable killing machine. Robots could even conclude that all of mankind should be destroyed, such as in the *Terminator* movies. Finally, robots might turn on humans if they conclude they are being used by us for our benefit rather than in their individual or collective interest, the latter resulting from a spirit of camaraderie with each other, as soldiers often form in battle.

Emotions help people to form their value systems in a process that develops human intelligence and our rational decision making. In contrast, robotic values are programmed. We could, however, imagine a process whereby robots developed something akin to human emotions over time as part of their learning process. Although emotional learning might result in greater decision making for robots, it might also have dangerous consequences if not employed appropriately. Emotions are not rational in themselves, but they can help with rationality if they are used to assess values and are applied in an intellectually reasoned comparison of future costs and gains that result in maximized utilitarian outcomes. Evaluating only certain lives as worthy in a cost-benefit analysis as is done by each side in a military conflict is ultimately flawed since we should still respect all humanity despite political and cultural differences. Even in the midst of a military operation, we thus provide appropriate medical treatment to wounded enemy soldiers and prisoners of war.

5. Conclusion

As demonstrated by the data presented, military leaders acted with a sense of economic rationality during the Vietnam, Iraq, and Afghanistan Wars. In the future, humankind should create international rules for using AI in warfare. Although policies might be employed in the business world to create worker-AI parity for hiring, robots cannot so easily create their own policies regarding warfare. Russia plans to have 30% of its military equipment AI-controlled by 2030 (Allen 2017), and China has its New Generation Artificial Intelligence Development Plan.

In contrast, President Donald J. Trump is calling for a 10% cut in public research on intelligence systems (Kania 2017). Since people will probably not value the loss of robotic casualties, one would expect that conflicts will dramatically increase in a world fought primarily by AI. Although many of these problems are still a long way off, we need to develop policies regarding proper use of AI in warfare before it is too late.

Acknowledgements:

I would like to thank my colleague Melissa K. Aho for valuable feedback on the essay.

References

Allen, G. C. (2017) Putin and Musk are right: Whoever masters AI will run the world, CNN, available at www.cnn.com/201709/ 05/opinions/russia-weaponize-ai-option-allen/ index.html, accessed on 02.11.2018.

Bernstein, P. L. (1998) *Against the Gods*. New York: John Wiley and Sons.

DCAS (2008) Vietnam Conflict Extract File record counts by INCIDENT OR DEATH DATE, available at www.archives.gov/research/ military/vietnam-war/casualty-statistics.html, accessed on 2.14.2018.

iCasualties (2018) Operation Freedom and Operation Enduring Freedom Casualties, available at http://icasualties.org/, accessed on 2.13.2018.

Leigh, D. (2010) Iraq War Logs Reveal 15,000 Previously Unlisted Civilian Deaths- Leaked Pentagon File Contain Records of More than 100,000 Fatalities Including 66,000 Civilians, *The Guardian* available at www.theguardian.com/world/2010/oct/ 22/true-civilian -body-count-iraq, accessed on 2.14.2018.

Kania, E. B. (2017) Artificial Intelligence and Chinese Power. *Foreign Affairs*, available at https://www.foreignaffairs.com/ articles/china/2017-12-05/artificial-intelligence-and-chinese-power, accessed on 2.11.2018.

Kasparov, G. (2017) *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins. New York:* Public Affairs.

Kolko, G. (1985) *Anatomy of a War,* New York: Pantheon Books.

Rafferty, A. (2017) The War in Afghanistan: By The Numbers, *NBC News*, available at www.nbcnews.com/politics/politics-news/war-afghanistan-numbers-n794626, accessed on 2.13.2018.

Rummel, R. J. (1997): Statistics of Vietnamese Democide Estimates, Calculations,

and Sources, in *Statistics of Democide: Genocide and Mass Murder since 1900*. Charlottesville: Center for National Security Law, School of Law, University of Virginia, available at www.hawaii.edu/powerkills/ SOD.TAB6.1A.GIF, accessed on 2.14.2018.

Schott, J. P., Scherer, L. D., and Lambert, A. J. (2011) Casualties of war and sunk costs: Implications for attitude change and persuasion. *Journal of Experimental Psychology* 47: 1134-1145.

Shiller, R. (2000) *Irrational Exuberance*. Princeton: Princeton University Press.

Todd J. Barry: PhD student
University of Southern Mississippi

Email: tjbarry00@yahoo.com