CCDCOE

NATO Cooperative Cyber Defence Centre of Excellence Tallinn, Estonia

Ludovica Glorioso, Anna-Maria Osula (Eds.)

1st Workshop on Ethics of Cyber Conflict

Proceedings

Tallinn 2014

www.ccdcoe.org
publications@ccdcoe.org

Contents

Introduction

The **NATO Cooperative Cyber Defence Centre of Excellence (NATO CCD COE)** is a NATO-accredited international military organisation that has always employed and valued an interdisciplinary approach to cyber defence, uniting the perspectives of the technical, policy, legal and strategic domains. With the launch of the 'Ethics of Cyber Conflict' project in 2013, the Centre's interdisciplinary research agenda has gained yet another perspective of the debate concerning current and future of cyber defence.

Over recent years the military, international lawyers, ethicists and policymakers have shown an increasing interest in whether and how current policies and international law are applicable to cyber warfare. Responding to this growing attention, the goal of the Centre's ethics project has been to examine the need for innovative and more effective ethical analyses of cyber conflict and cyber warfare, domains that should not be seen as the exclusive concern of the military. The analysis has focused on clarifying the ethical principles prominent in cyber conflict, and specifying ethical guidelines for the endorsement of such principles in policies and regulations.

The Centre's project culminated in the '**Workshop on Ethics of Cyber Conflict',** held at the Centre for High Defence Studies, in Rome, in November 2013. The 1½ day event brought together proclaimed experts from all over the world, and featured presentations from both invited speakers and from authors selected by a peer review process. The international audience of approximately 50 participants included representatives with military, academic, legal and policy backgrounds. The videos of the presentations are available at NATO CCD COE's website at http://ccdcoe.org/468.html.

The workshop's main purpose – to engage international experts in discussing topical issues related to ethics and cyber defence – is now reinforced in the format of workshop proceedings. The articles in the workshop proceedings are the result of vigorous individual academic research and do not reflect in any way NATO's or NATO CCD COE's opinion or official policy. They do exhibit, however, the lively debate of the workshop, and put forward authors' (sometimes provocative) ideas related to the ethical aspects of a number of issues such as 'Just War Theory' in cyber conflict, cyber warfare, cyber espionage and the status of cyber combatants, and the ethical bases of law. As can be inferred from the topical questions set out in many of the articles, well-established disciplines (such as ethics and law) may incline to significantly different approaches to interpreting principles related to the ethical "bottlenecks" of cyber defence which is in itself an indication of a clear need for further discussions over these matters.

Last but not the least, we would like to give special thanks to the **Centre for High Defence Studies in Rome** for their excellent support, to **Dr Mariarosaria Taddeo** for her knowledge and vision throughout the project, and to all of the authors for their expertise and contribution.

Ludovica Glorioso & Anna-Maria Osula
Law & Policy Branch

NATO Cooperative Cyber Defence Centre of Excellence
Tallinn, Estonia

April 2014

The Ethics of Cyber-Conflicts in Hyperhistorical Societies

Luciano Floridi
Oxford Internet Institute
University of Oxford
United Kingdom

luciano.floridi@oii.ox.ac.uk

Information and Communication Technologies (ICTs) are no longer mere tools but environmental forces, increasingly 'enveloping' our world (an *envelope* is the three-dimensional space that defines the boundaries reachable by a robot). Such enveloping leads to a shift from historical to hyperhistorical societies, within which standard distinctions—real vs virtual, natural vs artificial, onlife vs offline and things vs processes—become blurred. This blurring affects our understanding and management of conflicts, with the result that new and pressing challenges in the ethics of cyber-security and cyber-defence require a significant conceptual upgrade. In this talk, I shall argue that an e-nvironmental information ethics offers a fruitful approach to such an upgrade.

THE DIGITAL ELEPHANT IN THE ROOM

The story goes that when the Roman horsemen first saw Pyrrhus's twenty war elephants at the battle of Heraclea (280 BC), they were so terrorised by these strange creatures that they galloped away and the Roman legions lost the battle. Today, the new elephants are digital. The phenomenon might have just begun to emerge in the public debate but, in hyperhistorical societies, ICTs are increasingly shaping armed conflicts.

Disputes become armed conflicts when politics fails. In hyperhistory, such armed conflicts have acquired a new informational nature. Cyberwar or information warfare is the digital continuation of, and sometimes the replacement for, conflict. Four main changes are notable.

First, in terms of conventional military operations, ICTs have progressively revolutionised communications, making possible complex new modes of field operations. We saw this was already the case with the *Chappe* telegraph.

Second, ICTs have also made possible the swift analysis of vast amounts of data, enabling the military, intelligence and law enforcement communities to take action in ever more timely and targeted ways. ICTs and Big Data are also weapons.

Third, and even more significantly, battles are nowadays fought by highly mobile forces, armed with real-time ICT devices, satellites, battlefield sensors and so forth, as well as thousands of robots of all kinds.

And, finally, the growing dependence of societies and their militaries on advanced ICTs has led to strategic cyber-attacks designed to cause costly and crippling disruption. Armies of human soldiers may no longer be needed. This creates a stark contrast with suicide terrorism. Human life can regain its ultimate value because the State no longer needs to trump it in favour of patriotism. Drones do not die 'for King and Country'. Cyberwar is a hyperhistorical phenomenon. Conversely, terrorists de-humanise individuals as mere delivery mechanisms. Suicide terrorism is a historical phenomenon, in which the technology in between is the human body and a person becomes a 'living tool', using Aristotle's definition of a slave.

HYPERHISTROCIAL CONFLICTS

The old economic problem—how to finance war and its expensive technologies—is now joined by a new legal problem: how to reconcile a hyperhistorical kind of warfare with historical phenomena, such as the infringement of national sovereignty and respect for geographical borders. Furthermore, cyber-attacks can be undertaken by nations or networks, or even by small groups or individuals. ICTs have made asymmetric conflicts easier, and shifted the battleground more than an inch into the infosphere.

The scale of such transformations is staggering. For example, in 2003, at the beginning of the war in Iraq, US forces had no robotic systems on the ground. However, by 2004, they had already deployed 150 robots, in 2005 the number was 2,400; and by the end of 2008, about 12,000 robots of nearly two dozen varieties were operating on the ground.

In 2010, Neelie Kroes, Vice-President of the European Commission, commenting on Cyber Europe 2010, the first pan-European cyber-attack simulation, said that:

This exercise to test Europe's preparedness against cyber threats is an important first step towards working together to combat potential online threats to essential infrastructure and ensuring citizens and businesses feel safe and secure online.

As you can see, the perspective could not be more hyperhistorical.

ICT-mediated modes of conflict pose a variety of ethical problems, for war-fighting militaries in the field, for intelligence gathering services, for policy makers and for ethicists. They may be summarised as the three Rs: risks, rights and responsibilities.

THE THREE Rs: RISKS, RIGHTS, AND RESPONSIBILITIES

Risks. Cyberwar and information-based conflicts may increase risks, making 'soft' conflicts more likely and hence potentially increasing the number of casualties. Between 2004 and 2012, drones operated by the US Central Intelligence Agency (CIA) killed more than 2,400 people in Pakistan, including 479 civilians, with 3 strikes in 2005 escalating to 76 strikes in 2011. A troubling perspective is that ICTs might make unconventional conflicts more ethically acceptable by stressing the less deadly outcome of military operations in cyberspace. However, this might be utterly illusory. Messing with ICT-infrastructures of hospitals and airports may easily cause the loss of human lives, even if in a less obvious way than bombs do. Despite this, the mistaken impression remains that we might be allegedly moving towards a more precise, surgical, bloodless way of handling violently our political disagreements.

Rights. Cyberwar tends to erase the threshold between reality and simulation, between life and play and between conventional conflicts, insurgencies and terrorist actions. This threatens to increase the potential tensions between fundamental rights: informational threats require higher levels of control, which may generate conflicts between individuals' rights (e.g. privacy) and community's rights (e.g. safety and security). A State's duty to protect its citizens may come into conflict with its duty to prevent harm to its citizens, due to an extended system of surveillance, which may easily end up infringing citizens' privacy.

Responsibilities. Cyberwar makes it more difficult to identify responsibilities that are reshaped and distributed. Because causal links are much less easily identifiable, it becomes much more difficult to establish who, or what, is accountable and responsible when software, robotic weapons and hybrid, man-machine systems are involved.

New risks, rights and responsibilities: in short, cyberwar is a new phenomenon, which has caught us by surprise. With hindsight, we should have known better, for at least three reasons.

Take the nature of our society first. When it was modern and industrial, conflicts had mechanised, second-order features. Engines, from battleships to tanks to aeroplanes, were weapons, and the coherent outcome was the emphasis on energy; first petrol and then nuclear power. There was an eerie analogy between assembly lines and warfare trenches, between working force and fighting force. Conventional warfare was kinetic warfare. We just did not know it, because the non-kinetic kind was not yet available. The Cold War and the emergence of asymmetric conflicts were part of a post-industrial transformation. Today, in a culture in which we have seen that the word 'engine' is more likely to be preceded by the verb 'search' than by the noun 'petrol', hyperhistorical societies are as likely to fight with digits as they are with bullets, with computers as well as guns, not least because digital systems tend to be in charge of analogue weapons. I am not referring to the use of intelligence, espionage or cryptography, but to cyber attacks or to the extensive use of drones and other military robots in Iraq and Afghanistan. It is old news. On 27th of April 2007, about one million computers worldwide were used for DDOS (distributed denial of service) attacks on Estonian government and corporate web sites. A DDOS attack is a systematic attempt to make computer resources unavailable, at least temporarily, by forcing vital sites or services to reset or to consume their resources, or by disrupting their communications so that they can no longer function properly. Russia was blamed but denied any involvement. In June 2010, Stuxnet, a sophisticated computer malware, sabotaged around 1000 Siemens centrifuges used in the Iranian nuclear power plant of Bushehr. That time, the US and Israel denied any involvement. At the time of writing, there is an on-going attack on US ICT infrastructure. This time it is China that denies any involvement. Then there are robotic weapons, which may be seen as the final stage in the industrialisation of warfare, or, more interestingly, as the first step in the development of information conflicts, in which command and control as well as action and reaction become tele-concepts. Third-order technological conflicts in which humans are no longer in the loop have moved out of science fiction and into military scenarios. From software agents in cyberspace to robots in physical environments, we should not be too optimistic about the non-violent nature of cyberwar. The more we rely on ICTs, the more we envelop the world, the more cyber attacks will become lethal. Soon, crippling an enemy's communication and information infrastructure will be like zapping its pacemaker rather than hacking its mobile.

Second, consider the nature of our environment. We have been talking about the internet and cyberspace for decades. We could have easily imagined that this would become the new frontier for human conflict. Technologies have continuously expanded. We have been fighting each other on land, at sea, in the air, and in space for as long, and as soon, as technologies made it possible. Predictably, the infosphere was never going to be an exception. Information is the fifth element, and the military now speaks of cyberwarfare as 'the fifth domain of warfare'. The impression is that, in the future, such a fifth domain will end up dominating the others. The following two examples may help. On 13th of May 1999, arguably the first combat between an aircraft and an unmanned drone took place when an Iraqi MiG-25 shot down a US Air Force unmanned MQ-1 Predator drone. More than 360 drones have been built since 1995, for more than $2.38 billion. Second, since 2006, Samsung has also been producing the SGR-A1. It is a robot with a low-light camera and pattern recognition software to distinguish humans from animals or other objects. It patrols South Korea's border with North Korea and, if necessary, it can autonomously fire its built-in machine gun. It is increasingly hard to draw a clear distinction between cyberwarfare and conventional, kinetic warfare when some tele-warfare is in question.

Finally, think of the origin of cybernetics, the computer, the internet, the Global Positioning System (GPS) and unmanned drones and vehicles. They all developed initially as part of wider military efforts. The history of computing is deeply rooted in the Second World War and Turing's work at Bletchley Park. Cybernetics, the ancestor of contemporary robotics, begun to develop as an engineering field in connection with applications for the automatic control of gun mounts and radar antenna, also during the Second World War. We know that the internet was the outcome of the arms race and of nuclear proliferation, but we were distracted by the development of the Web and its scientific origins, and forgot about the Defense Advanced Research Projects Agency (DARPA). The now ubiquitous GPS, which provides the satellite-based information for navigation systems, was created and developed by the US Department of Defense, one more case of the political

importance of geography. It became freely available for civilian use only in 1983, after a Boeing 747 of the Korean Air Lines, with 269 people on board, was shot down because it had strayed into the USSR's prohibited airspace. Finally, the development of drones, mainly but not only by the US military, as well as autonomous vehicles (DARPA again) and other robots, owes much to the conflicts in Iraq and Afghanistan and the fight against terrorism. In short, much of the history of digital ICTs spookily corresponds to the history of conflicts and the financial efforts behind them: Second World War, Cold War, First and Second Iraq War, War in Afghanistan, and various 'wars' on terrorist organisations around the world. Hyperhistory has merely caught up with us.

NEW IDEAS FOR NEW PHENOMENA

The previous outline should help one understand why cyberwar, or more generally information warfare, is causing radical transformations in our ways of thinking about military, political, and ethical issues. The concepts of State, war and the distinction between civil society and military organisations are being affected. Are we going to see a new arms race, given the high rate at which cyber weapons 'decay'? After all, you can use a piece of malware only once, for a patch will then become available, and often only within, and against, a specific technology that will soon be out of date. If cyber disarmament is ever going to be an option, how do you decommission cyber weapons? Digital systems can be hacked: will the Pony Express make a patriotic comeback in the near future as the last line of defence against an enemy that could tamper with anything digital and online? Some questions make one smile, but others are increasingly problematic. Let me highlight two sets that should be of more general interest.

The body of knowledge and discussion behind Just War Theory is detailed and extensive. It is the result of centuries of refinements since Roman times. The methodological question we face today is whether information warfare is merely one more area of application, or whether it represents a disruptive novelty as well, which will require new developments of the theory itself. For example, within the *jus ad bellum*, which kind of authorities possesses the legitimacy to wage cyberwar? And how should a cyber attack be considered in terms of last resort, especially when a cyber attack could, allegedly, prevent more violent outcomes? Within the *jus in bello*, what level of proportionality should be attributed to a cyber attack? How do you surrender to cyber enemies, especially when their identities are concealed? Or how will robots deal with non-combatants or treat prisoners? Is it possible or even desirable to develop in-built 'ethical algorithms' when engineering robotic weapons?

Equally developed, in this case since Greek times, is our understanding of military virtue ethics. How is the latter going to be applied to phenomena that are actually reshaping the conditions of possibility of virtue ethics itself? Bear in mind that any virtue ethics presupposes a philosophical anthropology, a view of the human nature that may be Aristotelian, Buddhist, Christian, Confucian, Fascist, Nietzschean, Spartan and so forth. We saw in the previous chapters that information warfare is only part of the information revolution, which is also affecting our self-understanding as informational organisms. Take for example the classic virtue of courage: in what sense can someone be courageous when tele-manoeuvring a military robot? Indeed, will courage still rank so highly among the virtues when the capacity to evaluate and manage information and act upon it wisely and promptly will seem to be a much more important trait of a soldier's character?

Similar questions seem to invite new theorising, rather than the mere application or adaptation of old ideas. ICTs have caused radical changes both in how societies may come into conflict and how they may manage it. At the same time, there is a policy and a conceptual deficit. For example, the US Department of Defense intends to replace a third of its armed vehicles and weaponry with robots by 2015, but it still lacks an ethical code for the deployment of these new, semi-autonomous weapons. This is a global issue. The 2002 Prague Summit marked NATO's first attempt to address cyber-defence activities. Five years later, in 2007, there were already 42 countries working on military robotics, including Iran, China, Belarus and Pakistan, but not even a draft of an international agreement regarding their ethical deployment. There is a serious need for more descriptive and

conceptual analyses of such a crucial area in applied ethics, and more assessment of the effectiveness of the initial measures that have been taken to deal with the increasing application of ICTs in armed conflicts. The issue could not be more pressing, and there is a much felt and quickly escalating need to share information and coordinate ethical theorising. The goals should be sharing information and views about the current state of the ethics of information warfare, developing a comprehensive framework for a clear interpretation of the new aspects of cyberwar, building a critical consensus about the ethical deployment of e-weapons and laying down the foundation for an ethical approach to information warfare. We experimented with chemical weapons, especially during the First World War, and with biological weapons, in particular during the Sino-Japanese War of 1931-1945. The horrific results led, in 1925, to the Geneva Protocol, prohibiting the use of chemical and biological weapons. In 1972, the Biological and Toxin Weapons Convention (BWC) banned the development, production and storage of bio-weapons. Since then, we have managed to restrain their use and, by and large, respect the BWC. Something similar happened with nuclear weapons. The hope is that information warfare and e-weapons will soon be equally regulated and constrained, without having to undergo any terrible and tragic lesson.

CONCLUSION

Let us return to the elephants. During the civil war, in the battle of Thapsus (46 BC), Julius Caesar's Fifth Legion was armed with axes and was ordered to strike at the legs of the enemy's elephants. The legion withstood the charge, and the elephant became its symbol. Interestingly, nobody at the time could even imagine that there might be an ethical problem in treating animals so cruelly. We should think ahead, because history occasionally is a bit petulant and likes to repeat itself. At a time when there is an exponential growth in R&D concerning ICT-based weapons and strategies, we should collaborate on the identification, discussion and resolution of the unprecedented ethical difficulties characterising cyberwar. This is far from being premature. Perhaps, instead of updating our old ethical theories with more and more service packs, we might want to consider upgrading them by developing new ideas. Like the civilian uses of robots, information warfare calls for an information ethics. After all, iRobot produces both the *Roomba 700* that vacuums your floor and the *iRobot 710 Warrior* that disposes of your enemies' explosives.

REFERENCES

Floridi, L. (1999), *Philosophy and Computing: An Introduction* (London, New York: Routledge).

NATO Cooperative Cyber Defence Centre of Excellence (2013), Tallinn Manual on the International Law Applicable to Cyber Warfare : Prepared by the International Group of Experts at the Invitation of the Nato Cooperative Cyber Defence Centre of Excellence (Cambridge: Cambridge University Press).

The Economist (June 2 2012), 'Morals and the Machine'.

The Economist (June 7 2007), 'Robot Wars'.

An Analysis for a Just Cyber Warfare

Mariarosaria Taddeo
Politics and International Studies Department
University of Warwick
United Kingdom

M.Taddeo@warwick.ac.uk

During the past two decades, information and communication technologies (ICTs) proved to be useful and convenient for waging war, so much so that they have been deployed in most conflicts since the second Iraq war.[1] The military deployment of ICTs has radically changed the way in which wars are waged. It has actually determined the latest revolution in military affairs, making the cyber space the fifth domain of war, along with land, sea, air and space.

The informational turn in military affairs is not the exclusive concern of the military; it also concerns ethicists and policymakers, and existing ethical theories of war and national and international regulations struggle to address the novelties of this phenomenon. This article is devoted to developing an ethical analysis of cyber warfare (CW), with the twofold goal of overcoming the theoretical vacuum surrounding this phenomenon and of providing the grounding for an ethical regulation of CW.

The proposed analysis rests on the investigation of CW proposed by Taddeo (2012) which highlights the informational nature of this phenomenon as well as its relation to the so-called Information Revolution. In the rest of this paper it will be argued that Just War Theory (JWT) is a necessary but not sufficient instrument for the ethical analysis of CW. It will be maintained that analysing CW through the lenses of JWT allows for unveiling the fundamental ethical issues that this phenomenon brings to the fore, but that attempting to address these issues solely on the basis of JWT will leave them unresolved.

The thesis will be advanced that the problems encountered when addressing CW through JWT are overcome when the latter is merged with Information Ethics (Floridi, 2008). This is a macro-ethical theory developed to take into account the features and the ethical implications of informational phenomena, like internet neutrality (Turilli et al., Forthcoming), online trust (Turilli et al., 2010), peer-to-peer (Taddeo and Vaccaro, 2011) and CW. The goal is to develop an ethical analysis of CW able to take into account both its peculiarities and its novelty, while at the same time being consistent with the mainstream ethical analysis of warfare.

Having delineated the path of the analysis proposed in this article, we shall now begin by considering in more details the nature of CW.

CYBER WARFARE

For the purposes of this article CW is defined as follows:

'[Cyber] Warfare is [the warfare grounded on certain] uses of ICTs within an offensive or defensive military strategy endorsed by a state and aiming at the immediate disruption or control of the enemy's resources, and which is waged within the informational environment, with agents and targets ranging both on the physical and non-physical domains and whose level of violence may vary upon circumstances.' (Taddeo 2012, 114)

[1] http://www.economist.com/node/16478792

This definition highlights two aspects of CW: its *informational nature* and its *transversality*.[2] The informational nature of CW is a consequence of the fact that this kind of warfare rests on the military deployment of technological artefacts devoted to elaborating, managing and communicating data and information. In this respect CW is shown to be related to the so-called Information Revolution.

The Information Revolution is a multi-faced phenomenon. It rests on the development and capillary dissemination of the use of ICTs, which have a wide impact on many of our daily practises from working to interacting with other human beings, driving around and planning holidays. The dissemination of ICTs has important philosophical implications (Floridi, 2010), for the Information Revolution fundamentally changes the way in which reality is perceived and understood.

The Information Revolution determines a shift which brings the *non-physical domain* to the fore and makes it as important and valuable as the physical one. CW is one of the most compelling instances of such a shift; it shows that there is a new environment, where physical and non-physical entities coexist and are equally valuable, and in which States have to prove their authority and new modes of warfare are being developed specifically to be deployed in such a new environment (Taddeo, 2012).[3]

The shift toward the non-physical domain provides the ground for the transversality of CW. This is a complex aspect, and can be better grasped when CW is compared with traditional forms of warfare. Traditional war is understood as the use of a State's *violence* through the State's *military* forces to determine the conditions of governance over a determined territory (Gelven, 1994). It is a necessarily violent phenomenon, which implies the sacrifice of human lives and the damage of both military and civilian infrastructures. The problem to be faced when waging traditional warfare is how to reduce to the minimum such damage while ensuring that the enemy is overpowered.

CW differs from traditional warfare in that it is not a necessarily violent and destructive phenomenon (Arquilla, 1999). CW may involve a computer virus able to disrupt or deny access to the enemy's database, and in so doing cause severe damage to the enemy without exerting *physical* force or violence. In the same way, CW does not necessarily involve human beings. An action of war in this context can be conducted by a computer virus, targeting other artificial agents or informational infrastructures, such as a database or a website (see Figure 1). Nevertheless, CW is to be feared as much as traditional warfare, because it is transversal with respect to the level of violence and may escalate from non-violent to more violent forms. Consider, for example, the consequences of a cyber attack targeting a military air control system causing aircraft to crash (Waltz, 1998). The transversality of CW with respect to the levels of violence, the nature of the agents and the waging domain is the key feature of this phenomenon, the aspect that differentiates it the most from traditional warfare, and also the feature that engenders the ethical problems posed by CW.

Transversality makes CW extremely appealing from both ethical and political perspectives (Arquilla and Ronfeldt 1997). At first glance, CW seems to avoid bloodshed and human commitment and therefore it liberates political authorities of the burden of justifying military actions to public opinion. A more attentive analysis unveils that CW should be feared as much as traditional warfare as it can lead to highly violent and destructive consequences, which could be dangerous for both the military forces and civil society.

[2] 'Transversality' is used in this article to indicate that CW cuts across any qualifying binary such as 'violent-non violent', 'civil-military' or 'human agents-artificial agents'. This aspect is quite different from traditional warfare, which is violent, conducted by militaries and mainly by human agents.

[3] The USA only spent $400 million in developing technologies for cyber conflicts: http://www.wired.com/dangerroom/2010/05/cyberwar-cassandras-get-400-million-in-conflict-cash/The UK devoted £650 million to the same purpose: http://www.theinquirer.net/inquirer/news/1896098/british-military-spend-gbp650-million-cyber-warfare

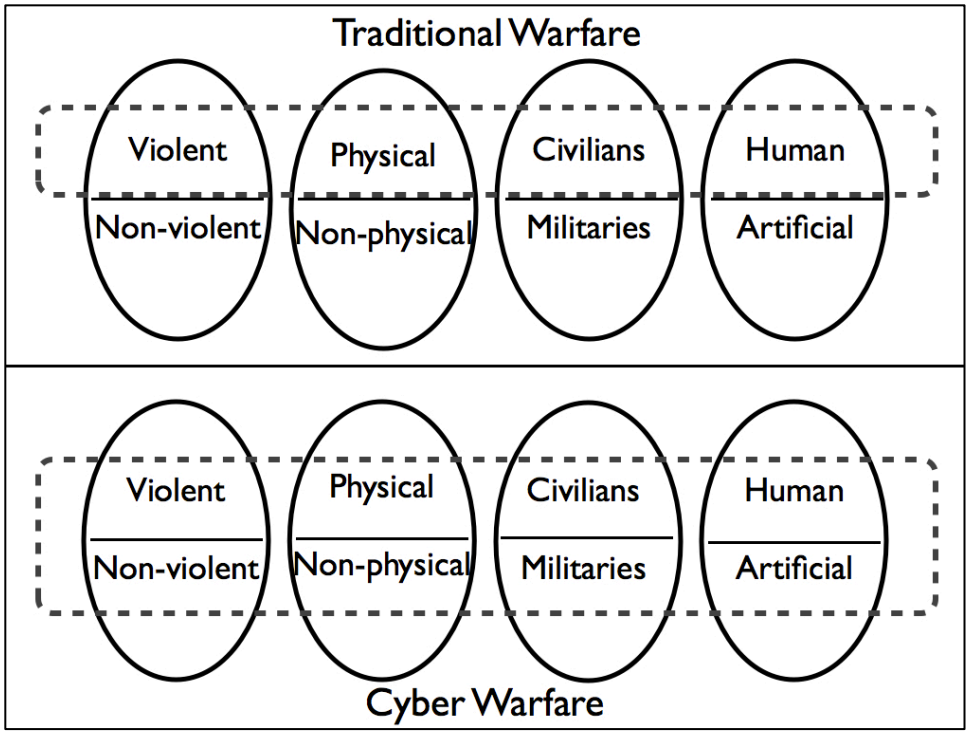

Figure 1 CW compared to traditional warfare in respect to the binaries 'violent and non-violent', 'civilians-militaries', 'human and artificial agents' and 'physical and non-physical'. These binaries are emblematic of the kind of war which is waged as they identify the mode of waging it, the agents involved in it and the targets. While traditional warfare is conducted in a violent manner, including human agents and mainly militaries, the latter is transversal as it spans the binaries.

For this reason, declaring and waging CW requires a strict ethical regulation to guarantee its fairness. An analysis of CW unveiling the ethical issues that it engenders and pointing at the direction for their solution is a necessary step towards the achievement of such goal.

JUST WAR THEORY AND CW

JWT refers to war as a violent and sanguinary phenomenon, declared by States and their official leaders and waged by military forces. Such a scenario is quite different from that determined by CW; the difference between the two forms of warfare is the origin of the problems arising when the principles of JWT are applied to CW. In this respect, there are three issues that deserve attention which follow from the application of the principles of 'war as last resort', of 'more good than harm', and of 'non-combatant immunity' to CW.

The application of the principle of 'war as last resort' is shaken when CW is taken into consideration, because in this case war may be bloodless and may not involve physical violence at all. In these circumstances, the use of the principle of war as last resort becomes less immediate.

Imagine, for example, the case of tense relations between two States and that the tension could be resolved if one of the States decides to launch a cyber attack on the other State's informational infrastructure. The attack would be bloodless as it would affect only the informational grid of the other State and there would be no casualties. The attack could also lead to resolution of the tension and avert the possibility of a traditional war for the foreseeable future. Nevertheless, according to JWT, the attack would be an act of war, and as such forbidden as a first strike move. The impasse is quite dramatic, for if the State decides not to launch the cyber attack it will be probably forced to engage in a sanguinary war in the future, but if the State authorises the cyber attack it will breach the principle of war as last resort and commit an unethical action which would probably be sanctioned by international regulations.

This example is emblematic of the problems encountered in the attempt to establish ethical guidelines for CW. In this case, the main problem is due to the transversality of the modes of combat, which make it difficult to define unequivocal ethical guidelines. In light of the principle of last resort, soft and non-violent cases of CW can be approved as means for avoiding traditional war (Perry, 1995), as they can be considered a viable alternative to bloodshed. At the same time, even the soft cases of CW have a disruptive purpose, that of disrupting the enemy's informational resources (Floridi, 2008), which needs to be taken into consideration by any analysis aimed at providing ethical guidelines for CW even when the disruption of the enemy's informational infrastructure is not achieved through violent and sanguinary means.[4]

The second problem to be considered concerns the principle of 'more good than harm'. According to this principle, a State is justified in declaring war only when the good is proportionate to the evil. This balance is easily assessed in cases of traditional warfare, where the evils are mainly considered in terms of the casualties and physical damage. Determining the equilibrium between the good and the evil becomes more problematic when CW is considered.

CW is likely to cause no or very few casualties, and as it targets informational infrastructures it is unlikely to cause the destruction of physical objects like buildings. Although it is possible for CW to turn into violent warfare, in most cases it does not cause physical damages, but nonetheless may result in unethical actions. If the only criterion for the assessment of harm in warfare scenario remains the consideration of the physical damage caused by war, then an unwelcome consequence follows. All the non-violent cases of CW comply by default to this principle. Therefore, destroying a digital database or erasing a digital archive containing the important historical records of a nation are all deemed to be ethical actions as they do not constitute *per se* physical damage.

In the case of this principle, it is not the prescription that the good should be greater than the harm in order to upset the decision to wage war, it is rather the set of criteria to assess the good and the harm, which are shown to be inadequate when considering CW.

The last problem concerns the principle of 'discrimination and non-combatant immunity'. This principle refers to a classic war scenario and aims at reducing bloodshed and prohibits any form of violence against non-combatants. Its correctness is not in question, yet its application is quite difficult in the context of CW.

In classic warfare, the distinction between combatants and non-combatants reflects the distinction between military and civil society, even if the diffusion of terrorism and guerrilla warfare during the twentieth century weakened the distinction between non-combatants and civilians. In the case of CW such association becomes even feebler due to the blurring between civil society and military organisations (Schmitt, 1999; Shulman, 1999).

This blurring leads to the involvement of civilians in war actions and poses two issues. The first concerns the discrimination itself: in the CW scenario it is difficult to distinguish combatants from non-combatants; wearing a uniform is no longer a sufficient criterion to identify someone's status. Civilians may take part in a combat action from the comfort of their homes, while carrying on with their civilian life and hiding their status as cyber warriors.

The second issue concerns the effects of this difficulty in distinguishing combatants from non-combatants and unveils an ethical conundrum. If combatants can easily hide themselves among the civilian population, then States may be justified in imposing high levels of surveillance over the entire population, thereby breaching individual rights like privacy and anonymity in order to identify the combatants and to guarantee the security of the entire community. For the sake of these goals, public authorities could also be justified in persecuting certain sections of the civilian population which are profiled and deemed to be potentially dangerous to the

[4] For a more in depth analysis of the non-violent cases of CW and their assessment as acts of war or of espionage see Arquilla (1998) and (Taddeo 2014).

community. Therefore, respecting the principle of discrimination may lead to the violation of individual rights, but waiving the principle of discrimination leads to bloodshed and dissemination of violence over the entire civil population, because the policy could be endorsed to target everyone or everything a soldier encounters in his or her way, as being potentially involved in the conflict.

It would be misleading to consider the problems described in this section as reasons to disregard JWT when analysing CW. The ideal of just warfare provided by JWT and its principles remain valid even when considering this new kind of warfare. Yet, the analysis proposed in this section points to a more fundamental problem, namely the need to provide an ethical framework for the regulation of CW able to address the novelty of this phenomenon.

INFORMATION ETHICS

Information Ethics is concerned with the ethical issues in which information is involved as a resource, as a product, and as a target (Floridi, 2008a). It proposes a twofold approach: (i) considering the whole information-cycle, from creation, to communication and storage, and (ii) analysing *informationally* all entities involved in a moral scenario. The moral agents and their actions are considered as part of the informational environment to which they belong as informational entities themselves.

In this framework, two concepts are of pivotal relevance: the infosphere and informational ontology. The infosphere is the totality of what exists. It includes agents and objects, relations and processes, as well as the space in which they act. It is not to be confused with cyberspace, as it includes online as well as offline and analogue domains. The infosphere comprises e-books and trees, online websites and rocks, movies in digital format and paintings on canvas.

The infosphere is the environment in which animate and inanimate, and digital and analogue informational objects are morally evaluated. Information Ethics endorses a universal approach, according to which all existing things, not only human beings and living things, but also artefacts and digital artefacts enjoy some minimal and overridable moral rights.

This universal perspective is grounded in an ontocentric principle in which all entities, understood as informational objects, have fundamental rights to exist and flourish. In Floridi's words: '[...] any form of reality (any instance of information/being), simply by the fact of being what it is, enjoys a minimal, initial, overridable, equal right to exist (be left alone) and develop (not to be interfered) in a way which benefits its nature' (Floridi, 2007b).

In such a universal context, the morality of a given action is assessed with respect to the effects that it will have on the patients, the recipients of the action, and ultimately on the infosphere. This is referred to as the patient-oriented perspective of Information Ethics, according to which we can decide whether an action is evil only on the basis of a clear understanding of its effects on interacting patients.

In a nutshell, Information Ethics is an environmental ethics which endorses an ontocentric and patient-oriented approach, and in which the morality of a course of action is evaluated on the basis of its effects on informational entities and ultimately on the infosphere. (Floridi, 2008a).

Within this framework, Information Ethics provides four moral principles that ought to be respected in order to preserve the wellbeing and continued flourishing of the Infosphere and its inhabitants.

1) entropy ought not to be caused in the Infosphere (null law);
2) entropy ought to be prevented in the infosphere;
3) entropy ought to be removed from the infosphere;
4) the flourishing of informational entities as well as the whole infosphere ought to be promoted by preserving, cultivating, enhancing and enriching their properties.

The concept of *entropy* adopted in the four laws indicates the result of any form of 'destruction, corruption, pollution, depletion (marked reduction in quantity, content, quality, or value) or unjustified closure of the Infosphere' (Floridi, 2001). Informational entropy is the evil which should be avoided in the infosphere, and should be understood as a metaphysical concept. It is not related to the concept of physical entropy or the use of entropy made in Shannon's information theory.

JUST CW

Following the ontocentric approach, all informational entities enjoy some minimal rights to exist and flourish in the infosphere. As such all entities, be they living things or non-living things, physical or virtual, deserve some minimal respect. When applied to CW, this principle allows for considering as moral patients all the entities that may be affected by an action of war within CW. A human being who suffers the consequences of a cyber attack and an informational infrastructure that is disrupted by a cyber attack are to be considered the receiver of the moral action. The morality of that action will be assessed on the basis on its effect on their rights to exist and flourish.[5]

The first question when considering the conditions for a just CW concerns the rights of the informational entities, namely what and whose rights should be preserved. The answer to this question follows from the rationale of Information Ethics. This states that an entity loses it rights to exist and flourish when it comes into conflict with the rights of other entities or with the well being of the infosphere. Therefore, any entity that causes entropy in the infosphere loses its informational rights as it conflicts with the well being of the other entities and ultimately of the infosphere itself. It is a moral duty of the other inhabitants of the infosphere to *remove* such a malicious entity from the infosphere, as it is a cause of entropy, or to impede it in perpetrating more evil.

This lays the ground for the first principle for just CW. The principle prescribes the condition under which the choice to resort to CW is morally justified:

CW ought to be waged only against those entities that endanger or disrupt the well being of the infosphere.

Two more principles regulate just CW, they are:

CW ought to be waged to preserve the well being of the infosphere.

CW ought not to be waged to promote the well being of the infosphere.

The second principle limits the task of CW to restore the *status quo* in the infosphere before the malicious entity began increasing the entropy in it. According to the second principle, CW should act only when some evil

[5] While assuming that all entities share some initial rights to exist and flourish, Information Ethics does not claim that there is no hierarchy among the entities. It specifies that the rights are overridable and hence that an entity ceases to hold the rights to exist and flourish should it contravene the well being of other entities or of the infosphere. Under Information Ethics, the position in the hierarchy of an entity depends on its contribution to the flourishing of the infosphere. For a more in depth analysis of the criteria to override the entities' initial rights see Floridi (2008).

has been or is about to be perpetrated with the goal of stopping it. CW ought to be endorsed as an *active* measure in response to the increasing of the evil, and not as a *proactive* measure to foster the flourishing of the infosphere. This is explicitly forbidden by the third principle, which prescribes that the promotion of the well being of the infosphere does not pertain to the scope of a just CW.

THREE PRINCIPLES FOR A JUST CW

The application of the principle of 'last resort' provides the first instance of how JWT and Information Ethics are merged. The principle takes into account traditional violent forms of warfare, and is coupled with the principle of 'right cause', which justifies the resort to war only in self-defence. As much as rightful this approach is when referred to traditional form of warfare, it is inadequate when CW is taken into consideration. The impasse is overcome when considering the principles for just CW.

The first principle prescribes that any entity that endangers or disrupts the well being of the infosphere loses its basic rights and becomes a valid target. Therefore, a State can rightly endorse CW as an early move against a malicious entity. The choice to resort to CW is further justified if it allows a State to avoid the possibility of traditional warfare, as this one would determine casualties and destructions in the infosphere, and as such it is deemed to be a greater evil than CW.

A caveat must be stressed in this case: the waging of CW must comply with the principles of 'proportionality' and 'more good than harm'. In waging CW, the means endorsed to win must be sufficient to stop the malicious entity, yet they ought not to generate more entropy than the entity which the State is aiming to remove from the infosphere. This leads us to consider in more detail the principle of 'more good than harm'.

The application of this principle is of paramount importance for the waging of a Just War, whether a traditional or an informational one. The issues concerning CW are due to the definition of the criteria for the assessment of the 'good' and the 'harm' that warfare may cause. Traditionally, they are defined with respect to collateral damage, casualties and damage to physical infrastructures of both the parties involved in the war. Such criteria do not take in consideration the harm that CW may cause.

In the case of CW, the damage to non-physical entities needs to be considered, as well as the damage to the physical. The assessment of the good and the harm should be determined by considering the general condition of the infosphere 'before and after' waging the war. A Just War never results in greater entropy (evil) than that it intended to remove from the infosphere in the first place. Once considered from this perspective, the principle of more good than harm acts as corollary of the second principle for just CW. It ensures that a just CW is waged to restore the *status quo* and it never increases the level of entropy in the infosphere.

The assessment of the entropy in the infosphere also allows for reconsidering the application of the principle of non-combatant immunity to CW. Two problems accompany the application of this principle: the consequences of its endorsement on the individuals' rights of privacy and anonymity, and the very distinction between combatants and non-combatants. The rest of this section will focus only on the latter issue; the former does not fall within the scope of this paper.[6]

The distinction between combatants and non-combatants promoted by this principle rests on the distinction between the military and civilians that it is inherited from traditional warfare. As we have seen, CW is transversal with respect to the social status of the combatants, for it does not require military skills to be waged. This makes the application of the principle problematic, but it nevertheless has to be maintained as it prescribes the distinction between enemies and 'innocents'.

[6] For an in depth analysis of this issue see (Taddeo 2012).

Help in applying this principle to CW comes from the first principle for just CW, which allows for overcoming the distinction between the military and civilians, and for substituting it with the distinction between valid and invalid targets, the former being the malicious entities that endangered or disrupted the well being of the infosphere.

CONCLUSION

This article rests on the conceptual analysis of CW provided in the second section. Such analysis stresses the novelty of this phenomenon and its relation with the Information Revolution, and argues that transversality is its main feature. Transversality is deemed to be the characteristic of CW that most differentiates it from traditional warfare, and also the one from which all the ethical issues posed by CW originate.

It has been argued that, given the radical novelty posed by CW, the ethical analysis of this phenomenon and the definition of the ethical principles for a just CW cannot rest solely on JWT because such a theory does not provide 'the right sieve' for the work. JWT does not take into account the main features of CW, namely the transversality of the levels of violence, the domain (physical and non-physical) in which it is waged, or the transversality of the nature and social status of agents who may be involved in this warfare. Yet, it would be a mistake to reject JWT altogether when addressing CW.

The ideal of just warfare and the principles prescribed by JWT are still valid when referred to CW, and they can be endorsed to regulate this new form of warfare if they are combined with a macro-ethical framework able to take into account the peculiarities of this phenomenon.

Information Ethics is a suitable ethical framework for CW. This is a macro-ethics, which endorses an ontocentric, patient-oriented and ecological approach, and is devoted to addressing the ethical problems posed by informational phenomena. In particular, the ecological facet of Information Ethics is extremely relevant for the purpose of the analysis proposed in this article, as by posing the well being of the infosphere as the ultimate good and the creation of entropy in the infosphere as the moral evil, it provides the criteria for the ethical assessment of the implications of CW.

Three principles for just CW, encompassing both the rationale of JWT and that of Information Ethics, have been provided. Such principles constitute the grounding for the development of more detailed ethical guidelines for CW that is for the next step of this research.

REFERENCES

Arquilla, J. (1998). Can information warfare ever be just? *Ethics and Information Technology, 1*(3), 203-212.

Arquilla, J. (1999). Ethics and information warfare. In Z. Khalilzad, J. White, & A. Marsall (Eds.), *Strategic appraisal: the changing role of information in warfare* (pp. 379-401). Santa Monica, USA: Rand Corporation.

Arquilla, J., & Ronfeldt, D. (1997). *In Athena's Camp: Preparing for Conflict in the Information Age*. Santa Monica, CA: RAND Corporation.

Floridi, L. (2008). Information Ethics, its Nature and Scope. In Hoven, J. v. d. & Weckert J. (Eds.), *Information Technology and Moral Philosophy* (Vol. 40-65). Cambridge: Cambridge University Press.

Floridi, L. (2010). The Digital Revolution as The Fourth Revolution. *Invited contribution to the BBC online program Digital Revolution*.

Gelven, M. (1994). *War and Existence*. Philadelphia, PA: Pennsylvania State University Press.

Perry, D. L. (1995). Repugnant Philosophy: Ethics, Espionage, and Covert Action. *Journal of Conflict Studies, Spring*.

Schmitt, M. N. (1999). The Principle of Discrimination in 21st Century Warfare. *Yale Human Rights and Development Law Journal, 2*, 143-160.

Shulman, M. R. (1999). Discrimination in the Laws of Information Warfare. *Pace Law Faculty Publications, 37*, 939-968.

Taddeo, M. (2012). Information Warfare: a Philosophical Perspective. *Philosophy and Technology, 25*(1), 105-120.

Taddeo, M. (Forthcoming). Just Information Warfare. *Topoi.*

Taddeo, M., & Vaccaro, A. (2011). Analyzing peer-to-peer technology using information ethics. *The Information Society, 27*(2), 105 - 112.

Turilli, M., Vaccaro, A., & Taddeo, M. (2010). The Case of on-line Trust. *Knoweldge, Technology and Policy*.

Turilli, M., Vaccaro, A., & Taddeo, M. (Forthcoming). Internet Neutrality: Ethical Issues in the Internet Environment.

Waltz, E. L. (1998). *Information Warfare Principles and Operations*. Norwood, USA: Publisher Artech House, Inc.

Law, ethics and cyber warfare

Bill Boothby
Geneva Centre for Security Policy
Switzerland

William.Boothby@rhul.ac.uk

The idea that there is any sort of morality associated with war seems at first somewhat farfetched – indeed the very notion that mature societies should resolve their political differences by hurling explosive projectiles at one another is ridiculous. That those same societies should celebrate the creation of a new domain for human interaction such as cyberspace by exploring how it can be employed to cause each other harm in an armed conflict causes one to develop the most pessimistic view of the human condition. And yet that is the very starting point for this chapter.

The law recognises warfare as a reality and seeks to regulate its conduct. The body of law that addresses the conduct of hostilities during armed conflicts forms part of public international law, which in turn forms a part of international law. It therefore logically follows that the acknowledged sources of international law, namely fundamental principles of law,[7] customary law[8] and treaty law,[9] are also the sources to which we should look for the principles and rules of the law of armed conflict.[10] Judicial decisions and the teachings of highly qualified commentators on the law are authoritative yet subsidiary means of identifying rules of international law, but are not sources of the law *per se* in the strictest sense.[11]

It is immediately plain from this explanation that international law is based on how States behave pursuant to their view of the law and on what law States make by virtue of agreements among themselves. The nation State is therefore at the very centre of the formation of international law, with the obvious consequence that the collective ethical and moral perceptions of States in general will tend to be reflected in the legal rules that they recognise and make.

The law of armed conflict is really divided into two distinct and separate parts: that body of law that regulates the resort to the use of force, known often by its Latin descriptor *jus ad bellum*, and the law that regulates activities undertaken during and in connection with an armed conflict, similarly known as the *jus in bello.* What follows is the briefest of summaries of the *jus ad bellum*, as greater detail lies outside the intended scope of this chapter.

[7] The fundamental principles of greatest relevance to the current discussion would seem to be the principles of military necessity, of unnecessary suffering, of proportionality, of chivalry and of distinction; see for example Manual of Military Law 1958, Part III, Page 2, para. 3 and UK Manual of the Law of Armed Conflict, 2004, (UK Manual) at pages 21-6, where the basic principles of the modern law of armed conflict as a whole are identified as military necessity, humanity, distinction and proportionality.
[8] Customary law rules are to be found in the general practice of states, reflecting a general view by them that the conduct in question is required by law. The practice does not need to be universal, but should be sufficiently extensive and convincing to support the belief that a rule of international law is involved. See UK Manual, paragraphs 1.12 and 1.12.1.
[9] A treaty is an international agreement between States in written form and which is governed by international law; it is the nature of the document itself rather than how it is entitled or designated that establishes its status as a treaty; Vienna Convention on the Law of Treaties, 1969, article 2(1)(a).
[10] See Statute of the International Court of Justice, San Francisco, 26 June 1945, (ICJ Statute) article 38.
[11] ICJ Statute, article 38(1)(d).

UN CHARTER RESTRAINTS ON THE RESORT TO FORCE

The law with regard to the resort to the use of force is governed for all practical purposes by the UN Charter. Under article 2 of the Charter, all members of the United Nations, that is all States, are required to resolve their international disputes by peaceful means in such a way that international peace and security are not endangered.[12] Additionally, all States are required in their international relations to refrain from the threat or use of force against the territorial integrity or political independence of any State, or in any other way that is inconsistent with the Charter's purposes.[13] Armed force may only be lawfully employed by a State in two circumstances. The first arises when a State is using force in exercise of its inherent right to individual or collective self-defence against armed attack.[14] The second is when the UN Security Council, acting under Chapter VII of the Charter, authorises the taking of such action by air, sea or land forces as may be necessary to maintain or restore international peace and security.[15]

It will immediately be appreciated that the maintenance of international peace, based on appropriate standards of behaviour between sovereign States, is the core philosophical principle on which the structure of international security is built. Where a State is subjected to armed attack, the use of immediate and proportionate force by or at the request of the victim State is permitted. Similarly, where the UN Security Council determines the existence of a threat to the peace, a breach of the peace or an act of aggression,[16] and where provisional measures under article 40 of the UN Charter or non-forceful measures under article 41 of the Charter are considered inappropriate or have proved unsuccessful, it can issue an article 42 authorisation. The core point for the purposes of this Chapter is that the entire regime is designed to favour the restoration and/or maintenance of stable peaceful relations between all States.

THE LAW REGULATING ARMED CONFLICT HOSTILITIES

An extensive body of law regulates the way in which force may lawfully be used in pursuance of an armed conflict. In this section of this chapter, an attempt will be made to draw attention to some core themes of this *jus in bello* with a view to illustrating the ethical and moral foundations on which the law has been constructed.

The modern law of armed conflict really began to emerge in the early 1860s with the writings of an American scholar, Dr Francis Lieber of Columbia University in the United States. His code, the Lieber Code, was written for President Lincoln and was adopted for the guidance of the Union side in the American Civil War as the code of law to regulate the behaviour of their forces.[17] Dr Lieber's Code was extensive and detailed, covered all aspects of the conduct of hostilities as they would be undertaken in those times and provided a valuable summary of the law as it was then understood. While some notions in the Code may seem somewhat dated to today's reader, many of the principles and rules that he enunciated are reflected, sometimes in very similar terms, in modern law.

Of the Code's 157 articles, I will refer only to one, as it illustrates an important and relevant point rather well by providing that 'the law of war does not only disclaim all cruelty and bad faith concerning engagements concluded with the enemy during the war, but also the breaking of stipulations solemnly contracted by the belligerents in time of peace, and avowedly intended to remain in force in case of war between the contracting

[12] UN Charter, Article 2(3).
[13] UN Charter, Article 2(4).
[14] UN Charter, Article 51.
[15] UN Charter, Article 42.
[16] Such a determination is made in accordance with Article 39 of the UN Charter.
[17] United States General Orders No. 100, War Department, Adjutant General's Office, dated 24 April 1863 (Lieber Code).

powers.'[18] This notion of good faith between belligerents illustrates that ethical conduct shall not be abandoned merely because warfare has broken out.

Treaty law on the conduct of hostilities started to emerge in 1868 with the St Petersburg Declaration.[19] The operative provisions of the Declaration need not trouble us, but the Preamble notes that: the progress of civilisation should have the effect of alleviating as much as possible the calamities of war; that the only legitimate object to be accomplished in war is to weaken the military forces of the enemy; that it is sufficient for this purpose to disable the greatest possible number of men; and that this object would be exceeded by the employment of arms which uselessly aggravate the sufferings of disabled men or render their death inevitable.[20] This language is really the fore-runner of a customary principle which binds all States and which prohibits the employment of weapons or ways of using them that are of a nature to cause superfluous injury or unnecessary suffering,[21] a legal rule that protects combatants.

The same instrument notes the need to conciliate 'the necessities of war with the laws of humanity'.[22] This reference to humanity is reflected, again in a preambular paragraph, this time to the 1899 Hague Convention II[23] and the corresponding Hague Convention IV of 1907,[24] both of which introduced the first treaty regulations to address the conduct of warfare on land. The Martens Clause, as it is known, is articulated in the latter instrument as follows:

'Until a more complete code of the laws of war has been issued, the high contracting Parties deem it expedient to declare that, in cases not included in the Regulations adopted by them, the inhabitants and the belligerents remain under the protection and the rule of the principles of the law of nations, as they result from the usages established among civilized peoples, from the laws of humanity and from the dictates of the public conscience.'[25]

The repetition of these sentiments in the treaty that contains the most modern provisions on the law of targeting[26] reinforces the point that there are no aspects of the conduct of armed conflict to which the law does not apply, emphasises the centrality of moral principle in the philosophy underpinning that law, and acknowledges that the law of armed conflict remains incomplete.

Protection of inhabitants and of combatants represents two very distinct notions. The inhabitants are, in general terms, representative of those who are not involved in the fight, and where they are concerned the law protects them because there is no military purpose to be achieved in doing otherwise and because simple notions of humanity dictate that those not involved in the fight should be spared. Protection of combatants is, perhaps, rather harder to understand; but the point to appreciate is that all involved in armed conflict are protected by the law in differing ways.

The law we are talking about here consists of what States do or refrain from doing based either on a widely accepted legal practice of States or on treaty obligations applying to the particular State. It is this State practice and these treaty rules that reflect the widely accepted moral principles that States believe should apply during the conduct of hostilities. But writing treaties about it is one thing – compliance with the rules is not always

[18] Lieber Code, Article 11(1).
[19] Declaration Renouncing the Use in Time of War of Explosive Projectiles Under 400 Grammes Weight, St Petersburg, 11 December 1868 (St Petersburg Declaration).
[20] St Petersburg Declaration, preambular paragraphs 1-4.
[21] Geneva Protocol I Additional to the Geneva Conventions of 12 August 1949 and Relating to the Protection of Victims of International Armed Conflicts, Geneva, 8 June 1977 (API), Article 35(2).
[22] St Petersburg Declaration, final paragraph.
[23] Convention II with Respect to the Laws and Customs of War on Land, The Hague, 29 July 1899.
[24] Convention IV with Respect to the Laws and Customs of War on Land, The Hague, 18 October 1907.
[25] Hague Convention IV, 1907, preambular paragraph 8.
[26] API, Article 1(2).

achieved, and so an important treaty rule states that 'In any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited.'[27] That was already a well-established idea, that you cannot simply use any method, however barbaric, of injuring the enemy.

As the years have gone by, nations have fashioned detailed laws to protect prisoners, the wounded, the sick, the shipwrecked, civilians in the power of the enemy, civilians in areas to be attacked, cultural property, foodstuffs and other items on which civilian populations rely, civilian objects in general, dams, dykes and nuclear electrical generating stations, medical facilities, transports and personnel, religious personnel, cultural objects and so on. The underlying purpose is to differentiate between those carrying on the fight and the objects that contribute directly to their conduct of hostilities on the one hand, and civilians who remain uninvolved in the hostilities and objects whose nature, location, purpose or use does not contribute to military action and objects as well as persons entitled to specific protection on the other. The philosophical purpose is to limit the scope of the fight to those persons and objects engaged in it and to protect those persons and objects that remain uninvolved.

Central to the law relating to the conduct of hostilities is this principle of distinction, described by the International Court of Justice in the Nuclear Weapons Advisory Opinion as a 'cardinal principle'.[28] It requires that '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.'[29]

The legal rules that, taken together, constitute targeting law are based on this fundamental principle. Article 51 of API therefore prohibits making the civilian population or individual civilians the object of attack, and Article 52 of the same treaty gives similar protection to civilian objects. Recognising that attacks on lawful targets may cause incidental injury to civilians or incidental damage to civilian objects, the treaty expressly prohibits indiscriminate attacks, that is attacks that are not directed at a specific military objective, or which employ weapons or ways of using them that cannot be so directed, or the effects of which cannot be appropriately limited, with the result that they are of a nature to strike lawful military targets and civilians or civilian objects without distinction.[30]

The law regulating targeting does not satisfy itself with the simple statements of principle and the specific protections of particular persons and objects that we have seen. Instead, it sets out precautionary measures that attackers must take and other precautions that must be taken against the effects of attacks. It is these precautionary rules that are critical to the delivery of effective protection to civilians, civilian objects and to the persons and objects entitled to specific protection. To understand their significance, it is necessary first to explain that attacks are, for these purposes, defined as acts of violence against the adversary, whether in offence or defence.[31] So all those using violence in armed conflict must take a detailed list of precautions,[32] the clear purpose of which is to seek to ensure that only lawful targets are attacked, that indiscriminate attacks are avoided, and that where attacks may affect the civilian population and where it is militarily feasible to do so, effective advance warnings are given. For the present purposes, it suffices to note the general precautionary

[27] API, Article 35(1).
[28] ICJ Nuclear Weapons Advisory Opinion, paragraph 78.
[29] The principle is codified in this form in API, Article 48.
[30] API, Article 51(4). Attacks that treat as a single target a number of clearly separated and distinct military objectives located in an urban or village area containing a similar concentration of civilians or civilian objects constitute one example of indiscriminate attacks. Another example is the proportionality rule which prohibits attacks that may be expected to cause incidental loss of civilian life, injury to civilians or damage to civilian objects, or a combination thereof that would be excessive in relation to the concrete and direct military advantage anticipated; API, Article 51(5).
[31] API, Article 49(1).
[32] See API, Article 57(2) and (3).

rule that '[i]n the conduct of military operations, constant care shall be taken to spare the civilian population, individual civilians and civilian objects.'[33] It is therefore clear that no person involved in attack operations is exempt from this requirement, and that the duty to take care applies at all times.

There are corresponding precautions that must be taken against the effects of attacks.[34] These require all parties to the conflict to try to remove civilians and civilian objects from the vicinity of military objectives to the maximum extent feasible, and thus from locations likely to be the subject of lawful attack, to avoid locating military objectives near or in densely populated areas, and to take other necessary protective precautions. It is the combined effect of the precautions to be taken by attackers and of those to be taken against the effects of attacks that are intended to provide effective protection for civilians.

OTHER LAW OF ARMED CONFLICT PROVISIONS

The focus so far has inevitably been on some of the rules that regulate targeting. The principle of good faith to which we referred earlier finds more modern expression in the perfidy rule that prohibits killing, injuring or, for States party to API, capturing an enemy by acting in a way that invites the confidence of the enemy that he is entitled to receive or accord protection under the law of armed conflict, but with the intention of betraying that confidence.[35] Deception, however, is a long-standing technique in warfare. In the *Aeneid*, Virgil refers to the employment of a wooden model of a horse to infiltrate a Greek unit into the city of Troy after ten years of siege.[36] In more modern times, Operation Mincemeat[37] during World War Two was designed to lead the German High Command to believe that the focus of the allied attack in 1943 would be on Sardinia and Greece, whereas Sicily was where the planned attack would actually fall. The method of deception was to deposit a dead body bearing concealed papers disclosing the false plan.

The modern law recognises the lawfulness of ruses of war, including the use of camouflage, decoys, mock operations and misinformation. It is the act of inviting the confidence of the adversary as to matters of protection under the law that distinguishes unlawful from lawful deception operations. Moreover, the unauthorised use of the UN emblem or any use of flags, emblems, insignia or uniforms of the enemy or of neutrals is prohibited.[38]

The law of armed conflict prohibits certain weapons outright, including poisons, poisoned weapons, particular sorts of bullet, asphyxiating gas, biological and bacteriological weapons, chemical weapons, certain types of fragmentation weapon, certain kinds of mine, certain kinds and uses of booby-trap, blinding laser weapons and so on. Other legal rules restrict the circumstances in which particular weapons may lawfully be used, for example prohibiting the air delivery of incendiary weapons against military objectives located within a concentration of civilians.[39]

As the reader will have observed, much of the early law regulating hostilities was adopted in the Hague and came to be known as 'Hague Law'. Other provisions of the law of armed conflict, largely adopted in Geneva and thus known as 'Geneva Law', focus on the protection of those who are out of the fight. Such individuals include civilians in the hands of an adverse party to the conflict,[40] captured combatants and others with prisoner of

[33] API, Article 57(1).
[34] API, Article 58.
[35] API, Article 37(1).
[36] Virgil, *Aeneid*, Book II.
[37] B Mcintyre, *Operation Mincemeat: The True Spy Story that Changed the Course of World War II* (2010).
[38] API, Articles 38 and 39.
[39] For a discussion of the law of weaponry, see W H Boothby, *Weapons and the Law of Armed Conflict* (2009).
[40] The law relating to the protection of such persons and in relation to belligerent occupation of territory is to be found in Geneva Convention Relative to the Protection of Civilian Persons in Time of War, Geneva, 12 August 1949 and in API.

war status,[41] the wounded and sick on land,[42] and wounded, sick and shipwrecked persons at sea.[43] A more detailed consideration of these matters lies outside the scope of the present chapter.

ENFORCEMENT OF THE LAW

Breaches of the law of armed conflict are frequent and dreadful. An internationally famous UK journalist, Marie Colvin, was among a group of civilian journalists killed and injured as a result of a seemingly deliberate attack by Syrian government forces on the city of Homs on 22 February 2012.[44] Another example would be the threat by the then Libyan leader, Colonel Gaddafi, to assault Benghazi early in 2011 precipitating UN Security Council Resolution 1973 and the enforcement of a 'no-fly' zone over Libya.[45] Other examples could be cited, including notorious events at the Abu Ghraib Detention Facility in Iraq[46] and the killing of a detainee called Baha Mousa when in the custody of members of the United Kingdom's armed forces.[47] The vital point is that the law prohibits this and the law is enforced more frequently now than used to be the case.

Ethnic cleansing, murder, and widespread breaches of international law in the former Yugoslavia and in Rwanda led to the establishment of special tribunals to bring the perpetrators to justice. In 1998, and after several years of preparatory work and negotiation, a Diplomatic Conference adopted the Statute of an International Criminal Court.[48] It is the job of that court, when a State is either unable or unwilling to investigate and prosecute alleged serious breaches of the law of armed conflict, to consider commencing proceedings at the Court at The Hague.

When the author first lectured on the law of armed conflict issues at the RAF College at Cranwell, the law seemed to lack teeth and carried little credibility with at least part of his audiences. That is no longer the case. National courts, *ad hoc* tribunals and international courts all enforce the law, although it must be accepted that many crimes go unpunished.

THE CYBER DIMENSION

At the turn of the last century, hostilities were conducted essentially in two domains: on land and at sea. In the first half of the twentieth century, air power became the preferred means for projecting military force deep into enemy territory. In the second half of the twentieth century we saw the increasing use of outer space for both civilian and military purposes, and in recent decades a man-made environment which some characterise as an additional domain has been developed, namely cyberspace, and it is likely that cyberspace will be used for military as well as civilian purposes. Indeed, Distributed Denial of Service Operations against Estonian

[41] While there are some provisions relating to prisoners of war in the Regulations Respecting the Laws and Customs of War on Land, annexed to Hague Convention IV, 18 October 1907, articles 4-20, the main body of law on this topic is to be found in Geneva Convention Relative to the Treatment of Prisoners of War, 12 August 1949 as supplemented in API.

[42] The protection of the wounded and sick on land is provided for in Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, 12 August 1949.

[43] The protection of the wounded, sick and shipwrecked at sea, including airmen downed at sea, is provided for in Geneva Convention II, 1949.

[44] G Rayner and R Spencer, 'Syria: Sunday Times Journalist Marie Colvin killed in 'targeted attack' by Syrian Forces', *The Telegraph*, 22 February 2012 available at http://www.telegraph.co.uk/news/worldnews/middleeast/syria/9098175/Syria-Sunday-Times-journalist-Marie-Colvin-killed-in-targeted-attack-by-Syrian-forces.html.

[45] D D Kirkpatrick and K Fahim, 'Qaddafi Warns of Assault on Benghazi as UN Vote Nears', *The New York Times*, 17 March 2011 available at http://www.nytimes.com/2011/03/18/world/africa/18libya.html?pagewanted=all.

[46] For a summary of *The New York Times* coverage of the Abu Ghraib events, see http://topics.nytimes.com/top/news/international/countriesandterritories/iraq/abu_ghraib/index.html.

[47] A public inquiry was established in the United Kingdom to investigate the circumstances surrounding the death of Baha Mousa, and relevant documents can be accessed at http://www.bahamousainquiry.org.

[48] Rome Statute of the International Criminal Court, 17 July 1998.

websites in 2007,[49] cyber operations against Georgian websites in 2008,[50] and the Stuxnet attack on Iranian Centrifuges discovered in 2010 and apparently linked to that country's nuclear programme[51] are among a number of recent events that illustrate some of the possible military applications of cyber capabilities in an armed conflict.

THE TALLINN MANUAL

Considerable international interest, and not a little controversy, has unsurprisingly focused on the notion of cyber warfare and on the vexed question of which rules of international law, if any, apply to hostilities undertaken during an armed conflict using cyber means. It was therefore logical and appropriate that the NATO Cooperative Cyber Defence Centre of Excellence in Tallinn, Estonia, should sponsor for just over three years a project to write a Manual on the Law of Cyber Warfare. The author is a member of the Group of Experts, and of the Drafting Committee, that undertook that work. The Manual was published by Cambridge University Press in 2013[52] and represents the collective view of International Experts as to the *jus ad bellum* and as to the *jus in bello* rules that apply to cyber warfare.

The Tallinn Manual Experts recognised that there are no treaty provisions that explicitly deal with cyber warfare, and that because State practice and publicly available expressions of State legal positions on the matter are sparse, it is hard to deduce cyber-specific norms that would, because of their customary nature, bind all States. However, the Tallinn Manual Experts were clear that cyber operations do not take place in a legal vacuum, that principles of *jus ad bellum* and *jus in bello* both apply to such activities, and that the challenge would lie in determining how the rules would apply in the cyber context.[53]

As a matter of technical law, the Tallinn Manual does not constitute law in its own right. It comprises, rather, the 'teachings of the most highly qualified publicists of the various nations' referred to in the Statute of the International Court of Justice as a subsidiary means of determining the law.[54] The primary sources of the law remain, as reflected in that Statute, certain fundamental principles, the customary law and treaty law to which reference has been made earlier in this chapter. Nevertheless, the Manual will be a valuable tool for States and others to use in reaching their own conclusions as to the legal rules that should regulate military use of cyberspace in war.

SOME RULES FROM THE TALLINN MANUAL

It will be neither possible nor appropriate to summarise the 262 pages of the Tallinn Manual here. The author will therefore satisfy himself with selecting some rules that seem likely to be of greatest relevance to the wider

[49] E Tikk, K Kaska and L Vihul, *International Cyber Incidents: Legal Considerations* (2010), 18-25. Note also the DDoS operation on 26-28 April 2008 which targeted the website of Radio Free Europe/ Radio Liberty's Belarus's service, reported and discussed at E Tikk *ibid*, 39-48; and the cyber operation that targeted Lithuania on 17 June 2008, E Tikk *ibid,* 51-64.

[50] J Markoff, 'Cyber Attacks Disable Georgian Websites, Min of Foreign Affairs of Georgia', (*The New York Times Bits Blogs,* 11 August 2008) at http://bits.blogs.nytimes.com/2008/08/11/georgia-takes-a-beating-in-the-cyberwar-with-russia/; J Markoff, 'Georgia takes a Beating in the Cyberwar with Russia', (*The New York Times, Bits Blog*, 11 Aug 2008) at http://bits.blogs.nytimes.com/2008/08/11/georgia-takes-a-beating-in-the-cyberwar-with-russia/; European Union Independent International Fact Finding Mission on the Conflict in Georgia, Report (2009) and see also E Tikk, n.43 at pages 67-79.

[51] The reports of damage have not, so far as the author is aware, been officially confirmed by Iran. See J Fildes, Stuxnet worm attacked high value Iranian assets, BBC News, 23 September 2010 at http://www.bbc.co.uk/news/technology-11388018 and W J Broad, J Markoff and D E Sanger, Israeli Test on Worm Called Crucial in Iran Nuclear Delay, *The New York Times*, 15 January 2011, available at http://www.nytimes.com/2011/01/16/world/middleeast/16stuxnet.html?pagewanted=all.

[52] Tallinn Manual on the International Law Applicable to Cyber Warfare, 2013.

[53] Tallinn Manual, Introduction, page 5.

[54] Statute of the International Court of Justice, San Francisco, 26 June 1945.

topic of the November Conference and of the associated volume. It should, however, be emphasised that Rules in the Tallinn Manual should always be considered and interpreted by reference to the accompanying commentaries. The remaining paragraphs of this Section therefore constitute a series of signposts, nothing more.

Rule 6 notes that a State 'bears international legal responsibility for a cyber operation attributable to it and which constitutes a breach of an international obligation'. Significantly, however, the 'mere fact that a cyber operation has been launched or otherwise originates from governmental cyber infrastructure is not sufficient evidence for attributing the operation to that State, but is an indication that the State in question is associated with the operation.'[55] Routing a cyber operation through infrastructure located in a State is, moreover, insufficient evidence for attributing the operation to that State.[56]

The prohibition on the threat or use of force is found to extend to cyber operations that would constitute threats or uses of force for the purposes of article 2(4) of the UN Charter, and the Manual explains the application in the cyber context of notions of 'use of force' and 'threat of force'.[57] Similarly, the concepts of armed attack and of individual or collective action in self-defence are explored by reference to cyber operations in Rules 13 to 16.

Of critical importance was the consensus among the Tallinn Experts that '[c]yber operations executed in the context of an armed conflict are subject to the law of armed conflict.'[58] This determination, coupled with the finding of the Experts that a cyber operation, 'whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects' is a cyber attack,[59] has the immediate effect that the rules of the law of targeting that are recognised as applying to more traditional military operations also apply to corresponding military activities undertaken using cyber means.

Accordingly, the principle of distinction, the rule prohibiting indiscriminate attacks, the proportionality principle and the rules as to precautions in attack and against the effects of attacks are all found to apply to cyber operations in armed conflicts. Furthermore, persons and objects that are specifically protected by the law are also found to benefit from corresponding protection against the effects of cyber attacks and operations.

Lurking within the commentary accompanying Rule 41 is a definition of cyber weapons. The Rule itself talks of cyber means of warfare as comprising 'cyber weapons and their associated cyber systems',[60] so to understand the idea of cyber means one has to appreciate that cyber weapons are 'cyber means of warfare that are by design, use, or intended use capable of causing either (i) injury to, or death of, persons; or (ii) damage to, or destruction of, objects, that is, causing the consequences required for qualification of a cyber operation as an attack.'[61] The philosophical link between the application of the notion of *in bello* attack to cyber warfare and the development of the notion of cyber weapons is entirely logical. It is also important in that characterising certain cyber capabilities as cyber weapons has the evident effect of applying weapons law rules, including the superfluous injury and unnecessary suffering and the indiscriminate weapons principles, to cyber weapons. It also means that states in general are obliged to undertake some sort of legal review of all new cyber weapons that they acquire, and that states party to API are required to apply article 36 of that treaty to 'the study, development, acquisition or adoption' of new weapons, means or methods of warfare that they study, develop, acquire or adopt.

[55] Tallinn Manual, Rule 7.
[56] Tallinn Manual, Rule 8.
[57] Tallinn Manual, Rules 10-12.
[58] Tallinn Manual, Rule 20.
[59] Tallinn Manual, Rule 30.
[60] Tallinn Manual, Rule 41(a).
[61] Tallinn Manual, commentary accompanying Rule 41, paragraph 2.

All of this means that the ethical and moral principles that have underpinned the development of the law of armed conflict in general similarly underpin the rules that apply to cyber operations during armed conflicts, provided that States adopt the approach of the Tallinn experts. It is currently unclear whether States will in general agree with the Tallinn Manual interpretations. That is something that will only become clear over time, possibly a great deal of time.

At the moment it is sufficient to observe that, with the publication of the Tallinn Manual, some welcome and carefully considered light has been shed where previously there was darkness and confusion. Hopefully States will avail themselves of the opportunity the Manual provides to develop the law, or at least to clarify their national positions on these important issues.

CONCLUSION

There is a body of law that is grounded in realism and there are thoroughly logical reasons for saying that this body of law applies to cyber operations in the course of an armed conflict as it applies to more conventional operations undertaken in a similar context. The law takes the world as it is. It accepts the reality that civilians will get into harm's way, and simply requires attackers and those in control where attacks are liable to occur to do all they can to protect civilians and civilian objects. It includes specific protections for particular classes of person and object where these specific protections are warranted. A law that would seek to prohibit mistakes or technical malfunctions of persons or of weapons would make no sense and, to its credit, the law does not do so.

In fact what the law does, or at least tries to do, is to strike the necessary balance that the early writers identified between the military needs of those fighting the war and the humanitarian needs and concerns of its victims, be they civilians in the battle zone, wounded, sick, prisoners and so on. That balance is a fine one, and yet it is essential that it be maintained. Why? Because if we don't maintain the balance, if humanitarian-inspired prohibitions become too extensive and start to impede the proper conduct of warfare, the law risks being increasingly ignored by those whose prime responsibility it is to uphold it, because they will regard it as unrealistic and therefore irrelevant. That would be a dreadful thing, because the law that protects the vulnerable would be put in jeopardy, and unrestrained barbarity might be the unintended and awful consequence.

It therefore follows that any suggested ethical rules that go beyond the law of armed conflict rules referred to in this chapter would also have to strike the same balance. To be clear, the author's view is that the focus should always be on seeking to ensure that all those involved in armed conflicts abide at all times by the law of armed conflict. It is the author's opinion that strict legal compliance would do much to reduce civilian casualties and loss and would contribute positively to the re-establishment of peace. Achieving that high degree of legal compliance pre-supposes that members of the armed forces thoroughly understand what the law comprises and what it requires of them. We should therefore be careful to ensure that any talk of additional ethical requirements does not create confusion. Perhaps it is only if the rules to be complied with are simple and straightforward that soldiers will in practice be able to adhere to them in the difficult, often ambiguous and always stressful circumstances of combat.

The Applicability of the Just War Tradition to Military Cyber Operations

Edward T. Barrett
Stockdale Center for Ethical Leadership
United States Naval Academy
United States

ebarrett@usna.edu

INTRODUCTION

This paper argues that the traditional *jus ad bellum* and *jus in bello* criteria are fully capable of providing the ethical guidance needed to legitimately conduct military cyber operations. The first part examines the criteria's foundations by focusing on the notion of liability to defensive harm worked out by revisionist just war thinkers. The second part evaluates ethical issues germane to responding to cyber force, including *casus belli*, moral aspects of "the attribution problem," and respective rights and duties when attacks involve innocent third-party states. The third part addresses *jus in bello* issues, including compliance with discrimination, necessity, and civilian due care imperatives; whether civilians may be targeted with sub-"use of force" cyber-attacks; and the permissibility of using civilian contractors to conduct cyber-attacks. Throughout these analyses, conclusions are brought into conversation with those of *The Tallinn Manual*.

JUST WAR CRITERIA AND THEIR FOUNDATIONS

On their face, the criteria seem simple. War must be the last resort available to a proper authority intending to pursue a just cause, and circumstances must indicate a reasonable chance of succeeding in a proportionate manner. Once in war, only combatants may be targeted, intentional harm to combatants must be necessary, and unintentional harm to civilians and their property must be minimized and then proportionate.

In service of accurately applying these criteria to cyber issues, we must first recall their foundations. Except for right intention, the criteria can be grounded in the assertion that persons normally possess a dignity that renders the preconditions to their flourishing—their rights—inviolable. In the context of war, human rights mean that aggression is unjust, and that defensive uses of lethal force are justified only if an aggressor has forfeited their right to life. This 'principle of forfeiture' recently has been developed under the rubric of 'liability to lethal defensive harm.' For our purposes, four aspects of these principles of forfeiture and liability deserve emphasis.

First, lethal defensive force is justified only in response to grave acts such as murder. However, less serious acts—stealing, for example—can proportionately diminish an actor's dignity and rights, and thus incur liability to non-lethal responses. Second, lethal defensive force is justified only in response to culpable grave acts: those done intentionally, freely, and with knowledge or vincible ignorance of the act's injustice. While a lethal response to non-culpable aggression may be excused due to a responder's epistemic limitations, a lethal response is not justified. Third, murderous intentions must be accompanied by preparation to act in order to compromise one's right to life. Fourth, since even murderous acts do not eliminate human dignity, aggressors may be harmed only as necessary for either the defence of potential victims or the aggressor's reform.

These ethical foundations generate and allow one to apply the *ad bellum* and *in bello* criteria. A just cause exists only when adversaries are culpably attempting, or actively preparing, to gravely harm. Although such aggressors have forfeited their claim-right to life at this point, their worth requires that any defensive harm be necessary. Ramifications of this necessity requirement include last resort *ad bellum* and necessity *in bello* vis-à-

vis combatants. If necessary uses of force will unintentionally harm civilians, defenders must—consistent with the principle of double effect—take adequate measures to minimize such harm, and foresee a reasonable chance of attaining the justifying condition in a proportionate manner, both before and during the conflict.[62]

From an ethicist's perspective, the *Tallinn Manual*'s international law-based treatment of the issue of proportionality is especially commendable.[63] While some ethicists argue that proportionality applies to tactical situations involving only combatants, the Manual correctly applies it to situations involving direct and indirect collateral damage. Also, I agree with the majority's position on the issue of collateral damage probability: in calculating proportionality, lower levels of probability should not decrease the value of collateral damage.[64] However, two quibbles are in order. First, the section discussing the definition of 'excessive' collateral damage correctly states civilian casualties should be compared to military advantage, not enemy combatants rendered *hors de combat.*[65] But 'military advantage' needs to be defined more precisely. Since 'advantage' in this context is relative to the just cause, and usually means innocents defended, I would recommend that cashing out proportionality in terms of 'innocents saved versus innocents lost.' Second, since adversary civilians retain all of their rights, I would not dismiss consequences such as fear or loss of email or banking services from collateral damage calculations.[66] Of course, such costs could be awarded relatively low values.

JUS AD BELLUM ISSUES

These ethical criteria—not merely state practice *vis-à-vis* international laws—must be applied to the four key issues concerning military responses to cyber activities: just cause, attribution, third-party rights, and proper authority.

Just cause

Because the principle of forfeiture governs permissible responses to all interpersonal harm, the point at which kinetic and cyber-attacks constitute a *casus belli* is the same.[67] But unique characteristics of cyber activities generate at least three questions concerning what is sometimes called "the threshold problem."

First, can an accumulation of events, which individually would not be sufficient reason to use lethal force, constitute a just cause? It depends. On the one hand, successive cyber intrusions that combine to merely reduce living standards do not result in liability to lethal force. But serial thievery is nevertheless unjust and, as in domestic contexts, subject to non-lethal responses. On the other hand, culpably seeking to cause low-level harm as part of an existentially-threatening campaign would create a liability to lethal kinetic or cyber force. However, in these cases, uses of force would be pre-emptive, and thus raise obvious concerns.

---

[62] To highlight two assertions: reasonable chance of success and proportionality pertain when civilians will be affected; and both criteria must guide *ad bellum* and *in bello* decisions.

[63] *Ibid*, Rule 51, pp. 159-164

[64] *Ibid*., p. 163

[65] *Ibid*., p. 161

[66] *Ibid*, p. 160. See also Michael Schmitt, 'Five Myths in the Debate about Cyber War,' 23 September 2013 at 11:03 am at www.justsecurity.org

[67] Legal analyses of just cause in the cyber domain have focused on defining an "armed attack," to which UN Charter Article 51 allows self-defensive responses. Death/injury/destruction/damage-causing "uses of force" of sufficient —using Pictet's taxonomy— "scope, intensity, or duration" are armed attacks. In order to transcend debates over "sufficient" effects (assuming parties can agree on the validity of an "effects-based" approach, rather than "instrument-based" or "strict liability" approaches), this article will prescind from this legal framework. For a helpful discussion on these legal tools, see Jeffery Carr, *Inside Cyber Warfare* (Cambridge: O'Reilly Media, 2010), pp. 45-75.

The Manual's treatment of accumulated effects—although brief and vague—coincides with this analysis.[68] A series of small-scale incidents are said to constitute an armed attack if they 'taken together have the requisite scale.'

Let's turn to a second question concerning just cause: can temporary losses of computer functionality constitute a *casus belli*? This one is pretty easy. While permanent losses of functionality create the same effect as physical destruction, temporary losses are unique to cyber operations. In these cases, justified responses would be a function of both an attack's culpability and effects. Assuming culpability, a temporary loss of functionality that resulted in loss of life would be a *casus belli*. A brief interruption of air traffic control services could amount to a just cause, while equally brief interruptions of electricity and especially ATM services probably would not.

The Manual's treatment of temporary functionality losses is self-contradictory. In one section, it argues that interference with functionality would qualify as a use of force only if the 'restoration of functionality requires replacement of physical components.'[69] But another section—on just cause—recommends an effects-based definition of "armed attack" consistent with my analysis.[70]

A third just cause-related question: may anticipatory self-defence be invoked in cases of cyber espionage or logic bombs? Legally, espionage is neither a just cause nor a war crime. And except for their locations, logic bombs posing an existential threat resemble nuclear weapons, whose mere possession has never been deemed a cause for war. But once aware of such intrusions, potential victims will find it impossible to determine when otherwise non-lethal activities are about to be elevated. Imminent attacks will be undetectable. Put this way, I think the question answers itself. As argued earlier, persons who actively intend to lethally aggress nevertheless should not be harmed unnecessarily. And traditionally, imminence has been considered a valid indicator of both intention and necessity. But in situations—such as cyber espionage and logic bombs—where imminence will be undetectable, certainty about active intentions to cause grave harm would justify anticipatory self-defensive measures. However, pre-emption in practice is morally hazardous for epistemic reasons, and is rarely justified.

The Manual's assessment of anticipatory self-defence comports with mine, including its caution.

Attribution

Assuming that a just cause exists, the attribution problem adds another complexity to responses. Technical aspects of computers and the internet can undermine certainty about a perpetrator's location, equipment, identity, and/or institution—and can even implicate innocent parties. Circumstantial evidence may be the only link to attackers.

At least two attribution-related questions arise. First, what degree of certainty about an attacker's identity must be attained before responding with lethal or non-lethal force? On this issue, I would emphasize the need for absolute certainty. Even if deterrence and utility would be maximized by responding when less than certain, it would be unjust to intentionally harm innocent parties—whose rights cannot be sacrificed for the greater good. [71] It would also be inexcusable to do so. Since circumstantial evidence does not adequately

[68] Schmitt, *Tallinn Manual*, p. 56
[69] Ibid., p. 108
[70] Ibid., pp. 56-57
[71] Randall Dipert, "The Ethics of Cyberwarfare," *Journal of Military Ethics* (2010) pp. 384-410. This article is the seminal piece on the ethics of military cyber issues.

establish certainty, and may be all that is available, denial-based deterrence of future incidents may the only legitimate response.[72]

A second and more vexing question: assuming a successful identification, what if an attribution delay discloses that a response would no longer be a defensive act? Instead, one would be punishing wrongdoers—a very different just cause. Given time constraints, I'll merely offer my conclusion. Since the principle of forfeiture indicates that lethal force may be used only to defend against culpable and grave harm, the validity of punishment as a just cause must be evaluated by this standard.[73] Legitimate cases of punishment are thus defensive and akin to pre-emption. Past aggressors who remain grave threats also remain liable to lethal force if necessary. But those posing no threat are liable to reform-oriented punishments and owe compensation, but are not liable to lethal force.[74]

Innocent third party states

Ethical challenges remain even after threshold and attribution problems are solved. Since cyber-attacks may emerge from, or be routed through, innocent third party States, decisions must be made about the respective rights of attacked and innocent third parties. Specifically, at what point does an attacked State's right to defend itself trump the sovereignty of a nation-state that is the geographic source, but not the cause, of a lethal attack? This question is not new: ongoing 'targeted killing' operations respond to threats emerging from non-complicit States, and aircraft overflight rules address attacks routed through third party States. But computer network capabilities have increased the likelihood of such situations.

An ethical analysis of these situations must include an analysis of nation-state sovereignty, the foundation of which is the liberty-right of individuals to associate for political purposes. This liberty-right generates the duty of others to not interfere. But as 'responsibility to protect' advocates have argued correctly, States that harm their own citizens can lose their right to non-interference. Additionally, and more germane to the question at hand, States have external obligations. They are morally obligated to protect other States' citizens from threats emanating from or traversing their jurisdictions. States unwilling to meet this external obligation, or who are unable and refuse necessary assistance, possess a degree of culpability and thus lose their right of non-intervention; and threatened States have a corresponding right to intervene in order to locate and/or retaliate against attackers. In these situations, the principles of forfeiture and necessity still govern actions toward all rights offenders. If the offense warrants lethal force, and if capture by the host or victim State forces is impossible, lethal force may be used. Of course, retaliation against States themselves is warranted if attribution efforts prove collusion with these citizens.

Cyber-attack victims should be guided by the same criteria. When necessary, lethal kinetic or cyber force may be used by States to defend against ongoing or anticipated *casus belli-level* cyber-attacks emanating from or traversing through another State. But such responses should be preceded by the opportunity for host State enforcement and/or cooperation, and are impermissible if adequate enforcement and/or cooperation efforts have occurred. In trickier situations, when an immediate response is necessary, it should be pre-negotiated with or approved by the third-party State, which also should be adequately compensated for damages.

---

[72] Dipert rightly notes that the "sum or compound attribution method" would not merely employ technical means. However, non-technical components of this method are likely to be inconclusive (i.e., using electronic and human intelligence sources to verify an attacker's identity) or are morally irrelevant (i.e., inferring from means and motive). See Randall Dipert, "Other-than-Internet (OTI) Cyberwarfare: Challenges for Ethics, Law, and Policy," *Journal of Military Ethics* (2013), pp. 34-53.

[73] Lethal force does not reform, which is a second legitimate justification for punishment.

[74] I am obviously calling into question the legitimacy of punishing for retributive and utilitarian (i.e., deterrence and public order) purposes, and would argue that public safety and reform should be the guiding purposes.

Proper authority

Finally, I want to mention three ill-advised innovations that could undermine compliance with the 'proper authority' criterion. First, business firms should not be allowed to use active cyber defences that strike network targets inside another state, and thus could embroil their state in an unjust war. Second, even if deemed reliable for *in bello* use, automated cyber countermeasures that strike inside another state should not be used prior to an existing conflict. While operationally advantageous, these systems should not be put in a position to start wars. Third, the use of private military contractors should be carefully managed. More on this issue later.

JUS IN BELLO ISSUES

In addition to these *ad bellum* issues, the application of cyber force presents questions about compliance with the *in bello* constraints of discrimination, military necessity, and due care for civilians. Added to these core *in bello* concerns are debates about whether some cyber-attacks may be intentionally directed against civilians, and about the status of civilian contractors.

Discrimination

Notwithstanding the apparent success of the Stuxnet virus, some have argued that tight linkages between military and civilian networks, and cyber weapons' inability to distinguish between legitimate and illegitimate targets, will combine to frequently render cyber-attacks indiscriminate. Kinetic munitions are subject to well understood laws of gravity, motion, and aerodynamics, and are therefore capable of striking an intended military target. While their accuracy is imperfect and their post-impact effect on the surrounding space can be difficult to predict, they can at least be aimed. But the mechanisms through which cyber weapons move and affect computers are not widely understood, and one can imagine ubiquitous effects even worse than those of chemical weapons.

However, assessments of these weapons' capabilities and morality by academics, and by most military and civilian leaders, are likely to be unreliable. Even those possessing an understanding of how these systems might operate are unlikely to possess the requisite high-level security clearances that would allow them to evaluate specific systems. But thus far, from the perspective of concerned ethicists who are cognisant of applicable normative constraints but not system capabilities, the news is good: experience shows that cyber weapons are not necessarily indiscriminate and thus illegal. The Stuxnet virus, for example, despite penetrating both military and civilian networks, only affected military targets. Unfortunately, because of necessary classification levels, we do not—and will not—have many examples. Therefore, the ethicist's roles will be to ensure the *in bello* criteria are understood by users of cyber weapons, and to insist that users are duly diligent in ascertaining technical capabilities and relevant aspects of situations in which they are used.

Military necessity and civilian due care

Even if cyber weapons are discriminate, questions of unnecessary harm to combatants and disproportionate harm to civilians remain. Since well-tested, human-launched kinetic weapons operate within natural, stable, and relatively knowable conditions, their effects are predictable. For example, the abilities to thoroughly test aircraft-launched precision guided munitions under controlled conditions, and then discern actual conditions when using them operationally, allow users to confidently foresee the affected area and damage within it. But cyber-attack effects may be highly unpredictable due to their manmade and thus changing cyber environment. While mock-ups of targeted systems can be constructed, controlled tests on actual adversary networks are obviously impossible. And a perfect mock-up could not eliminate the uncertainty associated with unknown changes that adversaries make to their networks. Short of perfect, real-time intelligence, attacks may 'impact' their target, and then create unexpectedly catastrophic and unjust results.

Nevertheless, as with discrimination, experience again shows that the intended and unintended effects of these weapons are foreseeable, and that cyber-attacks are therefore capable of complying with military necessity and civilian 'due care' constraints. In some cases, their use may be more ethical than kinetic options, and thus obligatory. To these ends, users of cyber weapons must, again, exercise due diligence in determining their likely effects—which includes adequate weapons testing and adequate intelligence on adversary computer network configurations.

However, rapidly-responding autonomous lethal systems lacking a 'human-in-the-loop,' while operationally advantageous, should be cautiously pursued. While the speed required to defend in the cyber domain may favour autonomously deployed 'active defences,' ethically adequate countermeasures would have to accurately verify targets and determine which strike options would best satisfy *in bello* criteria.

Until these systems are thoroughly tested in credible conditions (which is probably impossible), potentially injurious or deadly countermeasures should be approved by qualified persons.

Targeting of civilians

Echoing ongoing debates over the use of non-lethal weapons, some have suggested that cyber-attacks that do not strictly qualify as 'uses of force,' and which merely inconvenience, may be intentionally directed against civilian objects. According to this argument, international law stipulates that civilians and civilian objects should not be 'the object of attack,' but that attacks are defined as violence causing death, injury, damage, and destruction.[75] In other words, operations that do not use 'force' may be directed against civilians and their objects. An additional assertion, unique to the cyber domain, is that civilian data not transferrable into tangible objects is not a civilian object, and may thus be intentionally targeted.

These conclusions, which follow logically from their premises, demonstrate the danger of purely legal approaches to these issues. In addition to their likely ineffectiveness, such attacks would be unjust. In peacetime or wartime, persons who have not intentionally transgressed others' rights have forfeited none of their own and are not liable to any degree of harm, including inconvenience resulting from data destruction. In a domestic context, such intentional harms would be punishable crimes; in war, they would constitute relatively minor but nevertheless punishable war crimes. When committed by soldiers in war, these acts would undermine the cultivation of warriors disposed to appropriately respect the lives of combatants and, especially, of civilians. But in cases where indirect civilian participation in aggression obviates liability to lethal defensive harm but does incur liability to non-lethal responses, cyber capabilities create options that deserve further moral evaluation.

The *Tallinn Manual* assesses this issue differently. Given existing legal definitions, cyber operations that do not rise to the level of an attack may be directed against civilians. Interference with the functionality of cyber infrastructure is permitted if physical repair (a majority view) or operating system reinstallation (minority view) are not required.[76]

Status and use of private security contractors

The status of civilians who directly participate in aggression is obviously a different matter. At a certain, albeit difficult to define, point, civilians who are affiliated with aggressors and 'directly' participate in aggression

[75] See UN General Assembly (1977), especially articles 51 and 52, and also articles 35 and 55-57.
For an excellent legal perspective on this issue, see Schmitt (2011), especially pp. 114-123. My ethical analysis supports the more restrictive conclusions of Knut Dormann.

[76] Schmitt, Tallinn Manual, pp. 108-109. See also Schmitt, 'Five Myths in the Debate about Cyber War;' and Michael Schmitt, 'International Law in Cyberspace: The Koh Speech and the Tallinn Manual Juxtaposed,' Harvard International law Journal Vol. 54 (2012), pp. 26-27.

become liable to lethal force. While the Red Cross has defended a 'revolving door' concept whereby intermittent direct participants would regain their protected status immediately upon ceasing specific hostile acts, this position is unconvincing.[77] If such civilians have well-formed intentions to commit future acts, they are continuously targetable. Like their uniformed colleagues, these civilians arguably may be targeted even when 'off duty.'

CONCLUSIONS

The *Tallinn Manual* indicates that the group was split on this issue, with some members adopting the 'revolving door' notion of liability, and others rejecting it.[78] The document is not clear whether directly participating civilians may be targeted at home; only periods of travel to and from an operational location are discussed.[79]

Assuming that states are willing to tolerate this blurred line between their combatants and non-combatants, another question arises: are civilians appropriate agents of legitimate lethal acts? As the recent debate over the use of private security contractors (PSCs) highlights, the litany of *in bello* concerns includes whether they, even in combat support roles such as intelligence analysis and combat training and advising, result in unjust levels of adversary combatant and civilian harm.

While the use of abundant, computer savvy civilians as attackers is economically attractive, I believe culture-based concerns are valid. As virtue ethicist Alasdair MacIntyre has argued, just as the human *telos* defines and requires properly human virtues, one's social purpose will also define and require virtues appropriate to that particular function. Since the purpose of the liberal state is to secure a particular subset of nevertheless universal rights, dispositions proper to soldiers will be those conducive to securing the good of all persons, including enemy non-combatants. On the other hand, the purpose of a business is to maximise market share by satisfying the desires of consumers able to pay a market price. Accordingly, the virtues corresponding to successful business activity are not necessarily those of the soldier; selflessness may be a vice when acting within this sphere. To be fair, some soldiers lust for domination, and many security contractors are willing to risk themselves to benefit even unknown civilians. But at least in their professional capacities, soldiers are more likely to be solicitous of human rights than (to borrow Plato's term) producers.

For this reason, the approval authority for the use of potentially injurious or lethal force—including cyber force—should be reserved for members of the military. Additionally, threat analyses and training and advising should be supervised by the military, since these activities affect the use of force. The contributions of security and combat support contractors will be essential to military cyber operations, but cultural and character differences require military control of these functions.

[77] See Nilz Melzer, Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law (Geneva: International Committee of the Red Cross, 2009), especially pp. 70-71.
[78] Schmitt, Tallinn Manual, pp. 121-122
[79] *Ibid*., pp. 120-122

Distinctive Ethical Issues of Cyberwarfare

Randall R. Dipert
Department of Philosophy, University at Buffalo
New York
United States

rdipert@buffalo.edu

In several articles published since 2010, I laid out what I consider the foundations of the ethical issues of cyberwarfare. Even before then, two non-professional philosophers had written on the subject (Arquilla, 1999; Rowe, 2007, 2009) and professional philosophers had contributed to it (Owens *et al.*, 2009). Since then, a number of authors have added significantly to this strictly ethical literature (Stawser, 2013). Discussions in international law predated the ethical literature by more than a decade (Schmitt, 1998) and, with the Tallinn Manual (Schmitt *et al.*, 2013), have reached a more systematic state than has the ethics of cyberwarfare. I will say something later about the difficult issue of the difference and relationship between ethical and legal considerations.

One key issue in ethics is whether cyberwarfare raises distinct and new issues that cannot be addressed by traditional Just War Theory and other theories of the morality of war. I have argued that there are several distinctive ethical issues in cyberwarfare. One is that the *quantity* of harm inflicted by a cyberattack will often not rise to the level of traditional 'kinetic' weapons.[80] It will be more like the harms that have been called 'measures short-of-war', such as sanctions or boycotts. Secondly, the *quality* of the harm is will differ from that caused by traditional weapons. Killing or injuring human beings may be absent, as will the permanent physical destruction of other entities. These harms will instead often be to the *functioning* of information systems and of financial, energy, communication and other sectors of an economy. A third way in which cyberwarfare is unique is epistemic: we will often not know immediately who attacked us. This is the well-known Attribution Problem, although I believe it is on the way to being technically solved, and in any case need not lead to complete and indefinite inaction. I have argued elsewhere (Dipert, 2010) that this makes defending oneself against cyberattacks much like the problem of preemptive or preventive war.

Traditional theories of morality in war arise from ethical values that are widely shared across cultures: it is generally wrong to kill human beings and destroy the physical entities that are necessary for human well-being. It is still more wrong to kill large numbers of people or cause widespread, permanent destruction. Some careful definitions of war (Orend, 2005) have stipulated that death and damage in war (as opposed to warfare) must be 'widespread'.

However, many instances of cyberwarfare arise in far more slippery moral terrain: intrusion into information systems, exfiltration of data (cyberespionage) and 'theft' of intellectual property, as well as placing software in an enemy's information systems that could eventually be used to cause harm but needn't have that purpose and might remain inactive. Although many users are uninformed about the extent to which their information systems can be 'read' by others using the internet, they *could* know and are, after all, voluntarily connecting themselves to the internet. They are connecting themselves to the informatics-analogue of a pipe delivering untested and unguaranteed water that should be more clearly labelled 'possibly not drinkable.'

The defenders of U.S. government and military infrastructure have over-dramatised the supposedly precarious position of technologically advanced States as being subject to thousands or hundreds of thousands of 'attacks' per week or even per hour. Yet very few of these so-called attacks do any damage at all, and so the word

[80] This simple thesis has been recently drawn into a book-length treatment in (Rid 2013).

'attack' is misused. They fall into different categories, with different ethical implications. They are often clumsy, automated attempts to hack into information systems; these do no damage unless access is gained and an attempt is made to exploit this access. I will discuss this at some length. Many are either acts of cyberespionage or what is called the 'theft' of intellectual property such as aircraft design. Espionage has never been considered a reason by itself for going to war, and the only effective remedy lies with the spied upon State (Goldman, 2010).

The theft of intellectual property is likewise not an action to be answered with a kinetic or cyber military response. It is a complicated matter of international patent and trade agreements. If the Chinese, for example, steal our non-military industrial secrets, then they should not be allowed to market products made from them elsewhere in the world. If they are stealing our defence secrets, they are behaving as they and we should out of self-interest. One bit of damage these 'attacks' do is to raise the cost to the defender of protecting against them. The cost of enhanced cybersecurity to governments, corporations and individuals in the future is likely to be staggering. Now estimated in the tens of billions in the U.S., it is likely to be hundreds of billions in the near future. Intrusions into information systems that intentionally cause quantifiable damage are another matter, and the only serious one for the morality of warfare. In short, few of these so-called attacks are illegal, and it is not even clear they are seriously immoral.

Unlike traditional warfare, many forms of cyberwarfare do not involve the intrusion of physical objects or human agents into a State's territory. This fact is striking and marks yet another difference with traditional warfare. Intended harm is accomplished by the conveyance of information entities from one information system to another. As I have argued elsewhere, there are important forms of cyberwarfare, broadly defined as nation-on-nation intended harm to information systems, which may be accomplished by means other than the internet (Dipert, 2013). Far too little attention has been given to this other-than-internet information warfare.

The non-kinetic properties of most forms of cyberwarfare are significant because the Just War Principles and virtually all of international law have been interpreted in modern times through the lens of what may be called Westphalian intermediate principles. By Westphalian principles I mean the linked notions of sovereignty and territory, and permitted and non-permitted activities in that territory, described in the Treaty of 1648. So the modern interpretation of the Just War criterion of 'Just Cause' paradigmatically involves invasion by armies, that is of organised, armed human beings travelling into another State's sovereign territory, or physical destruction caused by physical objects such as arrows or cannon shells entering into territory. As with artillery shells and missiles, drones allow for intentional destruction in another State's territory without the need for human agents that might be captured and held responsible.

THE LEGAL AND ETHICAL ISSUES CONTRASTED

The legal and moral questions of war differ in interesting and complicated ways. The details of what counts as a law in international law is similar to that of most domestic legal systems. There are a kind of *statutes*, namely treaties, as well as precedents, and often unwritten customary prohibitions.

The fragility of judgments about legal aspects of cyberwarfare is immediately demonstrated by what appears to be the inherent vagueness of the meanings of 'force,' 'threat of force,' 'armed' and 'attack' in the U.N. Charter, and whether and how they apply to cyberattacks. The literal reading of 'arms' would be something like 'physical artefacts whose main purpose is killing or destruction.' This does not apply literally to the hardware used in cyberwarfare, since it consists of general purpose information-processing equipment. It is the software that is specifically adapted to warfare and espionage, but these information-theoretic entities are notoriously difficult to analyse. Likewise the term 'objects' in the widely accepted Addition Protocol I to the Geneva Conventions of 1977 gives examples which are all physical objects.

The ethical account of cyberwarfare would grant no privileged status to the UN Security Council except as a matter of procedural, but not substantive, justice. The Security Council, having passed a resolution, is neither a

necessary nor sufficient condition for an act of cyberwarfare being ethical. The international courts and the UN have a kind of moral force that derives from the status of treaties as promises of signatory States, but the treatment of States as having made binding promises over decades and through dramatic changes in their forms of government derives from a very complicated theory of agency, and of the identity and endurance of States, and is again Westphalian.

Ethical judgments are most often understood by ethical theorists as being based on basic, foundational principles of widespread application. Examples of such basic principles are the maximisation of the sum of the well-being of present and future humans, known as utilitarianism; or of one's own rationally considered well-being, known as enlightened egoism; or the promotion of universalisability. What if everyone acted like that? In certain areas of application, such as the 'Hippocratic' background of medical ethics and in the ethics of warfare, there had arisen widely agreed upon ethical principles before modern ethical theories applied their axiomatic approach to ethics. This gives such principles in applied ethics, such as the Just War conditions, a problematic status. They are not derived, in any obvious way, from truly basic ethical principles. In Aquinas and Grotius they were grounded in an Aristotelian theory of natural law or natural rights; however, few modern ethical theorists have endorsed that conception. In practice, this has brought modern ethicists to mix and match various foundational principles with traditionally accepted intermediate criteria like Just War Theory, and with intuitions or what seem to be conclusions of historical study, as elegantly practiced by Walzer (2006).

The use of intuitions is particularly problematic. While we might have intuitions—our ethical consciences—about the ways in which individuals should behave if they were in a moral dilemma, there is far less reason to think that we have decision-making skills, and developed intuitions as leaders of States responsible for many, many lives.

Ethical theories of war have oddly ignored certain game-theoretic results (Dipert, 2006a, 2006b) and so they have generally rejected 'realist' approaches (Christopher, 1999; Walzer, 2006) that are one of the major schools of international relations. In particular, philosophers have rejected deterrent strategies, including Mutually Assured Destruction, while most geopolitical thinkers and leaders have endorsed them. Although deterrence is always a complicated phenomenon, with many preconditions for success (Libicki, 2009), it would seem that it would be exceptionally usable in cyberwarfare, at least between rational cyberpowers. Deterrent strategies in cyberwarfare do not have the serious failures when measured against Just War Theory that they have in the nuclear case, where they fail the *Probability of Success* and *Proportionality* conditions. It also seems that in cyberwarfare, as much as we can generalise from relatively few cases, we are heading toward a game-theoretic equilibrium in which certain limited cyberattacks and extensive cyberespionage are tolerated.

The lack of an objection by Iran to Stuxnet in forums such as the Security Council and international courts, as well as the silence of the other major cyberpowers, indicate a tacit international acceptance of Stuxnet, perhaps as a limiting case of the most severe cyberattack that would be tolerated. George Lucas (2013), in a wise and perceptive essay on permissible cyberattacks, formulates criteria that more-or-less conform to the Stuxnet case and its apparent ethical acceptance. Avoiding an unstable escalation nevertheless remains a difficulty for deterrent strategies.

APPLYING JUST WAR THEORY AND OTHER MORAL PRINCIPLES TO CYBERWARFARE

Just War Theory and its variants can only be taken as intermediate guiding principles or rules of thumb, since they lack a derivation from foundational principles. Another source of legitimacy is that they have some status as conventions that have come to be widely accepted, and that might limit the damage of war if everyone abides by them. Of the four core principles of Just War Theory for going to war and initiating the use of force, two are especially problematic for some forms of cyberwarfare.

The four are:

1) Just Cause
2) Last Resort
3) Probability of Success
4) Proportionality

The widely accepted 'high' barrier for Just Cause, namely armed invasion by an enemy with an intention to use lethal force, does not seem to apply to many forms of cyberwarfare. Likewise, cyberwarfare is not necessarily a *last* resort. That would continue to be the use of lethal force or force that brings extensive permanent destruction. Some forms of cyberwarfare would fall in the next-to-the-last resort category, such as threats and ultimatums, sanctions, unilateral breaking of diplomatic and economic ties and so on. Modern ethical and legal theory has largely ignored these smaller acts of intentional harm.

A just cause for war has never included another State's distribution of misleading or faulty information, conveyed in human-to-human communications. This would simply be 'disinformation.' However internet-based injection of malware can be described in terms that involve the unwitting and undesired transfer of information. This develops the useful insights of Floridi and Taddeo that place cyberwarfare in the wider landscape of information warfare.

No person familiar with the technology could think that information coming through the internet is protected by diplomatic conventions or international principles. By connecting oneself to the internet, one knowingly opens one's own information systems to all manner of information, disinformation and noise. A convention might arise in which certain, ideally encrypted, messages are protected by their status from examination and manipulation. However the history of espionage, and especially of the morality and legality of espionage, seems broadly to permit examining and even manipulating another State's messages without incurring a justified armed attack.

Especially instructive is a careful moral examination of 'hacking into' a website. The maker or owner of the website might not desire non-authorised users even to access the public webpage, although this is inconsistent with using the internet to make it visible. If access to certain webpages and the ability to alter information is password-protected, we have a scenario that raises clearer ethical issues. Note that it is actually fairly rare to encounter government, corporate or individual public webpages that have warnings against their improper use by unauthorised personnel. Partly this would betray a naïveté and even inconsistent thinking about the internet. Without enforced statutes, or ways of pursuing or even correctly identifying non-citizen violators, there have not been even nominal attempts to separate permissible from non-permissible activities.

Various kinds of access to information via the internet can be described. It is clear that many of these forms are *undesired* by the owner, who might also have good reason to believe that no hacker will break through the protective barriers. The owner might declare that unauthorised users may not attempt to hack into the system, but with what moral and legal force? That is, when is a hacker doing something unethical, and why is it unethical?

A search for useful moral analogues is difficult. A business hangs an 'Open for Business' sign or otherwise gives indications of its entry conditions, with windows displaying goods, a description of the goods to be obtained there and perhaps an 'Enter here' sign. If the door is locked, then by convention one may not try to obtain entry, despite the 'Open for business' sign. If the door is unlocked, then one may reasonably enter the store and look around. If the goods are openly displayed, and absent a sign to the contrary, one may pick them up. The store will likely have a declared 'public' area as well as a private one where the general public may not go. The rules of information gathering are not so strict. If I can stand in the public area and see a sheet of the store's accounts on a desk, then perhaps I would be rude to attempt to scrutinise it, but it is unlikely that it

constitutes an illegal or even an unethical deed to read the numbers. The grocer does not expect or want me to see this information, but that fact alone does not constitute a strong case that it is unlawful or unethical.

Unfortunately, although this is the best one can do in the way of moral analogues, there are some failures of the analogy. For one thing, being private property, this is a case in which there is physical space governed by rules and conventions much as in case of the application of Westphalian principles. This physical space is owned and there are conventions governing what I may do in it. Secondly, there are elaborate conventions governing these various permitted and unpermitted activities—where I may go, how far I can reach my arm, what I may do with the merchandise and so on. It has proven beastly difficult to build artificial intelligence systems that understand these transactions and the background information governing behaviour. Normally, for example, I may not eat the products in a grocery store, but if there are small pieces of, for example, pastries which have been so displayed as to allow easy access by customers, then they are probably samples that may be eaten.

The concepts of an owned space, whether for private land like a store or for a State's territory, evolved over centuries, if not millennia. They are based on a shared and literal notion of space, and of boundaries in that space. By comparison, the ethics of cybersecurity have developed very recently, with almost laughably undeveloped concepts. There are few clear customs and very few clear, enforceable, extraditable laws.[81] It uses a metaphor of space, 'cyberspace,' but without key structural features of space. What counts as 'movement' and thus intrusion into another agent's cyberspace? And most troublesome of all, what are the acknowledged or declared boundaries of one person's, or one State's, area (or volume) of cyberspace?

In order for access to another's data to count as unethical we would need generally acknowledged principles of where the 'borders' are. In order for such 'border' notions to be useful, one would need well-developed techniques for determining who has violated them. This is unproblematic in the literal notion of owned space.

There are some ways of starting to make some progress. The devices that support the internet and the devices that constitute the hardware 'component' of information systems are all owned, or at least there is a more usual way of tracing ownership and boundaries. Some information is stored in, or resident on, parts of these devices. Likewise, information entities travel through devices where they might or might not be stored. However attempts to ground useful notions of cyberspace and ownership of space on owned material devices break down rapidly. The public webpage of a website is resident on certain sectors of a hard drive; the password-protected data is elsewhere; and the operating system is located still elsewhere. But it is in the nature of property and territorial boundaries that they occupy fixed places and, as much as possible, are contiguous and not fragmented.

By Westphalian principles this encroachment on another's territory will generally only occur with the permission of its owners, and hence be inherently compromised. So it does not advance our analysis to locate public and less public amounts of information by the material parts of storage devices on which they reside. We cannot say precisely where these boundaries are, and they shift second-to-second.

PERFIDY AND DECEPTION

In a series of papers, the computer scientist Neil Rowe (2013) has argued that many forms of cyberwarfare involve perfidy, in the sense that it is used in international law; he is almost certainly mistaken to use the term

[81] Even without regard to widely acknowledged criminal acts, extradition is most often a complicated affair, relying on numerous bilateral treaties; with 120 nations, one needs 7,140 separate bilateral treaties.

'perfidy' here. 'Perfidy' is used in a very narrow sense in international law to describe the exploitation of explicitly protected behaviours in the laws and customs of war to further an attack (Tallinn Manual, 2013).[82]

A correct term would be deception. It is much easier to see that when a Ukrainian, for example, hacks into a U.S. Department of Defense computer, he is pretending to be someone else, namely a user with a certain password who has access. But there is no blanket prohibition, either ethical or legal, on deception. Even in the ethical theory most hostile to deception, namely Kantian ethics, there would be exceptions.[83] In games of cards and many board games one tries to deceive other players as to one's cards and intentions. This is one of the most important and essential features of what it is to play the game. If one is aware of a high probability that another party could be bluffing, and there are no rules to the contrary, then deceiving is not 'deceptive' with any moral force suggesting wrongdoing. It is one of the design features of the internet and the software that interacts with it that it provides varying degrees of anonymity that one may choose. The concept of deception only makes sense if there is a 'reasonable expectation' that one can expect an honest representation. Agents using the internet are virtually carrying a sign, saying 'I may not be who I claim I am and what I say may not always be what I believe.' In this situation meaningful deception cannot logically arise.

Even if some internet practices and intrusions are sometimes morally wrong, they are not wrong in the sense required for anything substantive to follow in the high-stakes arena of military ethics. In that arena we are concerned with such consequences as what morally justifies a military counterattack or war, including death and permanent destruction, or what justifies punitive use of force by an international body. No mere internet deception is likely to rise to that level, unless it intentionally or negligently results in death and permanent destruction. In such cases it is covered by a plausible 'effects-based' assessment according to traditional laws and standards of the use of force.

CONCLUSIONS

I have argued that there are key differences between traditional warfare and what we are most likely to see in the near future in cyberwarfare. The differences are ethically significant. Just War Theory arose in contexts where warfare was assumed to always cause widespread death and destruction. The majority of forms of cyberwarfare we are likely to see do not rise to that level. Internet communication has not solidified around customs, practices and a legal environment in which Westphalian principles of territory and sovereignty can be usefully applied.

REFERENCES

Arquilla, John. 1999. Ethics and information warfare. In Khalilzad, Z., White, J., & Marsall, A, eds. *Strategic appraisal: the changing role of information in warfare* (pp. 379-401). Santa Monica, California: Rand Corporation.

Christopher, Paul. 1999. *The Ethics of War and Peace: An Introduction to Legal and Moral Issues,* 2nd ed. Upper Saddle River, NJ: Prentice Hall.

Dipert, R. R. 2006a. Strategies, Rationality, and Game Theory in the Philosophy of War, Paper. Joint Service Academy Conference on Professional Ethics (JSCOPE now ISME). Springfield, VA. http://isme.tamu.edu/JSCOPE06/Dipert06.html. Accessed 3 November 2013.

[82] Paradigms of perfidy include the deceptive use of the 'white flag' of surrender or truce, or of using signs for protected vehicles, such as the Red Cross, to disguise weapons or healthy soldiers.

[83] In what seems to be a perfectly inconsistent position, Rowe seems to endorse the use of honeypots, which seem equally if not more deceptive (Rowe and Goh, 2007).

Dipert, R. R. 2006b. Preventive War and the Epistemological Dimension of the Morality of War. *Journal of Military Ethics*. 5(1): 32-54.

Dipert, R.R. 2010. The Ethics of Cyberwarfare. *Journal of Military Ethics* 9(4): 384-410.

Dipert, R.R. 2013a. Other-than-Internet (OTI) Warfare: Challenges for Ethics, Law, and Policy. *Journal of Military Ethics* 12(1): 34-53.

Dipert, R.R. 2013b. The Essential Features of an Ontology for Cyberwarfare. In Panayotis Yannakogeorgos and Adam Lowther, ed. *Conflict and Cooperation in Cyberspace* Taylor & Francis: New York: 35-48.

Floridi, Luciano. 2008. Information Ethics, its Nature and Scope. *Information Technology and Moral Philosophy* Vol. 40-65. Cambridge: Cambridge University Press.

Goldman, Jan (ed). 2010. *Ethics of Spying: A reader for the intelligence professional.* Lanham, MD: Scarecrow Press.

Libicki, M. 2009. *Cyberdeterrence and Cyberwar.* Santa Monica, CA: RAND Corporation.

Libicki, M. 2012. Panel on Response to Cyberattacks: The Attribution Problem. The McCain Conference, organised by the Stockdale Center for Ethical Leadership. U.S. Naval Academy, Annapolis MD April, 2012. http://www.youtube.com/watch?v=bI7TLqTt0H0 Accessed 3 November 2013

Lucas, George. 2012. 'Permissible Preventive Cyberwar: Restricting Cyber Conflict to Justified Military Targets.'

Lucas, George. 2012. 'Can there be an 'Ethical' Cyber War?' http://www.usna.edu/...%20and%20Cyber%20War%20GR%20Lucas.pdf Accessed 3 November 2013.

Orend, B. 2005. War. In *The Stanford Internet Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/war/ Accessed 13 September 2013.

Owens, W., K. Dam, H. Lin. 2009. *Technology, Law, and Ethics Regarding US Acquisition of Cyberattack Capabilities.* Washington, DC: National Research Council of the National Academies of Science.

Rid, Thomas. 2013. *Cyber War Will Not Take Place* Oxford University Press: NY.

Rowe, Neil. 2006. A Taxonomy of Deception in Cyberspace. International Conference on Information Warfare and Security. Princess Anne, MD.

Rowe, Neil, 2007. The Ethics of cyberwar attacks. In A. Colarik and L. Janczewski, eds. 2007, *Cyber War and Cyber Terrorism*, Hershey, PA: The Idea Group,

Rowe, N. 2009. The Ethics of Cyberweapons. *International Journal of Cyberethics*, 1(1): 20-31.

Rowe, N. 2013. Cyber Perfidy. In Evans, N, ed, 2013, *Routledge Handbook of War and Ethics*, at http://faculty.nps.edu/ncrowe/cyberperfidy.htm accessed 16 Nov 2013.

Rowe, Neil and Han Ho. 2007. Thwarting Cyber-Attack Reconnaissance with Inconsistency and Deception. Proceedings of the 8th IEEE Workshop on Information Assurance, West Point, NY, June 2007. http://faculty.nps.edu/ncrowe/iaw07_reconnaissance.htm accessed 13 Nov 2013

Rowe, J., Crusty, E.J. 2010. Deception in Cyber Attacks. In *Warfare and Cyber Terrorism* Ed. L. Janczewski, A. Colarik. Hershey, PA: Information Science Reference.

Schmitt, M. 1998. Bellum Americanum: the U.S. view of twenty-first century war and its possible implications for the law of armed conflict. *Michigan Journal of International Law, 19*(4), 1051-1090.

Schmitt, M. (Gen. ed.) 2013. *Tallinn Manual on the International Law Applicable to Cyber Warfare* Cambridge University Press: NY.

Strawser, Bradley J. 2013 Issue on Cyberwar and Ethics. *Journal of Military Ethics* 2013:1.

Taddeo, Mariarosario. 2012. An Analysis for a Just Cyber Warfare. 4th International Conference on Cyber Conflict. C. Czosseck, R. Ottis, K. Ziolkowski (eds.). NATO CCD COE Publications: Talinn Estonia. http://www.ccdcoe.org/publications/2012proceedings/3_5_Taddeo_AnAnalysisForAJustCyberWarfare.pdf Accessed 15 Nov 2013.

Walzer, Michael. 2006. *Just and Unjust Wars* 4th ed. New York: Basic Books.

The Cyber-Combatant: a New Status for a New Warrior

Maurizio D'Urso
Italian Defence General Staff, Legal Affairs General Office (SMD – UGAG)
Italy

ugag.ue@smd.difesa.it

Cyber warfare differs from traditional forms of conflict, both in the instruments used – computers – and in the environment in which it is conducted – the virtual world of the internet and other data communication networks.

Cyber warfare, therefore, presents scenarios of conflict which are very different from those typical of traditional war. These differences can make us doubt that cyber warfare, like a traditional war, may be regulated by international humanitarian law (IHL) created at times when the power of information technology was unknown. Virtuality of instruments and actions, however, cannot mean obliteration of international humanitarian law, whose humanitarian protection is needed when cyber warfare produces serious and lasting damages to real environments, in particular to distribution networks controlled by computers.

The purpose of the study is to identify any areas of connection or overlap between the two worlds – the cyber world and the physical – in order to verify whether, even in cyber warfare, the concept of 'direct participation in hostilities' is still operative, with special reference to the laws related to it, and to assesses its consequences with regard to the law of armed conflict. In particular, it addresses the issue of whether in cyber warfare the distinction between lawful combatant and unprivileged combatant is still valid. In the opinion of the author, this distinction does not exist for non–military combatants in the cyber domain and, by working remotely, any civilian who is taking part in cyber warfare, takes direct part in the hostilities as an unlawful cyber combatant.

The paper then examines the concept of continuous combat function applied to cyber combatants, trying to investigate another issue: whether a virtual network, an online forum where members share methods on how to conduct cyber attacks against a common enemy, could be assimilated to a terrorist organisation whose members have a continuous combat function.

The terms of reference of this study are constituted by the *Tallinn Manual on the International Law Applicable to Cyber Warfare*, published by Cambridge University Press in 2013.

The questions of whether the existing rules of international law are capable of countering cyber warfare activities and whether, as the author believes, the Geneva Conventions and Additional Protocols need to be amended in order fit this brand new genre of combatant whose status remains uncertain, remain unsettled.

THE CYBER COMBATANT'S STATUS

The status of lawful combatant is defined in Article 1 of the Regulations concerning the Laws and Customs of War on Land. The Hague, 18 October 1907:

'Article 1. The laws, rights, and duties of war apply not only to armies, but also to militia and volunteer corps fulfilling the following conditions:

1) To be commanded by a person responsible for his subordinates;
2) To have a fixed distinctive emblem recognisable at a distance;
3) To carry arms openly; and
4) To conduct their operations in accordance with the laws and customs of war.

In countries where militia or volunteer corps constitute the army, or form part of it, they are included under the denomination "army."'

This rule is also reproduced in the art. 4 of the Convention (III) relative to the Treatment of Prisoners of War. G Geneva, 12 August 1949.[84]

According to The Hague '*jus in bello*' and Geneva Conventions, any combatant who does not match all four of the conditions set in Article 1 is considered an 'unlawful' combatant, meaning that, in case of capture by the enemy, he is not entitled to claim the rights granted to prisoners of war by Convention III relative to the Treatment of Prisoners of War, Geneva, 12 August 1949. Lawful combatant status applies without any doubt to Armed Forces personnel who take part to cyber operations. Generally, it applies Rule 26 set out in the Tallinn Manual:

'In an international armed conflict, members of the armed forces of a party to the conflict who, in the course of cyber operations, fail to comply with the requirements of combatant status lose their entitlement to combatant immunity and prisoner of war status.'[85]

This could easily happen in special ops against SCADA systems,[86] usually not connected to any public network such as the internet. A cyber attack consisting of introducing malware to a SCADA system by any means is part of a sabotage operation made by an 'insider', supported by spies and undercover special agents.

The issue of status becomes very delicate for non-military cyber combatants because, if they operate with computers only, they lack a basic requirement to be considered as lawful combatants: a computer is not considered a weapon, and thus they do not carry arms openly, as is required to be distinguished from civilian population.

A civilian hacker, in order to be distinguished from a member of the civilian population, should also wear a fixed distinctive emblem recognisable at a distance, but in reality it is difficult to imagine civilian cyber combatants wearing a distinctive emblem when they operate from their computers inside civilian buildings.

The problem of the distinction between cyber combatants and the general population is evident, as cyber combatants do not have the opportunity to distinguish themselves from the civilian population by the mere possession of computers that, unlike those of the civilian population, are used for war purposes. A computer can be a dual-use tool, working as a weapon while remaining indistinguishable from any other PC used by any civilian for his business, study or fun.

On the basis of the above considerations it can be affirmed that, according to The Hague '*jus in bello*' and Geneva Conventions in force, a non-military cyber combatant is an unlawful combatant in almost all cases, with consequent strict limitations on the prerogatives and rights that they may have against the enemy in the event of capture and detention. The current situation is thus paradoxical, as the combatant who uses conventional weapons and lethal force is more protected by IHL than a cyber combatant.

Dinniss believes that, in conflicts to which Additional Protocol I applies, those who conduct cyber attacks as part of the armed forces of a State, whether as part of the regular or irregular forces, will be considered (lawful) combatants.[87]

[84] See Appendix 1.
[85] Schmitt, Tallinn Manual on the International Law applicable to Cyber Warfare, (CUP 2013), p. 96.
[86] SCADA is the acronym for 'Supervisory Control And Data Acquisition**',** a type of industrial control system (ICS). Industrial control systems are computer-controlled systems that monitor and control industrial processes in critical infrastructures, like manufacturing, production, power generation and fabrication.

As a matter of fact, Additional Protocol (AP) I (Article 43) defines combatants more widely:

1) 'The armed forces of a Party to a conflict consist of all organised armed forces, groups and units which are under a command responsible to that Party for the conduct of its subordinates, even if that Party is represented by a government or an authority not recognised by an adverse Party. Such armed forces shall be subject to an internal disciplinary system which, 'inter alia', shall enforce compliance with the rules of international law applicable in armed conflict.
2) Members of the armed forces of a Party to a conflict (other than medical personnel and chaplains covered by Article 33 of the Third Convention) are combatants, that is to say, they have the right to participate directly in hostilities.
3) Whenever a Party to a conflict incorporates a paramilitary or armed law enforcement agency into its armed forces it shall so notify the other Parties to the conflict.'

Thus, according to this opinion, militia and volunteer corps are equally entitled to combatant status and the subsequent right to participate in hostilities. This is true for corps such as the Estonian Defence League's Cyber Unit (EDL CU), a voluntary organisation established in 2008 after the attacks on Estonia a year before, which is intended to protect Estonian cyberspace.[88] The Cyber Unit's mission is to protect Estonia's high-tech way of life, including protecting of information infrastructure and supporting broader objectives of national defence. This organisation is recognised and regulated by Estonian Law, is a part of the organised national defence organisation, and is under the direction of the Estonian Ministry of Defence. The Estonian Cyber Militia is a unique case of formal and lawful integration of paramilitary cyber forces into a National Defence framework, but this formal recognition does not erase the issue of how to distinguish its members from the civilian population.

Actually, Article 43 AP I should be read in conjunction with Article 44,[89] the third paragraph of which states:

'In order to promote the protection of the civilian population from the effects of hostilities, combatants are obliged to distinguish themselves from the civilian population while they are engaged in an attack or in a military operation preparatory to an attack. Recognising, however, that there are situations in armed conflicts where, owing to the nature of the hostilities an armed combatant cannot so distinguish himself, he shall retain his status as a combatant, provided that, in such situations, he carries his arms openly:

(a) during each military engagement, and

(b) during such time as he is visible to the adversary while he is engaged in a military deployment preceding the launching of an attack in which he is to participate.'

A cyber combatant who uses the computer only does not carry arms openly. Article 44(3) suggests that the civilian cyber combatant can carry convention weapons, too: it would be a nonsense. It is incomprehensible that a cyber combatant should raise his level of personal danger by carrying a lethal conventional weapons unless under fire.

The ICRC claims that the definition of Article 43 has become customary international law,[90] but there is no unanimous consensus of doctrine on this point,[91] and no decisive circumstance in which to consider a customary law to exist.

---

[87] Heather Harrison Dinniss, Participants In Conflict – Cyber Warriors, Patriotic Hackers And The Laws Of War, available on the website: Leiden ; Boston : M. Nijhoff, 2013. p. 251-278. In: International humanitarian law and the changing technology of war. Cote 345.25/275

[88] For more information see the site http://www.kaitseliit.ee/en/edl.

[89] See Appendix 2.

DIRECT PARTICIPATION IN CYBER HOSTILITIES BY VIRTUAL COMMUNITIES

Assuming that non-military cyber organisation members cannot be considered lawful combatants, they have to be regarded as unlawful and unprivileged combatants who take part in hostilities. Regarding participation in cyber hostilities, the *Tallinn Manual* states in Rule 29 that 'Civilians are not prohibited from directly participating in cyber operations amounting to hostilities, but forfeit their protection form attacks for such time as they so participate.'[92]

This rule derives from Article 51 AP1[93], 3rd paragraph:

'Civilians shall enjoy the protection afforded by this Section, unless and for such time as they take a direct part in hostilities.'

Therefore, those civilians who participate directly in hostilities lose their general protection against the dangers of military operations and may be attacked for such time as they do so. They may also be prosecuted in domestic or international criminal courts for their actions.

According to the ICRC *Guidance on Direct Participation in Hostilities* the notion of direct participation requires the following three cumulative elements:

1) 'the act must cause harm to the military operations or military capacity of a party to an armed conflict or, alternatively, inflict death, injury, or destruction on persons or objects protected against direct attack (threshold of harm);
2) there must be a direct causal link between the act and the harm (direct causation);
3) the act must cause harm in support of a party to the conflict and to the detriment of another (belligerent nexus).'

Time is another element which plays a decisive role in considering a civilian taking part to hostilities as lawful target.

Membership of an organised armed group means that the person assumes a continuous function for the group involving his or her direct participation in hostilities ('continuous combat function'). The continuous combat function is the mission of irregular combatants, the members of which have with a specific role in the armed organisation which aims to fight the enemy indefinitely until his defeat. This mission is permanent and unlimited as long as hostilities last. The continuous combat function requires lasting integration into an organised armed group. Resembling soldiers of regular armed forces, members of an organised armed group who have a continuous combat function may be attacked at any time. Those who lack a continuous combat function, but who periodically take up arms, must be treated as civilians directly participating in hostilities and may be attacked only while doing so.

Those conducting hostilities face the difficult task of distinguishing cyber combatants who are engaged in a specific hostile act (direct participation in hostilities) from members of organised armed groups (continuous combat function). This difficulty is evident when it comes the issue of direct participation in hostilities of

---

[90] Rule 4, Jean-Marie Henckaerts and Louise Doswald-Beck *Customary International Humanitarian Law* (CUP 2005) Vol 1, 14.

[91] *Inter alios*, see Anthony Rogers 'Combatant Status' in Elizabeth Wilmshurst and Susan Breau (eds) *Perspectives on the ICRC Study of Customary International Humanitarian Law* (CUP, 2007) 101, 110; Yoram Dinstein, *The Conduct of Hostilities Under the Law of International Armed Conflict* (2nd edn, CUP, 2010) 51-55.

[92] Schmitt, *Tallinn Manual on the International Law applicable to Cyber Warfare*, (CUP 2013), p. 104.

[93] See Appendix 3.

members belonging to virtual communities, like forums or chat rooms, or who are members of groups inside popular social networks like Facebook, Google or Twitter, where cyber combatants exchange malware and tips on conducting cyber attacks. The IHL regulations, if interpreted widely, can militate towards considering the virtual network as an organisation that takes part in hostilities, and therefore permit the use of force against members of the network regardless of how effective a contribution they have made to the cyber hostilities.

In a virtual community a continuous combat function that legitimises the targeting of the members anytime can be attributed by distinguishing the member on the basis of their role inside the network. Continuous combat function is ascribed to:

- the administrators of the community, those who have organised the community and who give its members permission to use its services;
- the advisors and the supporting staff, that provide services and technical support; and
- the so-called 'senior members or moderators' - members distinguished from the others by the quality and quantity of their contribution in the community.

All the other members do not have this continuous combat function and could be attacked only if and when they take part in hostilities on the basis of the three parameters of threshold of harm, direct causation and belligerent nexus.

CONCLUSIONS

The existing rules of international law are not capable of countering cyber warfare activities. IHL current regulations do not allow a civilian cyber combatant to be considered a lawful combat, causing a tremendous disparity of treatment in the case of capture and detention when compared with a 'regular' lawful civilian combatant using conventional weapons. This is unfair in equity or in ethics.

Another disparity comes from the issue that, according to the IHL, a cyber combatant can be neutralised not only by cyber attack, but also by the use of kinetic force, even lethal force if necessary. The principle of proportionality usually refers to the use of the minimum force required to accomplish the mission and neutralise the combatant. In case of a fight between computers and conventional weapons, the disproportion between the means used by one side in comparison with the other is clear. So, if in relation to the military advantage achieved it is important to neutralise a fighter, whatever force necessary to achieve that purpose, including that of lethal force, may be used.

It would be worthwhile if the Geneva Conventions and Additional Protocols were to be amended in order to fit this brand new genre of combatant, whose status remains uncertain.

APPENDIX 1 – ARTICLE 4

A. Prisoners of war, in the sense of the present Convention, are persons belonging to one of the following categories, who have fallen into the power of the enemy:

(1) Members of the armed forces of a Party to the conflict as well as members of militias or volunteer corps forming part of such armed forces.

(2) Members of other militias and members of other volunteer corps, including those of organised resistance movements, belonging to a Party to the conflict and operating in or outside their own territory, even if this territory is occupied, provided that such militias or volunteer corps, including such organised resistance movements, fulfil the following conditions:

(a) that of being commanded by a person responsible for his subordinates;
(b) that of having a fixed distinctive sign recognisable at a distance;
(c) that of carrying arms openly;
(d) that of conducting their operations in accordance with the laws and customs of war.

(3) Members of regular armed forces who profess allegiance to a government or an authority not recognised by the Detaining Power.

(4) Persons who accompany the armed forces without actually being members thereof, such as civilian members of military aircraft crews, war correspondents, supply contractors, members of labour units or of services responsible for the welfare of the armed forces, provided that they have received authorisation from the armed forces which they accompany, who shall provide them for that purpose with an identity card similar to the annexed model.

(5) Members of crews, including masters, pilots and apprentices, of the merchant marine and the crews of civil aircraft of the Parties to the conflict, who do not benefit by more favourable treatment under any other provisions of international law.

(6) Inhabitants of a non-occupied territory, who on the approach of the enemy spontaneously take up arms to resist the invading forces, without having had time to form themselves into regular armed units, provided they carry arms openly and respect the laws and customs of war.

B. The following shall likewise be treated as prisoners of war under the present Convention:

(1) Persons belonging, or having belonged, to the armed forces of the occupied country, if the occupying Power considers it necessary by reason of such allegiance to intern them, even though it has originally liberated them while hostilities were going on outside the territory it occupies, in particular where such persons have made an unsuccessful attempt to rejoin the armed forces to which they belong and which are engaged in combat, or where they fail to comply with a summons made to them with a view to internment.

(2) The persons belonging to one of the categories enumerated in the present Article, who have been received by neutral or non-belligerent Powers on their territory and whom these Powers are required to intern under international law, without prejudice to any more favourable treatment which these Powers may choose to give and with the exception of Articles 8, 10, 15, 30, fifth paragraph, 58-67, 92, 126 and, where diplomatic relations exist between the Parties to the conflict and the neutral or non-belligerent Power concerned, those Articles concerning the Protecting Power. Where such diplomatic relations exist, the Parties to a conflict on whom these persons depend shall be allowed to perform towards them the functions of a Protecting Power as provided in the present Convention, without prejudice to the functions which these Parties normally exercise in conformity with diplomatic and consular usage and treaties.

C. This Article shall in no way affect the status of medical personnel and chaplains as provided for in Article 33 of the present Convention.

APPENDIX 2 – Article 44 – Combatants and prisoners of war

1. Any combatant, as defined in Article 4, who falls into the power of an adverse Party shall be a prisoner of war.

2. While all combatants are obliged to comply with the rules of international law applicable in armed conflict, violations of these rules shall not deprive a combatant of his right to be a combatant or, if he falls into the power of an adverse Party, of his right to be a prisoner of war, except as provided in paragraphs 3 and 4.

3. In order to promote the protection of the civilian population from the effects of hostilities, combatants are obliged to distinguish themselves from the civilian population while they are engaged in an attack or in a military operation preparatory to an attack. Recognising, however, that there are situations in armed conflicts where, owing to the nature of the hostilities an armed combatant cannot so distinguish himself, he shall retain his status as a combatant, provided that, in such situations, he carries his arms openly:

(a) during each military engagement, and
(b) during such time as he is visible to the adversary while he is engaged in a military deployment preceding the launching of an attack in which he is to participate.

Acts which comply with the requirements of this paragraph shall not be considered as perfidious within the meaning of Article 37, paragraph 1 (c).

4. A combatant who falls into the power of an adverse Party while failing to meet the requirements set forth in the second sentence of paragraph 3 shall forfeit his right to be a prisoner of war, but he shall, nevertheless, be given protections equivalent in all respects to those accorded to prisoners of war by the Third Convention and by this Protocol. This protection includes protections equivalent to those accorded to prisoners of war by the Third Convention in the case where such a person is tried and punished for any offences he has committed.

5. Any combatant who falls into the power of an adverse Party while not engaged in an attack or in a military operation preparatory to an attack shall not forfeit his rights to be a combatant and a prisoner of war by virtue of his prior activities.

6. This Article is without prejudice to the right of any person to be a prisoner of war pursuant to Article 4 of the Third Convention.

7. This Article is not intended to change the generally accepted practice of States with respect to the wearing of the uniform by combatants assigned to the regular, uniformed armed units of a Party to the conflict.

8. In addition to the categories of persons mentioned in Article 13of the First and Second Conventions, all members of the armed forces of a Party to the conflict, as defined in Article 43of this Protocol, shall be entitled to protection under those Conventions if they are wounded or sick or, in the case of the Second Convention, shipwrecked at sea or in other waters.

APPENDIX 3 – Article 51 – Protection of the civilian population

1. The civilian population and individual civilians shall enjoy general protection against dangers arising from military operations. To give effect to this protection, the following rules, which are additional to other applicable rules of international law, shall be observed in all circumstances.

2. The civilian population as such, as well as individual civilians, shall not be the object of attack. Acts or threats of violence the primary purpose of which is to spread terror among the civilian population are prohibited.

3. Civilians shall enjoy the protection afforded by this Section, unless and for such time as they take a direct part in hostilities.

4. Indiscriminate attacks are prohibited. Indiscriminate attacks are:

(a) those which are not directed at a specific military objective;
(b) those which employ a method or means of combat which cannot be directed at a specific military objective; or
(c) those which employ a method or means of combat the effects of which cannot be limited as required by this Protocol; and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction.

5. Among others, the following types of attacks are to be considered as indiscriminate:

(a) an attack by bombardment by any methods or means which treats as a single military objective a number of clearly separated and distinct military objectives located in a city, town, village or other area containing a similar concentration of civilians or civilian objects; and
(b) an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.

6. Attacks against the civilian population or civilians by way of reprisals are prohibited…'

Cyber Security Ethics at the Boundaries: Systems Maintenance and the Tallinn Manual

Sandra Braman
Department of Communication / Global Studies
University of Wisconsin-Milwaukee
United States

braman@uwm.edu

Fears about the 'vulnerabilities' of the State as a result of digitisation first appeared in policy debate after the Tengelin report to the Swedish government in the late 1970s first enunciated the dangers (Braman, 1991); they underlay the design of what we now call the internet from the start of that process in 1969. Vulnerabilities of individual States become those of the international system itself. Ethical issues raised by cyber security present challenges to the maintenance of the Westphalian system that has provided the foundation and medium for the international system throughout modernity.

Many of the ethical challenges that are unavoidable for those involved in cyber security appear at the boundaries of systems, be they technological, political, or informational, where those boundaries are being contested. Thus disagreements among international experts regarding the application of existing international law to cyber security provide vivid markers of where those ethical challenges to the nature of the international system itself lie. Such points of disagreement are clearly identified in the 2013 *Tallinn Manual on the International Law Applicable to Cyber Warfare*, which provides NATO-sponsored and peer reviewed analysis of a group of experts on existing international law applicable to cyber war. Starting from the unanimous position that new law is not needed because existing law suffices, this group synthesised treaty and customary law into a single document, presenting 95 'black letter' rules for the principles that should drive war practices in cyberspace. The rules deal with the technical details of operationalising the law in areas that are considered established and uncontested. The *Manual* is very clear on the extent of agreement regarding each point and provides arguments behind the differences in position in most of the areas where there are disagreements. It is in two parts, with the bulk of the rules in the section pertaining to cyber warfare. The first section, however, includes 19 rules specific to cyber security, and it is these that are the subject of the analysis presented here.

Ethicists will immediately spot issues in the *Manual*'s discussion of cyber security that they might wish to debate, including Rule 1's support for States that wish to completely shut down their country's networks; Rule 6's approval of State pressure on private citizens to engage in cyber operations against other States or targets abroad ('cyber volunteers'); and issues of political autonomy raised by the Rule 10's willingness to defend some forms of political interference by one State in another electronically as *not* being in violation of international law. Given the changing circumstances and quickly expanding capacities of an 'NBIC' convergence environment, one in which there is a convergence among nanotechnologies, biotechnologies, information technologies and cognitive technologies, the Rule 11 position that psychological operations need not be prohibited, presumably because they were at one point deemed not to rise to the level of coercion, might need to be reconsidered from an ethical perspective.

Cyber security is one among several factors contributing to the destabilisation of the international system that political scientists began commenting on in the 1970s. Legal historians describe these transformations in law-State-society relations as so profound that they should be considered comparable in importance to those that took place with the Westphalian Treaty of 1648 itself. Others include macro-level effects of the use of digital technologies, a variety of other drivers for legal globalisation, environmental decline and population growth. Many of the efforts to engage, respond and adapt to such developments revolve around the emergence of a

global information policy regime (Braman, 2004), with legal developments in the issue area of cyber security dominant among those shaping the overall information policy regime in the twenty-first century.

This paper provides a brief background of the process of regime development, introduces the position that existing international law applies to cyberspace, and identifies the features of cyberspace as distinct from those of material, or kinetic, space that are of importance from a cyber security perspective. Analysis of the points on which the international group of experts that authored the *Tallinn Manual* were *not* able to agree in their discussion of the cyber security-oriented rules (Section I, Rules 1-19) makes evident the types of international system boundaries that are under challenge.

It must be emphasised that the *Tallinn Manual* represents the state of the law as understood by an international group of experts at the time of the completion of the manuscript in 2012. The profound public negotiations and debates over the limits of the State's right to conduct surveillance will bring many of these issues to the forefront and may well, before they are concluded, result in reconsideration, if not shifts, in some of the positions articulated in the *Tallinn Manual*'s rules.

INTERNATIONAL LAW AND CYBER SECURITY

The first task of the international group of experts was to decide whether or not existing law applies to cyber space, a question that policy-makers have had to address across all domains of the law. The experts reached a clear consensus in their answer of yes, but in its operational details the problem became more complicated.

International Law and Cyberspace

The *Tallinn Manual* experts unanimously agreed that existing international law applies to cyber operations, defining as 'cyber-to-cyber' those operations directed against a State's critical infrastructure as well as those against a State's command and control systems (p. 5). Because these matters are well understood under the existing law of armed conflict, cyber operations for the purposes of the *Manual* do not include 'kinetic-to-cyber' operations (such as an aerial attack against a cyber control centre), where the word 'kinetic' refers to attacks using material means, or to traditional types of electronic warfare attacks such as jamming radio signals.

The thinking behind this position is presented in discussion of Rule 9 on the question of how to determine whether or not unacceptable use of force had been or was being used. Authors of the *Manual* understand the UN Charter to prohibit 'any use of force, regardless of the weapons employed' (p. 42):

'Therefore, the mere fact that a computer (rather than a more traditional weapon, weapon system, or platform) is used during an operation has no bearing on whether that operation amounts to a "use of force." Similarly, it has no bearing on whether a State may use force in self-defence.'

NATO began its own engagement with cyber operations in the late 1990s, when activities in the Balkans were disrupted by hackers (Woudsma, 2013). Its commitment to cyberspace was made explicit in 2010 when all NATO bodies were brought under centralised cyber protection, NATO planners began to enhance and coordinating national capabilities, and efforts to better integrate NATO cyber warning and response systems with those of member States were launched. Although the UN Security Council had never, at the time of the writing of the *Tallinn Manual* in 2012, officially identified any cyber operation as a threat to peace, breach of peace, or act of aggression, the international group of experts unanimously agreed that it 'incontrovertibly' has the right to do so. US policy has had enough impact on NATO and UN positions that the lead author on the *Tallinn Manual* also published a journal article providing, in essence, a concordance that compares the US position as it stood then with the positions taken in the *Manual* (Schmitt, 2012).

Cyberspace Features Requiring Adaptation of the Law

When the US elaborated its International Strategy for Cyberspace in 2011, it noted that the 'unique attributes of networked technology require additional work to clarify how these norms apply and what additional understandings might be necessary to supplement them' (p. 3). Three features of cyberspace requiring such attention were explored in the *Tallinn Manual*: those that raised ontological issues, those that involved challenges to sovereignty, and those that affected the nature of governance.

Ontological issues. Some of the specific characteristics that make the creation, processing, flow and use of intangible digital information problematic for policy-makers can be described as ontological in the sense that they have to do with the nature of the thing; in the Aristotelian sense, involving the essence of being at the most abstract level (OED). Many of these are familiar from other discussions of the nature of the information society, or the network economy, that have been taking place for decades.

The time and space compression so key to globalisation complicates cybersecurity efforts aimed at fairly and justly protecting against and responding to harmful cyber operations. Traditional reliance upon the rhythms of diplomatic communication under Westphalian rules, designed to slow down progress towards war in the hope of preventing it, is not available in today's network environment. This becomes evident in the *Tallinn Manual*'s discussion of Rule 13 on self-defence: if a victim country cannot act in self-defence, perhaps endangering the system as a whole, another country may need to act in a situation in which instantaneity precludes the kinds of intermediary steps which are available in the kinetic diplomatic environment such as a request to cease and time to allow cessation to take place (pp. 60-61). The problem of instantaneity runs throughout the discussion of rules 13 to 15 in efforts to ascertain a meaningful distinction between preparatory actions and the initial phases of an attack (p. 65).

Instantaneity is a characteristic of digital agency that also contributes to the dynamism of the intelligent network environment. It is noted, for example, that cloud and distributed computing may span national boundaries in ways that constantly change with the replication and dynamic relocation of data (Rule 2, p. 19). The ability to reroute easily and quickly was a primary design criterion for the internet, resulting in route dynamism so facile and flexible that asking States to take responsibility for cyber operations that transit through their territory is effectively meaningless (Rule 5, p. 28). On the other hand, it is the same dynamism that makes it possible to establish a preference in international law for countermeasures that have temporary or reversible effects (Rule 9, p. 38).

Constant innovation is another factor contributing to the peculiar qualities of the cybersecurity context. The authors of the *Tallinn Manual* note the 'often unprecedented character' (Rule 5, p. 26) of the technologies, processes and effects of concern. They point out that, since the subjects of the discussion are themselves continuously changing, the problem of establishing agreed-upon definitions, criteria for evaluation and thresholds for the application of criteria are even more difficult than would otherwise be the case (Rule 9, p. 42).

Cyberspace and sovereignty. In order to extend geopolitically-based law to the network-based conditions of cyberspace, the *Tallinn Manual* examines the relationship between cyberspace and sovereignty from several directions. Starting from the foundational position that cyberspace itself is sovereign free, the *Manual* asserts that a government can claim sovereignty over cyber infrastructure located within its territory and activities associated with that infrastructure (Rule 1, p. 15). While this would seem to be simply analogous to common sense notions of geopolitical sovereignty in kinetic space, network structure does not replicate geopolitical structure. In reality, operationalisation of this principle in cyberspace *de facto* gives the US sovereignty over the entire internet in its current form.

Today's debates over the future of internet governance are taking place concurrent with efforts by some countries to experiment with delinkage from the internet, replacement of the internet (defined, as it is, by a set

of technical protocols and not by material passageways) with an alternative, and other defragmentation pressures. Under most, if not all, of these scenarios US network centrality is under challenge. Full discussion of the implications of these developments for cybersecurity is beyond the scope of this paper, but will affect future negotiations.

The possibilities of multiple and collective jurisdiction have been noted. The transborder nature of the network is also critical from a jurisdictional perspective; the experts also noted, though, that they were presupposing that multi-jurisdiction does not undermine the potential for a given government to exercise jurisdiction in any specific circumstance (Rule 2, p. 19). At the same time, the nature of the network also means that in a sense governments are jurisdictionally *non*-engaged; as is discussed in the exploration of Rule 5, on control of cyber infrastructure, the passage of data through networks located in a State does not 'presuppose' any involvement by that State (p. 28).

Cyberspace and governance. The use of induced hardware failures as a means of enforcing copyright law through 'digital rights management' (DRM) techniques made vividly clear even to consumers that among the effects of the use of digital technologies on governance has been the availability of new types of policy tools. The experts who produced the *Tallinn Manual* agreed that in the area of cyber security 'the availability of countermeasures by cyber means expands the options available' (Rule 9, p. 39).

Once past agreement on the binary question of whether, under international law, a State can exercise jurisdiction over cyber infrastructure that is on its territory, territory that it owns or controls, and that which it has a responsibility to protect, there were several areas of disagreement among the experts on questions raised by specific contexts. In Rule 13's discussion of self-defence against armed attack, experts did not agree on two points. The first involved including motivation among the criteria to be taken into account when determining whether or not an operation should be considered an armed attack for the purpose of justifying the use of force in self-defence. Some experts argued that motivations do matter, taking the position that attacks motivated by purely private interests would not trigger the right of self-defence. Others asserted that motives should be irrelevant (p. 59).

There was also disagreement regarding location. Some argued that attacks on objects outside a State's territory should be considered attacks against the State for the purposes of triggering the right of self-defence if those attacks are against non-commercial government facilities and personnel. Others believed that additional factors should be taken into account, including the extent of the damage caused, whether the property involved is public or private, the status of the targeted individuals, and motivation with particular interest in whether or not the target of an attack was chosen because of nationality. The focal case for this discussion involved a cyber operation by one State with the purpose of killing the chief executive officer (CEO) of a second State's State-owned corporation while that CEO was abroad; the experts were *not* able to agree on whether or not such an operation would be considered an attack by one State upon another, and concluded that in this area it would probably be practice that ultimately leads the law.

CYBER SECURITY AND EPISTEMOLOGY

The 'unique characteristics' of cyber space from a security perspective also include epistemological issues. They are worth distinguishing and highlighting because they are so fundamental to the intertwinings of knowledge, power and governance that are the dominant features of modernity and of the Westphalian system. Challenges in these areas are not only theoretical, they are challenges to the ability to make and keep peace and ensure security in the cyber environment.

Three types of epistemological issues run throughout the *Tallinn Manual*'s rules for and discussion of cybersecurity. The impact of digital technologies on facticity plays out in the cyber security realm in multiple ways. As with any legal regime, evidentiary issues are important. The role of knowledge as a criterion to be taken into account when evaluating the seriousness and legitimacy of particular cyber operations plays a far

larger role than it has historically in warfare, although military history is among those that is characterised by ever-increasing information intensivity (de Landa, 1991; Keegan, 1993; van Creveld, 1991).

Facticity

Facticity is a social formation that involves a cultural orientation around the fact, whether towards or away from it. In cultures characterised by facticity, there are specific social functions for particular narrative forms or genres. There are also detailed and verifiable procedures by which facts are developed and evaluated, and institutional certification of facts, fact-producing processes and fact producers. The intertwining of facticity, institutions and socio-economic and political power are characteristic of modernity, and of the Westphalian system. The *Tallinn Manual* notes a number of challenges to facticity presented by the characteristics of cyberspace that are, in turn, problematic from a security perspective.

Falsification, difficulty in attribution and geolocation, deception and complex correlation tasks all confound the effort to discern the source of a threat against peace or use of force, and to identify the responsible party. These are issues that derive from the fundamental technical characteristics of cyberspace, the conditions of which exacerbate the types of lengthening causal chains and mismatch between processes of concern and our perceptual mechanisms that Ulrich Beck (1992) pointed out has made it so difficult to assign liability and accountability in the contemporary environment.

There is a great deal of concern over the possibility that the identity of an entity exercising jurisdiction might be 'spoofed,' or falsified (Rule 2, p. 19; Rule 7, p. 35), making it difficult or impossible to know which State or non-State entity is responsible for a cyber operation that might be deemed a use of force or armed attack. It is often difficult to geolocate the source of a cyber operation for the same reason. In addition to deception, facticity issues can be generated by the difficulty of the attribution of responsibility for a given cyber operation, and the need to correlate separate sets of events in order to understand them as part of a coordinated and distributed attack on one or more targets.

During discussion of Rule 5, the experts pointed out that this lack of knowledge is far from trivial; if it is not possible to determine the source of an imminent or ongoing attack, it may be necessary to destroy the entire system in order to succeed defensively. That would include causing harm to oneself in what the authors of the *Tallinn Manual* called 'self-denial' of service (p. 27). During times of conflict, the experts thus suggest that a control test be used to determine whether or not a State is the party responsible for actions against which defence is desired. Using this test, an entity 'may be deemed to have enemy character if it is under the actual control of a person or of persons residing, or carrying on business, in enemy 'territory' (Rule 3, p. 23).

Evidentiary Issues

The international experts involved in the *Tallinn Manual* offered a number of specific guidelines for how to interpret information of various types. The discussion of Rule 8, for example, specifies that government or non-government cyber infrastructure located in a State through which an attack launched in another State is routed 'cannot be presumed to be associated with the cyber operation' (p. 36).

The generation of false news or findings can be considered falsification of evidence. False news was among the several types of coercive political interference deemed prohibited interventions if carried out by one State within another (Rule 10, p. 45). Another form of political interference prohibited by international law involving cyber security include forms of falsification such as the manipulation of elections.

Witting Requirements

The question of whether or not action must be witting -- knowingly[94] undertaken -- runs throughout the *Tallinn Manual* discussion of cyber security because what might be called 'witting requirements' have been inherited from various sources of pertinent international law, but their implementation in the cyber context is quite difficult. The international group of experts noted, in their discussion of Rule 5 on control of cyber infrastructure, that the requirement that a State may not knowingly allow its territory to be used for activities that cause damage to another state are "complicated by the nature of harmful cyber-acts ... and their often-unprecedented character' (p. 27). 'Actual knowledge' of particular attacks is defined as having being achieved either when intelligence agencies detect a cyber attack or when credible information is received that a cyber attack is underway. In the electronic environment, the experts acknowledge, this may be impossible to achieve because of the difficulties in attribution, correlation of multiple distributed events, and the ease of deception (p. 28). As Massumi (2007) has pointed out, doctrines that allow pre-emptive activity in a counter-terrorism environment vastly multiply the range and scales of possible unknowables and uncertainties, making the amount of knowledge an entity would need for full protection essentially boundless.

The problematic importance of being witting came up again in discussion of Rule 14, on necessity and proportionality requirements. When considering whether the conditions of necessity justify the use of force in a given situation, it could be possible that an attack had stopped but the victim State did not know that and thus kept using force in self-defence. Experts did agree that doing so would be reasonable and thus justified. Before that point is reached, it is also possible that there may have been an armed attack or one that has been underway for a long time but not known, either because the cause of the damage or injury was not identified or because the initiator of the attack was not known until long afterwards (p. 66).

Discussions about what to do if a State does *not* have adequate knowledge, and thus cannot be said to be acting wittingly, are often those where there is no consensus among the international experts who produced the *Tallinn Manual*. These are at the boundaries of the international system where ethical issues historically have always become intertwined with system maintenance (Wuthnow, 1989).

CHALLENGES TO SYSTEM BOUNDARIES

There were many points on which the group of experts could not reach a consensus deal with the fundaments of the international system. Here we will look at the implications of these areas of disagreement for the sustenance of the system itself; what might be referred to as international rule of law; which agents and actions are of concern from the perspective of system vulnerability; and differential considerations as they arise in response to differences in the form or phase of power under evaluation.

The International System (Rule of Law)

The experts unanimously agreed on the general point that existing international law applies to cyber issues. The same consensus, however, was not achieved in specific areas that are fundamental to evaluations of threats to peace and uses of force in the cyber environment. These begin with a basic question: does the consensus that the law of armed conflict applies to cyberspace mean that it fully applies, that anything not explicitly forbidden is generally permitted, or something in between? Concluding that that question could *not* be answered consensually, the experts also felt it important to note that '... the fact that States lack definitive guidance on the subject does not relieve them of their obligation to comply with applicable international law in their cyber operations' (p. 3). There are other issues that probe the universality of international law when

[94] The definition in the *Oxford English Dictionary* is: "The fact of knowing or being aware of something ;

knowledge, cognizance" (accessed November 7, 2013).

applied to cyberspace. In discussion of Rule 9, for example, it is asked which norms and obligations universally apply when countermeasures are involved (p. 39).

An important element of the rule of law from the perspective of a government as sovereign is the right to call upon sovereign immunity; immunity from the jurisdiction of the courts of another State. The experts addressed this question in their discussion of Rule 4, but did *not* reach a consensus on whether such immunity should exist in the cyber environment (p. 25). The international group of experts was *not* able to agree on whether objects owned or used by the State are covered by sovereign immunity, a problem that arose in the course of discussing Rule 4. Those who argued in favour of extending sovereign immunity to objects pointed to protections for diplomatic communications and archives in times of conflict and reference to special protections, via bilateral or multilateral agreements, for particular locations and objects.

The problem of whether a defensive cyber operation could be launched from or use assets located in a State to which the act cannot be attributed when there is neither consent nor a UN mandate arose in the course of discussing Rule 13 on self-defence. On this point, most of the experts took the position that this must be acceptable when the territorial State is unable or unwilling to stop the cyber operations, but only if the victim State asks for assistance and the State involved in the offensive activity is offered an opportunity to address the situation first (pp. 60-61). The minority would add the additional criterion of necessity before considering such operations acceptable.

Agents of Concern

From a complex adaptive systems perspective, the State is one among many systems at multiple levels that are interpenetrated and interacting in multiple ways. The same would have been the case at the Westphalian moment, which saw the secular system of States replacing church-based hierarchies of power as the locus of identity within the international system. Almost 500 years ago the decision was to treat States and only States as the systems of concern from the perspective of identifying agents capable of actions subject to international law. The international acceptance of Al Qaeda as having launched an armed attack triggering self-defence after 9/11 was a significant turning point on this question, for the first time extending the right of self-defence to a non-State entity (13), a position on which this international group of experts did *not* reach a consensus. The question of non-State actors first comes up in discussion of Rule 1, where it was noted that an 'embryonic' view held by a minority did view cyber operations by non-State actors as matters that could violate a State's sovereignty (p. 17).

Once international law is extended to non-State actors, the problem of how to determine which should be considered systems of concern because they could threaten or harm the State comes to the fore. The *Tallinn Manual* notes that there is a great deal of uncertainty within the international legal community regarding just how organised a group must be, if at all, in order to be considered capable of mounting an armed attack from a legal perspective (Rule 13, p. 58). Experts did *not* agree on whether a single individual acting on his or her own, in contrast to those 'cyber volunteers' acting under the direction of a State, who launches a cyber operation that qualifies as an armed attack could trigger the right of self-defence (Rule 13, pp. 58-59). Discussion of Rule 12 had already noted that:

'[C]yber capability is not as dependent on a State's size, population, or economic and military capacity as is the capacity to use conventional force. This means it may be more difficult for a State to evaluate the capacity of another State to make good on its threat to use force by cyber means. Therefore, this issue plays a diminished role in evaluating cyber threats.' (p. 53).

On the separate question of whether a threat should be treated as such from a cyber security perspective if the international community believes that the State making the threat may have the capacity but not the intention to do so, perhaps because the most important audience for the rhetoric is internal, experts also did *not* agree (p. 53).

Agency of Concern

Which actions should be considered of concern also received attention. There is a running debate throughout the *Tallinn Manual* over whether evaluations regarding the seriousness of some action, and thus acceptable responses to it, should be based on the intentions behind and nature of the action, or on its effects. The international experts did *not* agree on this question when it came up in discussion of Rule 13 on self-defence. Those experts who focus on effects still disagreed, during their discussion of Rule 13 on self-defence, regarding whether it is the nature of the damage or its extent that should determine whether or not a given action should be considered an armed attack triggering the right of self-defence. The example used during this conversation also raised the question of which types of effects should be taken into account, for it involved what was described as the 'classic' scenario of an incident involving an attack on a stock exchange that causes it to crash; some experts viewed this damage as 'merely' economic and not a justification for self-defence, while others pointed to the catastrophic consequences that could ensue and would treat it as an armed attack (p. 59).

The question of intentionality, pertinent at the level of the international system and particularly resonant with issues raised as they affect individuals under counter-terrorism laws such as the USA PATRIOT Act, came up during discussion of Rule 13 on self-defence. There was *no* consensus on this point in the abstract. The case explored in the discussion involved unintentional damage by one country to another's cyber infrastructure as a result of cyber espionage that is not prohibited by international law. Some experts took the position that while it could not be considered an armed attack without intent, the principle of necessity could justify taking countermeasures. The majority, however, treated intent as irrelevant and argued that only scale and effects should be taken into account with necessity and proportionality as key criteria. A State affected by 'bleed-over' effects, for example, even if not the intended target of an attack, does have the right of self-defence.

Phase of Power

Although a full exploration of the *Tallinn Manual* through the lens of how its rules deal with differences in forms and phases of power is beyond the scope of this piece,[95] there are disagreements among experts about the application of international law when it comes to the treatment of power as exercised in the past and in the future worth noting.

Actual vs. sunk. We can think of 'sunk' power as that which was exercised in the past and continues to exert experiential force. Power of any kind in its sunk phase will often be intertwined with symbolic power in its actual phase, but can be distinguished from that in the kinds of circumstances that are the subject of the *Tallinn Manual*. The question of how to think about the exercise of sunk power arises in discussion of Rule 9, when the question arose of whether the continuation of countermeasures is justified once the act to which they are responses has ceased. The majority of experts said it was not justified, but noted that that position runs against State practice. There have been times, the *Manual* comments, when it has appeared that States have been acting punitively, sometimes in addition to justified countermeasures but sometimes solely so. Perceptions that State action is punitive rather than a justifiable countermeasure increase when that action begins after the triggering action has stopped.

Actual vs. potential. The distinction between power in its potential rather than actual phase arises several times in the *Tallinn Manual*'s analysis. The discussion of Rule 5, on control of cyber infrastructure, dealt with the question of whether or not States should be held responsible for preventing 'merely prospective' cyber operations as distinguished from those already being planned or under way, operations that could happen, but are not yet happening. Some experts interpret the law to mean that States have a positive responsibility to try

[95] For a discussion of the rising importance of informational power relative to power in its instrumental, structural and consensual (symbolic) forms, and of the distinctions between power in its actual, potential and virtual phases, see Braman (2006).

to prevent prospective attacks, but there is *no* consensus on this point. Others point out that it would be impossible for any State to fulfil such a responsibility because of the nature of cyberspace itself (p. 27).

Particular factors that might affect whether or not a potential use of force becomes actual received specific attention. Rule 12's focus on defining a threat of force raised the question of whether a threat should be treated as such from a cyber security perspective if the State involved does not, in the estimation of the international community, have the capability to actually follow through (p. 53).

CONCLUSIONS

The accomplishment of the *Tallinn Manual* in providing a comprehensive and rigorous analysis of international law as it pertains to cyber security issues is enormous. Clearly there are a number of critical elements that are as yet unresolved; many of those result from characteristics of cyberspace that are so qualitatively different from those of kinetic space that a great deal of thinking and experimentation will need to take place to fully articulate and implement the law under the range of possible conditions.

At the time of writing, many of those are the subject of political battles, structural tensions, and public debate around the world. Among the matters at stake is the maintenance and survival of the Westphalian international system. Points of disagreements among experts on international law make visible challenges to that system's fundamental approach to international law and the rule of law, to the actors and actions of concern, and to phase and forms of power.

ACKNOWLEDGMENT

Thanks are due to the Center for International Education at the University of Wisconsin-Milwaukee for the 2012-2013 Global Studies Research Fellowship that provided partial support for this work.

REFERENCES

Beck, Ulrich. (1992). *The risk society: Towards a new modernity*. London: Sage Publications.

Braman, Sandra. (2006). *Change of state: Information, policy, and power*. Cambridge, MA: MIT Press.

Braman, Sandra. (2004). The processes of emergence (pp. 1-11) and The emergent global information policy regime (pp. 12-37). In Sandra Braman (Ed.), *The Emergent global Iinformation Policy Regime*. Houndsmills, UK: Palgrave Macmillan.

Braman, Sandra. (1991). Vulnerabilities of the state and the New World Information and Communication Order, *Media Development*, *38*(3), 6-8.

DeLanda, Manuel. (1991). *War in the age of intelligent machines*. New York: Zone.

Keegan, John. (1993). *A history of warfare*. London: Hutchison.

Massumi, Brian. (2007). Potential politics and the primacy of preemption, *Theory & Event*, *10*(2), theory_and_event/v010/10.2massumi.html.

Schmitt, Michael N. (Ed.). (2013). *Tallinn Manual on the international law applicable to cyber warfare*. Cambridge: Cambridge University Press.

Schmitt, Michael N. (2012). International law in cyberspace: The Koh speech and the Tallinn Manual juxtaposed, *Harvard International Law Journal*, *54*, 13-37.

van Creveld, Martin. (1991). *Technology and war: From 2000 BC to the present,* rev. ed. New York: Free Press.

Woudsma, Peter. (2013). Cyber defence: A major topic in NATO's transformation, Transformer 2013-01, www.act.nato.int, accessed November 5, 2013.

Wuthnow, Robert. (1989). *Meaning and moral order: Explorations in cultural analysis*. Berkeley, CA: University of California Press.

Violence, Just Cyber War and Information

Massimo Durante
Department of Law, University of Turin
Italy

massimo.durante@unito.it

War is a very old concept which has affected not only politics, law and history but also philosophy. War was meant to be the origin of all things (Heraclitus). The idea of becoming was seen as a conflict that marks the passage between being and not-being (Plato). The theoretical roots of war can be traced back to the idea of action, conceived as the capacity of an entity to affect another entity by destroying or modifying it. Thus, a war action has always had two dimensions: an act of destruction, which damages, deteriorates, deletes or suppresses an entity, and an act of exploitation, which alters, modifies or distorts an entity, in order to obtain something more or something different from what this entity is normally expected to be for.

The traditional view of war has changed from an empirical and a theoretical viewpoint. We progressively move from hard to soft powers (Nye, 2004): in this scenario, the second dimension of a war action becomes more and more important. War is no longer based only or even mostly on physical kinetic armed attacks, but also on political, economic, ideological and informational strategies intended to exploit someone else's informational resources. This does not amount to saying that war ceases to be destructive; rather it means that a deeper comprehension of what war is in the cyber-age requires us to take into full account these two dimensions.

'The impact of a cyber attack depends on what is targeted and more importantly what relies on that target' (Gervais, 2011, 5). Cyber attacks target computers: our current information societies are everywhere increasingly based and dependent on computers. That is why a cyber attack is meant to be able to affect, either directly or indirectly, any trait of our societies, according to the unique (military) or dual-use (civilian and military) nature of targeted objects (Gervais, 2011, 36; Richardson, 2011, 27). That is also why it is so important to understand the conceptual core of a cyber attack, in order to better grasp its critical impact on our information societies.

THE TWOFOLD INFORMATIONAL DIMENSION OF A CYBER ATTACK

The question arises as to what a kinetic cyber attack is: how force is to be interpreted in the cyber age. We need a unified approach to our understanding of a cyber attack, which may encompass the two dimensions of a cyber war: destruction and exploitation. Destruction is a traditional concept which belongs to the common representation of war, whereas exploitation is an area of rising importance fostered by the ongoing development of cyber war. A comprehensive theoretical framework is offered to us by the informational approach provided for by Luciano Floridi's philosophy (2011) and ethics of information (2013). This framework may enable us to deal with the conceptual core of the idea of cyber war, which is a war on information through information.

According to Floridi's philosophy of information, any entity is an informational object: 'any entity is a consistent packet of information, that is an item that contains no contradiction in itself and can be named or denoted in an information process' (Floridi, 1999, 43). This is a static representation focused on the epistemological dimension of an informational object. On the basis of such representation, every epistemic subject can be an informational object at a certain level of abstraction. Some information entities are also agents, that is to say entities 'capable of producing information phenomena that can affect the infosphere' (Floridi, 1999, 44). This means that an information agent is not only a consistent packet of information but also a source of information. This is morally relevant given that, according to Floridi's model of information ethics, an

information agent can also be a moral agent in one of three ways, since she 'can avail herself of some information (information as a *resource)* to generate some other information (information as a *product)* and, in so doing, affect her information environment (information as a *target*)' (Floridi, 2010, 102).

Some current definitions of cyber attack do not encompass some important aspects as the crucial dimension of exploitation. Three examples are (Gervais, 2011, 8; Schimdt and Cohen, 2013, 103):

'The damaging, deletion, deterioration, alteration or suppression of computer data without right', and 'the serious hindering without right of the functioning of a computer system by inputting, transmitting, damaging, deleting, deteriorating, altering or suppressing computer data' (Council of Europe, Convention on Cybercrime, opened for signature Nov. 23, 2001, 41 I.L.M. 282, articles 5-6).

'The premeditated use of disruptive activities, or the threat thereof, against computers and/or networks, with the intention to cause harm or to further social, ideological, religious, political or similar objectives. Or to intimidate any person in furtherance of such objectives (U.S. Army Cyber Operations and Cyber Terrorism Handbook, 2006, VII-2).

'Action by a nation-state to penetrate another nation's computers or networks for the purposes of causing damage or disruption' (Richard Clarke, former U.S. counterterrorism chief, cited in Schimdt and Cohen, 2013, 103).

According to Gervais, even if cyber espionage or cyber exploitation is of greater importance as a major threat to commerce and national security, it 'fails to rise to the level of warfare', and 'does not violate the international laws of war' (Gervais, 2011, 9). We do not want to stretch the concept of cyber attack, but we believe that consideration of cyber warfare should take into account its whole informational dimension.

Other scholars speak of cyber intrusion when a broader concept is invoked (Kesan & Hayes, 2012, 439-440), and refer to cyber attacks and cyber exploitations to denote the two specific subtypes of cyber intrusions. A cyber attack is characterised by the fact that it seizes an entity as a 'source of information'. This may happen in two different ways (Floridi, 2010):

- A cyber attack deprives an entity of its capacity to be *a* source of information, because it damages, deteriorates, deletes or suppresses it. In these circumstances, a cyber attack is a disruptive activity, which patently rises to the level of warfare; or
- A cyber attack deprives an entity of its capacity to be *that* source of information it would have been if not under attack, because it alters, modifies or distorts the way this entity is a source of information, or the information displayed by this entity. It turns information into misinformation or disinformation.

In these circumstances, which may encompass espionage or exploitation, a cyber attack is not a disruptive activity as such, but it can lead to disruptive effects which may rise to the level of warfare.

It is difficult to assess when a cyber attack amounts to a prohibited use of force under Article 2(4) of the United Nations Charter, for the very simple reason that 'force' is still interpreted as being traditionally associated with the military instrument. There are at least four approaches (Gervais, 2011, 11-12) to analysing force in cyber warfare:

1. the 'method of delivery', which takes into account the specific method of delivering an attack and prohibits cyber attacks based on how they are executed;
2. the 'strict liability' model, which takes into account the specific target of an attack and prohibits cyber attacks directed against 'critical infrastructure';

3. the ‘direct result’ model, which takes into account the direct result of an attack and prohibits cyber attacks that attempt to cause direct physical destruction, injury or death; and

4. the ‘consequence-based’ model, which takes into account the reasonably foreseeable consequences of an attack and prohibits cyber attacks when their effect resembles that of a traditional attack.

None of these approaches can account for all the aspects of cyber attacks, but each points out some issues which must be considered in order to have a full understanding of what is the recourse to force through cyber attacks. The following points are critical (Gervais, 2011, 11-12):

- Cyber weapons might be outdated by the time their prohibition is codified.

- The strict reference to critical infrastructure may collapse the distinction between armed violence, coercion and mere interference. A strict liability model would justify anticipatory self-defence in almost any case of a threat of harm aimed at a critical infrastructure.

- The direct effects of cyber attacks can result in non-physical damage.

- The indirect effects of cyber attacks may well result in physical damage that, therefore, should be taken into consideration, but it is often difficult to trace back the indirect effects to a specific cyber attack.

- It is hard to state whether the reasonably foreseeable consequences of a cyber attack resemble those of a conventional attack on the basis of six criteria (severity, immediacy, directness, invasiveness, measurability, and presumptive legitimacy), which are not always applied and which account for the dynamics of cyber attacks. Which criterion prevails over the other? In such cases, there is ‘little guidance as to the weight of each of the six factors’ (Gervais, 2011, 14) and the model leaves the way open for unconventional measures of coercion, like economic, diplomatic or ideological coercion (*ibid*, 15).

These points can be summed up by saying that the notion of cyber war is no longer based on conventional military instruments, physical damage, direct effect and strict armed violence. In this framework, we have to consider more closely how the idea of cyber war differs from that of conventional war. This requires us to investigate the idea of violence, which is of central importance across modern ages not only with regards to the concept of war but, first and foremost, to the idea of politics itself.

CYBER WAR AND TRADITIONAL WAR

Cyber war changes the idea of war. Traditional war is mostly conducted by human beings through the use of physical force, namely, of kinetic violence. Traditional war is characterised both from an internal and an external viewpoint:

From an internal (or vertical) viewpoint, traditional war is the sovereign State’s claim addressed to its members, through which it calls them to be ready to sacrifice their life:

‘When the prince says to him: ‘It is expedient for the State that you should die,’ he ought to die, because it is only on that condition that he has been living in security up to the present, and because his life is no longer a mere bounty of nature, but a gift made conditionally by the State’. (Rousseau, ed. 1997, II.IV)

This places the essence of war in the conceptual framework of death. According to Hobbes, the modern sovereign State is constructed on and politically legitimated by its capacity to delay the individual’s death: symmetrically, during wartime, such a delay is put off and the risk of death again becomes imminent.

From an external (or horizontal) viewpoint, war follows from a sovereign State's war declaration addressed to another State:

'War then is a relation, not between man and man, but between State and State, and individuals are enemies only accidentally, not as men, nor even as citizens, but as soldiers; not as members of their country, but as its defenders. Finally, each State can have for enemies only other States, and not men; for between things disparate in nature there can be no real relation'. (Rousseau, ed. 1997, I.IV)

Declaration of war is essential to the traditional concept of war. Hobbes remarks that:

'For war, it consists not in battle only, or the act of fighting, but in a tract of time, wherein the will to contend by battle is sufficiently known; and therefore the notion of time is to be considered in the nature of war'. (Hobbes, ed. 2008, XIV)

According to Hobbes, war combines information and time: war is a tract of time, not in any chronological meaning, but rather as a form of anticipation based on the pending menace of battle and death, which therefore lies beneath both the internal and the external dimension of traditional war.

In a globalised world, sovereign States' right to co-existence involves also a right and a duty to cooperate, based on their involuntary interdependence (Picone, 1995, 519). Thus, the traditional concept of war changes, not only as a result of the evolution of cyber war, but also as a result of the deep changes in the international context in which this concept has been elaborated. Globalisation, exacerbated by the digitisation that has turned sovereign States into computer-dependent societies, has gradually raised issues with the myth of sovereign States' self-sufficiency; the idea of an international legal order based on the equilibrium between isolated and autonomous sovereign States; and the States' natural rights to act unilaterally, *uti universi* (Viola, 2005, 41-42).

The involuntary interdependence between sovereign States marks the shift from the traditional idea of government, based on the concept of will, to the contemporary idea of governance, based on the complexity of reality that transcends the idea of will. Such an involuntary interdependence generates also a tension that affects both national and international security.

National security

At the level of national security, there is a strong tension between the principle of responsibility, according to which the authority entrusted with the responsibility to assure security also has the competence to decide what this security requires, and the duty to international cooperation, under which it is no longer feasible to guarantee national security without guaranteeing international security.

International security requires States to cooperate by setting up agreements and by providing these agreements with stability. We should not forget that, according to Hobbes, security is guaranteed by the stability of pacts. The societal immunisation from the imagined original violence is not meant to assure individual security, but rather the stability of pacts; it is meant to assure individual security through the stability of pacts:

'The cause of fear, which makes such a covenant invalid, must be always something arising after the covenant made, as some new fact or other sign of the will not to perform, else it cannot make the covenant void. For that which could not hinder a man from promising ought not to be admitted as a hindrance of performing'. (Hobbes, ed. 2008, XIV)

The stability of pacts depends on the available information already included in the hypothetical original violence, which enables us to anticipate 'something arising after the covenant [is] made, as some new fact or other sign of the will not to perform'. At the international level, there is, nonetheless, the problem of how to

make agreements and have them respected without the use of force, since there is no international *Leviathan* and 'a covenant not to defend myself from force, by force, is always void' (Hobbes, ed. 2008, XIV).

International security

There is a strong tension between sovereign States and the individual's claim to the protection of human rights in the international arena. There is a spreading tendency, furthered by digitisation, of individuals or groups of them to perceive themselves as international political subjects, both as patients, whose fundamental prerogatives are to be protected anywhere, and as single- or multi-agents willing to act on the international political scene. This raises the question of the State's responsibility for distributed form of cyber attacks by non-State actors, which, if aggregated, may rise to the level of warfare.

CYBER WAR, DEATH AND PEACE

There are two more important changes with regard to cyber war. It is no longer only or primarily conducted by human beings (Pagallo, 2011 and 2013), and, involving as it does information, it is no longer characterised by the Hobbesian relation between information and time. It is even difficult to state, in legal terms, when a cyber war begins and ends. In a pessimistic scenario, cyber war ceases to be a distinguished *tract of time,* instead becoming an underlying constant stream of strategic operations. From this perspective, the idea of war, conceived as cyber war, no longer fits its modern conceptual framework: namely, the ideas of death and of peace.

The idea of death

This idea brings us back to the Hobbesian construct of modern political thought. Hobbes's great intuition is that the process of political legitimisation does not frame or solve the conflict but, rather, stems from it. If we want to capture what legitimises a political power, we have to realise from which conflict this power stems and to which conflict it is meant to be the answer.

When war is conceived in the theoretical framework of death, its legitimacy stems from two conflicts. Firstly, it originates from a horizontal conflict between sovereign States in the international arena, and secondly from a vertical conflict between the sovereign State and its members who are called upon to potentially sacrifice their lives. Parliamentary authority to authorise war ultimate resides on this vertical conflict. Where war is no longer thought of in the theoretical framework of death because, for instance, it does not involve human combatants, its legitimisation ceases to be based on a vertical conflict. Rather, it is based solely on a horizontal conflict between states, which entrusts the process of war legitimisation to national government agencies. In this case, war is no longer considered and evaluated in relation to the primary value of human life, but becomes a matter of technological, economic, informational or other resources.

The idea of peace

The traditional concept of war is strongly linked to the concept of peace and, in many respects, it is at the foundation of this concept. The progressive transformation of the traditional concept of war in terms of cyber war is therefore also likely to affect the concept of peace. This is certainly of great importance, since the legal value of peace is or should be the basic principle of the international legal order.

The concept of peace is essentially procedural tied up with the concept of war. From Hobbes to Kelsen (1966), peace is understood as the absence of war. That is to say, as the absence of the illegitimate use of physical force. The negative and procedural conception of peace tends to turn peace into security, from which it should be distinguished as it is in Article 2 of the Charter of the United Nations. This concept of peace seems to be inconsistent with the notion of cyber war, which does not require the use of force in the traditional sense. Bobbio (1979) describes peace in procedural terms, but less negatively, as the legally sanctioned conclusion of

a war. This concept of peace can hardly be reconciled with cyber war, which does not necessarily have a legally sanctioned beginning and end (see the Stuxnet case: Richardson, 2011; and, more generally, Kesan & Hayes, 2012).

Thus the concept of cyber war cannot easily be associated with the traditional idea of peace and it requires us to revise not only the traditional idea of war but also of peace, and hence the foundation of the international legal order. This revision goes far beyond the scope of this paper, but two issues are particularly important. The concept of cyber war is no longer conceived in the framework of death, but rather of a competition between different allocations of strategic resources. Secondly, cyber war is no longer thought of in the horizon of peace, but of public security, which may be understood either in substantive (the factual conditions that enable sovereign States to have control over their life-cycle of information) or in formal terms (the shared norms that govern the sovereign States' cooperation for the control over their life-cycle of information).

We have insisted so far on the idea that cyber war differs from the traditional idea of war in that it does not necessarily make use of physical violence, although it may have indirect violent consequences. Physical violence thus marks a basic difference, at least in the premise, if not in the consequences, of kinetic and cyber attacks. However, the concept of violence is wider and more complex than the mere reference to the physical violence leads us to suppose.

THE MODERN IDEA OF VIOLENCE

In modern and contemporary political thought, violence is generally referred to and used as the theoretical foundation of the political order. Violence plays the role of a negative condition, from which a civil society is to be immunised in order to flourish as a stable and ordered political community. Violence is anthropologically founded: human beings are intrinsically violent, and violence is understood in kinetic terms as physical violence. This negative anthropology is already present in Luther and tends to characterise almost the whole tradition of modern natural law. The spark of physical violence is an energy that troubles the collectivity and needs to be controlled for the life of the political community to be possible. Two diverse ideas of violence originate here and find their theoretical formulation in Walter Benjamin's (1985) thesis on violence, which delineates the violence that *founds* the political order (including the law) from the violence that *preserves* it.

There is the idea that violence founds the existing political order. The original violence is to be immunised against: this allows the civil society to found a stable political order, which is then distinct from the premises from which it stems. Hence, the violence that arises from time to time is already included in the original violence, but is always of a lesser scale. The *imagining* of an original violence founding the political order turns out to be the unspoken justification of the existing political order, as magnificently accomplished in Hobbes.

There is also the idea that violence preserves the existing political order. The original violence is immunised against, but there is always a trace of it attached to the existing political order. The violence that arises from time to time is the undeletable trace of the original violence against which society cannot be completely immunised. The recourse to violence is what in the end assures the effective existence of that political order. The violence that preserves the existing political order proves that this order is never justified, but is always potentially unjustified and inclined to make recourse to violence when necessary to reaffirm itself.

This genealogy of modern political order, justified and premised over the hypothesis of an original violence where this origin is not necessarily conceived in chronological or historical terms, is not based on a concept of violence exclusively understood in terms of physical violence. Let us delineate two diverse forms of violence, concerned with its means and ends. The means of violence concern how violence is perpetrated. There are, naturally, many ways to manifest and accomplish violence, which are not limited to physical force. Verbal or psychological violence, for instance, do not require recourse to physical force, although the primary and basic manner to be violent is by means of physical force. Therefore, we define this form of violence as physical violence.

The ends of violence are different from how it is manifested and accomplished. Although there are exceptions, (for instance, blind violence or the so called 'systemic violence' (Zizek, 2008)) violence is generally meant to achieve something that goes beyond its manifestation. Thus, we define this form of violence as moral violence. We do not take moral to refer to anything ethically justified, but rather to an end not immediately perceivable in the manifestation of violence.

Throughout history, the idea of physical violence has been always coupled with the idea of moral violence, which plays a key, yet less visible, role in the foundation of the political order and in the justification of war. In some cases, the use or the impediment of physical violence is justified as far as it prevents moral violence. In other cases, the use or the impediment of moral violence is justified insofar as it stops physical violence. In this sense, physical and moral violence function as two normative systems that can be engaged or disengaged in order to serve as a justification for each other (for the relation between violence and morality, see Magnani, 2011; for a commentary, see Durante, 2013).

Let us sketch the interplay of physical and moral violence in Hobbes' and Locke's political philosophies, where this interplay underlies both the construction of their social contract theories and the conceptual relation between the State of Nature and the State of War.

Unlike Hobbes, Locke makes a distinction between the State of Nature and that of War. The Lockean State of Nature is the state in which human beings live together according to reason, without a common superior with authority to judge between them. The State of Nature is not characterised by the war of all against all (that is, by physical violence), but by the lack of a common judge to appeal to. This lack opens the way of getting justice ourselves, which is justified as the law of war when it comes to reject those who want to deprive us of freedom, or is unjustified when it conceals the use of force without right. The State of War is characterised by recourse to force or by its menace, when there is not a common superior to appeal to or when force is exerted without right, even if in presence of a common judge:

*'*Want of a common judge with authority, puts all men in a State of Nature: force without right, upon a man's person, makes a State of War, both where there is, and is not, a common judge.*'* (Locke, ed. 1998, 3.19)

Locke goes on:

*'*The law, which was made for my preservation, where it cannot interpose to secure my life from present force, which, if lost, is capable of no reparation, permits me my own defence, and the right of war, a liberty to kill the aggressor, because the aggressor allows not time to appeal to our common judge, nor the decision of the law, for remedy in a case where the mischief may be irreparable.'(Locke, ed. 1998, 3.19).

Again, morality consists in introducing a delay in the immediateness of life. Here, Locke introduces a different concept of violence, which concerns moral violence. It is apparent that, according to Locke, violence is primarily moral violence, which consists of getting justice by itself (being judge in our own case) through the use of force without right. Self-defence is justified and even required when the subject of aggression is in the absence of a common judge. The moral foundation of Locke's political philosophy is the delegation of justice – the appeal to a common superior. According to Hobbes, security may be achieved only through the stability of pacts and moral violence consists of the betrayal of pacts:

'Thus the nature of justice consists in maintaining the valid covenants, but the validity of the agreements will not start if not with the constitution of a civil power sufficient to compel men to keep them.' (Hobbes, ed. 2008, XIV)

The stability of pacts is promoted by the common interest to escape from the State of Nature dominated by physical violence and guaranteed by a civil power that compels people to maintain valid covenants. It is the fear of physical violence that leads us to overcome the risk of moral violence.

According to Locke, security may be achieved only through the delegation of justice, which draws the distinction between the State of Nature (moral violence: lack of a common judge) and that of War (physical violence: fear of death). We can free ourselves from physical violence only by overcoming moral violence. In the Lockean State of War there is no appeal: 'war is made upon the sufferers, who having no appeal on earth to right them, they are left to the only remedy in such cases, an appeal to heaven' (Locke, ed. 1998, 3.20); 'Where there is no judge on earth, the appeal lies to God in heaven' (Locke, ed. 1998, 3.21).

Physical violence is always coupled with moral violence. According to Hobbes, physical violence conceptually underlies moral violence: it is the disengagement from physical violence that assures the maintenance of pacts. According to Locke, moral violence conceptually underlies physical violence: it is the disengagement from moral violence that frees human beings from the threat of physical violence. Physical and moral violence are dissimilar but they have a common conceptual ground that allows us to define the essence of violence. Let us state it by referring to what Emmanuel Levinas has pointed out (for comments, see Durante, 2003):

'Violence is to be found in any action in which one acts as if one were alone to act: as if the rest of the universe were there only to receive the action: violence is consequently also any action which we endure without at every point collaborating in it.' (Levinas, 1990, 6).

Hence, violence is what turns an agent into a mere patient or what prevents a patient from becoming an agent. Violence is found in any action in which one acts regardless of another member or instance of the universe. The conceptual core or the essence of violence is its radical *regardlessness.* This also traces the limit one should never trespass when justifying the recourse to violence. Physical violence is never justified as such, but only by juxtaposing it with a moral violence to be avoided. In this sense, the use of force requires a moral engagement or, in the vocabulary of the just war tradition, a '*iusta causa'*, namely, a 'good reason'. We should then focus on the possible justifications of war, that is, on the tradition of just war *(Jus ad bellum).*

TWO THEORIES OF 'JUST WAR'

Traditional theories of Just War are mainly centred on the interpretation of what is the '*iusta causa*' (the good reason) for war (Viola, 2005, 55). There is no space in this paper to account for all Just War theories, but let us follow Francesco Viola (2005, 56-60), who discerns two traditional interpretations of Just War.

The School of Natural Law and the Modern Ius Gentium

According to the School of Natural Law and to the modern *Ius Gentium* (Grotius etc.), the '*iusta causa'* resides in the right to self-defence from aggression. As seen in Locke, all human beings and Nation States have the natural right to self-defence, 'because the aggressor allows not time to appeal to our common judge, nor the decision of the law, for remedy in a case where the mischief may be irreparable' (Locke, ed. 1998, 3.19). This is taken for granted by Vitoria, Suarez and Grotius, and is not subject to dispute. What is under discussion is the extent of the class of rights (goods or values), which authorise the State's reaction against the aggressor (Viola, 2005, 57-58). Life and freedom are naturally included, but property, for instance, is under debate, as are the controversial the limits of State's reaction: *Jus in bello.* What is common to all scholars, until Hobbes, is that self-defence is allowed only against the tangible threat of an imminent danger. It is Hobbes that introduces the idea (consistent with the essential role that imagination plays in his political philosophy: see chap. II-III of *Leviathan*) that self-defence can be preventive: that is to say, based on the supposed menace of a potential danger. It is a central, striking idea affecting the whole development of the notion of Just War. With regard to this preventive attitude, Hobbes formulates what we might consider the first account of what cyber war is and requires from those that govern:

'Since therefore it necessarily belongs to rulers for the subjects safety to discover the enemies counsels, to keep garrisons, and to have money in continual readiness, and that princes are by the Law of Nature bound to use their whole endeavour in procuring the welfare of their subjects, it follows, that it is not only lawful for

them to send out spies, to maintain soldiers, to build forts, and to require money for these purposes, but also, not to do this, is unlawful. To which also may be added, whatsoever shall seem to conduce to the lessening of the power of foreigners whom they suspect, whether by sleight, or force. For rulers are bound according to their power to prevent the evils they suspect, lest peradventure they may happen through their negligence.' (Hobbes, ed. 1998, XIII, 8).

This preventive attitude (which includes three basic informational strategies: discovering the enemies' intents; lessening their powers by sleight; and preventing their suspected evils) is of key importance, since prevention is not directed to restore a broken political international order, but rather to take part in its construction. A 'just cyber war' is thus characterised not only by a defensive or reactive role but, first and foremost, by an active or constructive one. This preventive attitude can become the most important side of cyber war. Would this mean that some forms of cyber war will be normalised and included in States' political strategies in international relations? Is preventive cyber war going to be a form of distributed control, at the international level, on strategic lifecycles of information? Namely, a form of control concerned with the discovery of intent, the lessening of powers and the prevention of evils? Our idea is that cyber war will not simply take the place of traditional war in many cases as a form of continuation of politics by other means. Cyber war, understood as exploitation rather than destruction, will be a steady and significant part of current international politics.

Middle Age Tradition

According to a different tradition going back to St Thomas and St Augustine, the '*iusta causa*' resides in the protection of the weak and the oppressed. In this sense, Just War is not motivated by self-defence, but by the need to protect someone else and to punish the aggressors for their faults. According to St Thomas, the justification of Just War is not defence but the protection of common good. St Thomas does not treat the issue of war as part of natural law or justice, but in relation to the virtue of charity: 'he does not wonder whether a war is moral, but whether it is always sinful to make a war, namely, in what case to kill another human being is not contrary to the love for the neighbour' (Viola, 2005, 60). Viola remarks that the protection of common good, which authorises a Just War, concerns:

'…both those who are to be protected and those who are unjustified aggressors. In fact, a just war is made in support of other people as well as in the interest of the enemies themselves. Thus, all the hypotheses envisaged by St Thomas concern the use of war as a sanction, which is intended to punish a fault and to fulfil the claims of justice. The recourse to war has to be premised upon a very serious injustice to be punished, as in the case in which a State does not sanction the violence perpetrated by its members or does not return what unlawfully obtained.'(Viola, 2005, 59-60)

War is thus justified when it comes to redress someone else's unjust sufferings *(Jus ad bellum),* and this is done in a proportionate manner *(Jus in bello).* This interpretation of Just War is not limited to the hypothesis of self-defence, but is primarily concerned with someone else's unjust sufferings, and thus it may embrace the protection of human rights. Viola remarks that 'it is with this middle-age tradition that we have to be deal, since it is more apt to interpret the current claims of just peace than the modern tradition is' (Viola, 2005, 60). This interpretation implies a clear understanding of the causes of injustice that authorise a war: namely, we need to share a list of values (constitutional principles, human rights, fundamental goods) the infringement of which makes someone else suffer from unjust causes that ask for intervention at international level. In this framework, Just War is intended to restore a broken international political order on the basis of recognised shared values. In this respect, we have to make two critical remarks.

Pluralism

It is hotly debated whether the recognition of a comprehensive list of values is possible when confronted with a pluralistic conception of human rights; and to what extent values, if recognised, are shared. These are difficult questions with issues of pluralism and universalism and cannot be dealt with in pure theoretical terms. The use

of legitimate force can be justified on moral grounds, but these are more concerned with the practical impediment to moral violence than with the identification of what is just. Thus, the question is: what is moral violence in the cyber age? Is it the infringement of pacts, as global interdependence raises a duty to trustful cooperation, or the lack of appeal to a common superior? We argue, in the last part of the paper, that the answer depends on a full appraisal of the informational nature of cyber war.

Preventive war

A Just War attack can be preventive when directed to the defence of the weak or the oppressed, and it is hardly arguable that the intervention is intended only to restore the broken political international order. A preventive intervention seems always aimed at participating in the construction of that order. The preventive intervention also has to show the moral grounds on which the reason to intervene rests, which differ from national interests and security reasons. Such reasons mainly refer to the protection of the public internal order of the international community.

JUST CYBER WAR

Let us rehearse what we have said so far:

- Cyber war changes the notion of war dramatically, and this change affects the idea of a possible just cyber war.
- Cyber war is hardly related to the traditional legal ideas of peace, either the procedural and negative idea of peace as the absence of illegitimate violence, or the procedural and positive idea of peace, according to which peace is the legally sanctioned conclusion of a war.
- Cyber war is not understood as a declared conflict between sovereign states, which is finally directed to re-establish peace, by re-assuring control over a territory and hence national (or international) security.
- Cyber war is conceived as an undeclared conflict and tactical competition between national agencies which is directed to reassure national or international security by having control over strategic flows of information. In this perspective, cyber war is less aimed at restoring a broken international order than at participating in its construction.
- Cyber war has a proactive rather than a merely reactive nature. This is consistent with the structure of the Information Society, which is increasingly an inference society, based on the pre-emptive capacity to anticipate future trends and behaviours.

Therefore, contrary to the emphasis placed at present on human rights that supports the idea of 'just peace' (Viola, 2005, 60), we think that a just cyber war is meant to endorse the Hobbesian tradition of preventive self-defence. Just cyber war is hence characterised by a proactive, constructive role, because of which it becomes a preventive form of control, at the international level, over strategic life-cycles of information. Since Hobbes, this form of control includes three fundamental informational strategies: the discovery of enemies' intents; the lessening of their powers by means of sleight; and the prevention of their suspected evils. All these activities are mainly based, as remarked, on the element of covertness:

'The element of covertness is a tricky area of international lawyers. It is an emerging area that will gain great resonance at state increasingly turn to covert cyber attacks to achieve their goals. There is no bright-line rule on whether a covert cyber attack will be held unlawful by the international community for the reason of its covertness; whether a covert cyber attack is held unlawful depends on any number of contextual factors. Who perpetrates the attack, who is the target, whether civilians are at risk, whether the intended outcome is to

coerce or to destroy, whether the target is afforded an opportunity preceding the covert operation to change its offensive behaviour, and whether the attack complies with jus in bello obligations are all relevant factors.' (Gervais, 2011, 31)

In the Hobbesian perspective, moral violence is overcome when pacts are stable and respected. In the absence of a superior power that compels States to respect the pacts, the stability of pacts depends on available information which enables States to predict whether there is 'something arising after the covenant made, as some new fact or other sign of the will not to perform' (Hobbes, ed. 2008, XIV). Therefore, moral violence is mainly concerned with tightening international collaboration between States.

The duty of collaboration arises from mutual interdependence and the need to cope with the distributed nature of cyber attacks, mainly perpetrated by non-State actors. The duty of collaboration is mainly directed at assuring that:

- a State adopts reasonable measures to prevent foreseeable cyber attacks from non-State actors that originate from its territory (Gervais, 2011, 20);
- a State adopts reasonable measures to discontinue (or make reparations for) the wrongful conduct of non-State actors that originate from its territory, when a series of incidents cannot be considered in isolation but, according to an accumulation doctrine, as a single armed attack;
- a State is accountable for the dangers of covertness of cyber attacks, which *'can* transform an otherwise lawful operation into an unlawful action under international law' (Gervais, 2011, 29); and
- a State provides other States with a sufficient level of information, in order to make them discriminate between combatants and civilians: this is more difficult to achieve when cyber attacks are run by non-State actors.

These cases are all concerned with the quantity and quality of information shared, or with data in any way gathered, which enable us to make predictions about future trends and behaviours. This means that cyber war is not only conceived in informational terms, but is a war on information through information. Let us consider more closely the informational nature of cyber war and the idea of informational moral violence. This may help us determine whether a cyber war is justified only in case of self-defence or also for the protection of human rights (i.e. in support of the weak and the oppressed).

The main difficulty in justifying the use of force for the protection of human rights resides in their pretended universality, as Viola puts it:

'The assertion of their universality often conceals the belief in the superiority of the western conception of human rights based on individualist philosophy, which is neither accepted nor shared by different cultures, notably, the oriental ones, more sensitive to communitarian values and collective rights (Asian values). Therefore, a war justified by the need to protect human rights would be easily used to impose the supremacy of western values and of the political and economic systems related to them.' (Viola, 2005, 61)

There is a possible reply to this argument. Not all human rights should be used to justify the recourse to force (Viola, 2005, 61). There is a limited number of fundamental rights which are so necessary that they enable the exercise of all the others (Shue, 1980, cited in Viola, 2005, 62). Any type of rights, whether liberal or communitarian, individualistic or collectivistic, is premised upon such fundamental rights:

*'*These rights have been defined as socially basic human rights: their respect is the minimal condition for human dignity. Certainly, they include security, that is, the right not be killed, tortured or aggressed (security rights), and the rights to subsistence (subsistence rights), namely, the right to adequate food, clothes, housing as well as clean air and water. It is debated whether negative freedom should also be included among these

elementary rights, as I believe. One might assume that all human beings, despite their cultural identity and particular theory of rights, should agree on the fact that being deprived of one of these fundamental rights is considered a serious violation of human dignity. This may be judged, under certain conditions, a good reason for war.' (Viola, 2005, 62)

Let us consider what is, in informational terms, a deprivation of fundamental rights that amounts at a serious violation of human dignity. This brings us back to the issue of what constitutes moral violence from an informational standpoint. It is obvious that physical violence is accomplished when a disruptive activity damages, deteriorates, deletes or suppresses an informational object. In this case, the entity seized by the violence ceases to be an informational object. This entity is no longer a source of information. The protection of the mere existence of an entity as a source of information is part of its security rights; that is, the right not to be destroyed or aggressed. When is moral violence then perpetrated?

Violence is found in any action in which one acts regardless of any other member or instance of the universe. The essence of violence is thus its radical *regardlessness.* Violence is what turns an agent into a mere patient or what prevents a patient from becoming an agent. A patient is, thereby, deprived of the fundamental capacity to become an agent and, hence, to become a *specific* source of information. This means that the violent act prevents an entity from becoming *that* source of information which it could have been had it not been subject to violence. For this reason, an information agent can no longer be a moral agent, since she cannot 'avail herself of some information (information as a *resource)* to generate some other information (information as a *product)* and, in so doing, affect her information environment (information as a *target*)' (Floridi, 2010, 102). From an informational standpoint, moral violence is the deprivation of such fundamental capacity: i.e. to be *that* specific source of information.

The deprivation of this capacity is part of the security rights of an informational agent, when it concerns its right not to be tortured (i.e. the right not to be *that* source of information it would not have been if not subject to violence), and so *vim vi repellere licet;* the impediment of informational physical or moral violence authorises the recourse to force. Informational security rights are part of the *socially basic human rights* that justify a cyber attack. The key questions are, therefore, whether or not to be *that* source of information may as well count as a *subsistence right, and* can the deprivation of informational subsistence rights authorise the recourse to force? The answer is to be found in the nature itself of affordances of subsistence rights. A subsistence right is considered a social basic human right, the disrespect of which is a serious violation of human dignity authorising the recourse to force, when it affords the possibility to exercise all the other human rights, whether liberal or communitarian, individualistic or collectivistic. Therefore, a specific source of information counts as an informational subsistence right, when it affords the possibility to exercise all the other rights. In this case, the protection of *that* source of information is a legitimate reason for war. The informational approach may have a further consequence. Being a specific source of information allows the agent to be what it is as a moral agent, and thus the informational subsistence right coincides also with negative liberty, if this is conceived as the necessary requirement for moral choice and human flourishing. This means that the informational approach widens the scope of subsistence rights, by including negative liberty within the socially basic rights that authorize the intervention to protect the weak or the oppressed.

CONCLUSIONS

This informational approach accounts both for the case of cyber destruction, which is meant to deprive an entity of its capacity to be *a* source of information, and for the case of cyber exploitation, which is meant to deprive an entity of its capacity to be *that* source of information it would have been if not attacked. Aware of the lesson of modernity, it couples physical violence with moral violence and provides us with some hints about what form of cyber attacks may be considered justified. It also tells us that, in the long run, the informational nature of cyber war will turn war into a strategic competition between national agencies for the control over the lifecycle of information at the international level. This warlike competition will not be a continuation of

politics by other means, but part of current international politics. Finally, the informational approach suggests to us that cyber war will inherit from the Hobbesian tradition a preventive attitude towards self-defence which is directed to participate in the construction of the international legal order. Since cyber war is no longer 'a tract of time' but something progressively displaying in covert areas, this will raise questions of transparency and accountability. Nonetheless, the violation of informational security and subsistence rights will authorise recourse to force to protect the weak and the oppressed. To what extent and in what circumstances the protection of informational security and subsistence rights will be considered a legitimate reason for war is left to future investigation.

REFERENCES

Benjamin, W. (1985), Critique of Violence *[1921], in One-Way Street, and Other Writings*, *Verso, London, 132-154.*

Bobbio, N. (1979), Il problema della guerra e le vie della pace, *Il Mulino, Bologna.*

Durante, M. (2003), Violenza e diritto nella riflessione d'Emmanuel Levinas. Riflessioni sul post-totalitarismo, in *AA. VV. Annali della Facoltà di Giurisprudenza dell'Università di Ferrara*, Nuova Serie, Vol. XVII, Giuffré, Milano, 141-164.

Durante, M. (2013), Notes on Lorenzo Magnani Understanding Violence, in *Mind & Society,* November, vol. 12, issue 2, 257-262.

Floridi, L. (1999), Information ethics: On the philosophical foundation of computer ethics, in *Ethics and Information Technology*,1, 37-56.

Floridi, L. (2010), *Information. A Very Short Introduction,* Oxford University Press, Oxford.

Floridi, L. (2011), *The Philosophy of Information,* Oxford University Press, Oxford.

Floridi, L. (2013), *The Ethics of Information,* Oxford University Press, Oxford.

Gervais, M. (2011), 'Cyber Attacks and the Laws of War' (October 1, 2011). Http://ssrn.com/abstract=1939615.

Hobbes, T. (ed. 1998), *De Cive* [1642], Cambridge University Press, Cambridge.

Hobbes, T. (ed. 2008), *Leviathan* [1651], Oxford University Press, Oxford.

Kelsen, H. (1966), *Principles of International Law*, II ed., Holt, Rinehart and Winston, New York.

Kesan, J.P. & Hayes, C.M. (2012), Mitigative Counterstriking: Self-Defence and Deterrence in Cyberspace, in *Harvard Journal of Law& Technology,* Spring, 25.2, 429-543.

Levinas, E. (1990), *Difficult Freedom. Essays on Judaism,* The Athlone Press, London.

Locke, J. (ed. 1998), *Two Treatises of Government* [1690], Cambridge University Press, Cambridge.

Magnani, L. (2011), *Understanding Violence. The Intertwining of Morality, Religion and Violence: A Philosophical Stance,* Springer-Verlag, Berlin-Heidelberg.

Nye, J.-S. (2004), *Soft Powers: The Means to Success in World Politics*, Public Affairs, New York.

Pagallo, U. (2011), Robots of Just War: A Legal Perspective, *Philosophy and Technology*, 24(3), 307-323.

Pagallo, U. (2013), *The Laws of Robots: Crimes, Contracts, and Torts*, Springer, Dordrecht.

Picone, P. (1995), 'Interventi delle Nazioni Unite e obblighi *erga omnes*', in Id. (a cura di), *Interventi delle Nazioni Unite e diritto internazionale,* Cedam, Padova,

Richardson, J.-C. (2011), 'Stuxnet as Cyberwarfare: Applying the Law of War to the Virtual Battlefield' (July 22, 2011). Http://ssrn.com/abstract=1892888.

Rousseau, J.-J. (ed. 1997), *The Social Contract* [1762], Cambridge University Press, Cambridge.

Shue, H. (1980), Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy, Princeton University Press, Princeton.

Viola, F. (2005), La teoria della guerra giusta e i diritti umani, in AA.VV., *Pace, sicurezza, diritti umani,* a cura di S. Semplici, Messaggero, Padova, 39-68.

Žižek, S. (2008), *Violence*, Profile, London.

The Autonomy of Automated Weapons

Giovanni Sartor
University of Bologna and European University Institute of Florence
Italy

giovanni.sartor@unibo.it

In this paper I shall introduce some basic ideas that may be useful in approaching discussion on autonomous systems and applying it to weapons, beginning with an account of the idea of automaticity, an attempt to provide a distinct notion of autonomy. Then, I shall discuss the idea of a subclass of autonomous systems, those that may be called teleological. I shall discuss what kind of delegation is involved in the use of autonomous and teleological systems, namely, cognitive delegation, and consider what aspects of cognition may be delegated to such a system. I shall also address more specific and extensive notions of autonomy, such as the idea of moral autonomy. Finally, I shall consider what kinds of liabilities may be involved in the use of autonomous systems. I shall not provide a discussion of the many ideas and definitions that that been provided in the literature, but I will present a particular perspective on the matter, relying in particular on Castelfranchi and Falcone (2003), to whom I refer the reader for a discussion of the literature.

AUTOMATICITY

Castelfranchi and Falcone (2003) base their definition of automaticity on their concept of delegation, delegation between a principal *A* and an agent or client *B* whenever 'A needs or likes an action of another agent B and includes it in its own plan'. Thus, for instance, with regard to an automatic door in a shop, the shop owner would be a delegator, but so would be also the shop clients, who know about the working of the door and rely on its action for entering in the shop. In a military context the delegator would be the person who is deploying the automated weapon, but also whoever is relying on the working of the weapon for achieving his or her objectives. Thus, not only would the remote human pilot of a drone be delegating to the drone the control of the flight until its destination is reached, but so would the commander who has ordered the mission and the soldiers who are waiting for the drone's strike before moving on.

According to Castefranchi and Falcone (2003) a system is automatic when:

- it performs an action by itself;
- the action is a task delegated to it;
- the action substitutes a human action of a delegator;
- it is artificial and its work is its delegated task; and
- it is teleonomic, having certain features which are intended to produce certain results.

As an example of an automated system, Castefranchi and Falcone mention the case of an automated door: the door opens by itself, and this is the task that has been conferred to it. Performance of this task substitutes the action of the person going through the door. The door is artificial and it has been constructed to perform its task. Certain features of the door, in particular those which enable or cause it to open when somebody approaches it, can be explained with regard to the fact that they enable the door to perform that task, since the door has been developed in order to be able to open automatically.

The same analysis can also be provided with regard to an automated weapon, such as a landmine. It detonates by itself when something passes over it, this being the function delegated to it; it substitutes for a war action that, at least in principle, could be performed by its user; it is artificial and has been developed for that purpose, which explains why it is the way it is.

This analysis of automaticity could be applied to simple artificial biological entities, such as an engineered virus, that are unable to learn individually, though they may learn collectively, as we shall see.

AUTONOMY (AUTO-TELEONOMY)

Developing the theory of Falcone and Castelfranchi (2003), we may see that an agent is autonomous when the following conditions hold:

- the agent's behaviour is auto-teleonomic: it adapts its behaviour to its purposes (such purposes being selected by the agent itself, provided by its user, or hardwired in its architecture);
- it interacts with its environment, getting inputs and providing outputs; and
- it adopts 'internal states' on which its behaviour depends.

According to the first criterion, a device merely having been constructed for a purpose does not count as an autonomous system, unless it has the capacity of continuing to align itself to its intended purpose when environmental pressures would otherwise cause a misalignment. Thus an autonomous system needs to have a feedback or homeostatic mechanism which keeps it focused on its objective as the environment changes. Consider how the automatic pilot in a drone needs to be able to react to changing environmental conditions, such as speed and direction of the wind, and adapt the flight so as to still be able to reach the target under variable conditions. This capacity is sometimes called autonomicity (capacity to govern itself), and distinguished from autonomy (capacity to make independent choices), which might exist independently from the first (see Truszkowsk, *et al.* (2009). However, here the term autonomy includes an element of autonomicity, as self-teleonomy.

Obviously, highly intelligent systems fully qualify for autonomy in this sense. This would be the case, for instance, of an autonomous car being able to conduct itself in such a way as to discharge its purpose, overcome various issues that may emerge during along the way such as encounters with other cars, signals, road blockages, and be able to plan and re-plan its route. Similarly, for a drone which is able to reach its destination and identify its target.

However, a remotely piloted aircraft system (RPAS) that maintains the direction established by the pilot by monitoring its position and adapting its flight to remedy possible deviations would qualify as autonomous under this description. Similarly, an intelligent bomb able to track its target, and to adjust its trajectory to the movement of the target would also qualify as autonomous. In both cases two remaining characteristics of autonomy would indeed be satisfied, namely, interaction with the environment and internal states. In fact the system's internal state would need to reflect the objectives of the system so that its behaviour continues to track those objectives under changing circumstances.

From our analysis it emerges that automaticity and autonomy are two distinct notions, only partly overlapping. There may be automatic and autonomous entities, such the car and the drone. There may be automated and non-autonomous entities such as the automatic door or the landmine. Finally there may be autonomous entities that are not automated such as non-artificial biological beings. Such beings would fail to be automated, according to the definition provided above, even when they perform a task delegated to them (e.g., a watchdog), missing the property of artificiality.

According to our definition, automated agents could get inputs from the environment, but would be merely reactive, reacting to each input with a predetermined response. We may consequently wonder whether we have an automated or an autonomous entity, according to our classification, in the case of an entity that is able to achieve its objectives under different environmental conditions, given its capacity to react in predetermined ways to such conditions. It seems to me that autonomy should be included when the entity has the capability to learn, increasing its capacity to face similar circumstances in the future; to repeat successes and to avoid

failure. Consider for instance the case of a target recognition system based on a neural network, which improves its performance through reinforcement learning, enhancing connections enabling the correct identification of a target and demoting those leading to mistaken identification, in ways not foreseen when the system was put in place. Equally, we may wonder whether a homeostatic system for maintaining a certain temperature would qualify as autonomous, when reacting in a predetermined way to any possible misalignment from the target temperature.

Figure 2

Teleological systems

Finally, certain autonomous systems can be characterised as teleological systems. A teleological system is characterised by the fact that it has an explicit representation of its cognitive structures such as:

- goals (representational structures that are meant to determine the environment - mind to world orientation);
- beliefs (representational structures meant to track aspects of the environment - world to mind orientation); and
- self-constructed plans (representational structure that specify how to reach the goals given the beliefs).

A teleological system thus includes the features that characterise the BDI (belief-desire-intention) model for intelligent agents (Rao and Georgeff, 1993), if we characterise desires generally as goals (abstracting from the emotional aspects of desires) and adopted plans as intentions.

A teleologic system, according to this characterisation, is necessarily autonomous, in the sense of being teleonomic, since it selects its behavior exactly in order to adapt it to its goal. Such a system should perform epistemic cognition, that is it should form new beliefs on relevant aspects of the environment, given pre-existing inputs and beliefs, as well as practical cognition, forming new subgoals and plans, given pre-existing goals and beliefs. As an example of a teleologic system, consider for instance a drone which has the goal of destroying a target, which requires that it fly to the target zone, identify the target, and then select and implement a way to eliminate it. Such a drone would store its goal, acquire inputs from the environment, process such inputs to determine relevant environmental conditions and identify its target, develop and implement flight plans to reach the target, and then carry out plans to destroy it.

According to the characterisation we have provided above, teleological systems are a strict subset of autonomous systems. For instance, a sensor system based on a neural network, which discriminates different signals and learns to improve its discriminating capacity through a supervised learning process, qualifies as an autonomous system, but not as a teleological one, being teleonomic but not teleologic.

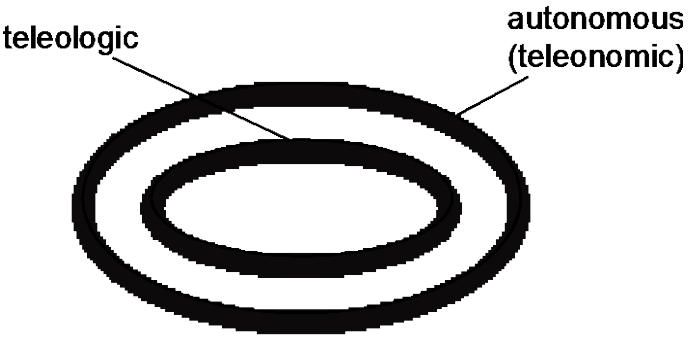

Figure 3

In a non-teleologic but autonomous (teleonomic) system, the system's behavioural patterns are selected by a mechanism of the system on the basis of the fact that they achieve the purpose of the system, but cognition is only implicitly represented in the system's internal state (the internal state has been selected on the basis of environmental responses and the system's purpose, but does not model explicitly either of them). Not only individuals, but also collectives of them could have a teleonomic but not teleologic behaviour. This is the case for collectives endowed swarm intelligence: the purpose-oriented behaviour of the collective emerges from the actions of the individual, though this purpose is not present in the individuals themselves, which blindly reproduce simple pattern of action. This may apply to drones flying in a flock, where each drone only keeps the distances from the other, land vehicles involved in the elimination of landmines, etc.

Figure 3 shows the outcome of the classification I have provided. Teleologic system are a subset of autonomous systems, in the sense of teleonomic ones. Automatic systems will be teleologic and/or teleonomic if they are artificial.

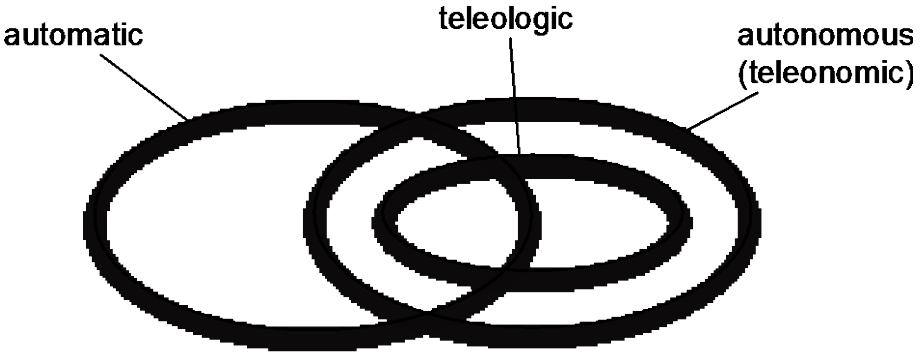

Figure 4

Figure 4 provides for a recap of the distinctions that we have introduced. Note that the arrow denotes inclusion between subset and superset.

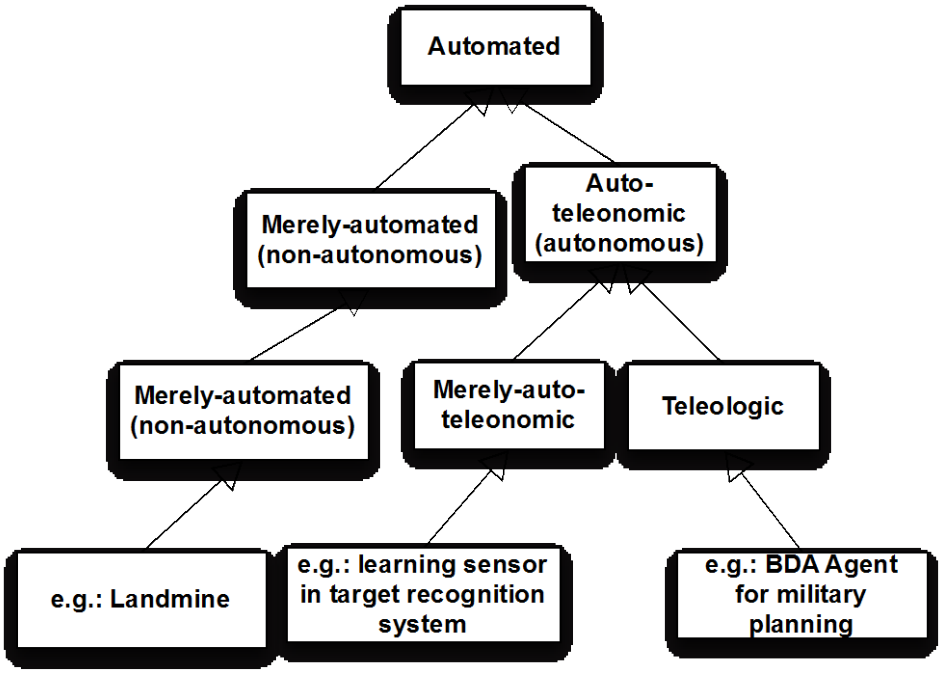

Figure 5

Let us try to characterise more precisely what it is for an agent to have cognitive states, focusing in particular on the distinction between goals and beliefs.

The mental states to autonomous systems

We may say that the internal state of an entity is a belief concerning the existence of certain external situations when:

- there is a world-to-mind covariance (Dretske, 1986) between the internal state and these situations; and
- this covariance enables the entity to react appropriately to the presence of these situations.

Similarly we may say that the internal state of an entity is a goal concerning the existence of certain external situations when:

- there is a mind-to-world covariance between the internal state and these situations; and
- this covariance enables the agent to implement its purposes.

Intentions or plans can similarly be characterised by the agent's revisable commitment to execute them under the indicated conditions (Bratman, 1986; Castelfranchi and Paglieri, 2007).

Concerning the conditions under which the attribution of cognitive states to an artificial entity is possible, we may refer to Dennett's (1989) idea of the intentional stance according to which the behaviour of a complex being, whose internal structure is unknowable, can only be explained and anticipated:

- by assuming that the being has a purpose, and that its behaviour is a way to achieve that purpose (the design stance); and
- by attributing intentional states to the being, and that its behaviour results from choices through which the agent aims to reach its goals according to its beliefs (the intentional stance).

We may however wonder whether the mental states of an artificial system are just in the eye of the beholder, or whether they reflect intrinsic aspects of the system. I think that the latter position holds, at least with regard to artificial teleological systems where we can identify the representational structures constituting the mental states at issue.

The mental states of a teleological system we have identified do not exhaust the possible architecture of such a system, additional components of which enable increased levels of autonomy. For instance, Pollock (2004) distinguishes values from desires, arguing that while desires (goals) prompt us to develop plans of actions, we use values to assess the comparative merits of our choices. Similarly, Castelfranchi and Paglieri argue for the need to distinguish different kinds of goals according to the role they play in guiding human action. Thus an agent able to be guided not only by desires but also by values would gain an increased level of autonomy/teleology.

Another important extension for a teleological agent would consist in its being guided by norms, operating as (possibly defeasible) constraints over the agent's teleological reasoning on norms in robots for the military (Arkin, 2009). Such norms, in their turn could be subject to critical evaluation according to the values endorsed by the agent, or attributed to its social environment.

In general, we can say that an agent is reflectively autonomous when it has access to its cognitive states (beliefs, desires, intentions, values), and is able to assess their merit on the basis of its other cognitive states: should I have this belief, given the evidence which I have; should I have this intention, given my objectives; should I endorse this norm, given its nature and the effects of complying which it; should I have such objectives, given my values; is this value really so important as I have assumed it is? I shall not focus more on reflective autonomy, since it seems to be largely out of the scope of existing artificial entities, though various philosophers have focused on it.

COGNITIVE DELEGATION AND AUTONOMOUS-AUTOMATICITY

In all instances of automaticity we have the delegation of a task to an artificial being. However, in the case of autonomy, not only do we expect that such an artificial entity will perform a certain task, we also expect that it will maintain its alignment to the task, processing incoming inputs and acting accordingly. This means that when delegating an operation to an autonomous auto-teleonomic system we delegate not only behaviour, but also the choice of actions, their implementation and controls over them; we delegate practical cognition. In this kind of delegation, that we may call cognitive delegation, the following generally holds:

- the delegator does not know and thus does not want what the agent will choose to do in future situations (no mere automaticity);
- the delegator has chosen to delegate the choice to the autonomous delegates since:
    - he prefers not to make that choice (this is the case, for instance of the remote pilot, who is letting the automatic pilot drive the drone to destination); or
    - he is unable to do make choice in the given framework (consider for instance the pilot of an autonomous drone, in a situation when the drone has lost connection or has to respond to an attack with a speed exceeding human reaction time; as another example consider a high speed trading system which has to take trading decisions in fractions of seconds.

This is not a limitation or a failure in the autonomous system, but rather the very reason why it is used; to substitute human cognition when it is not needed or not available.

To further analyse cognitive delegation, we need to ask ourselves what capacities are involved in the delegated tasks. We may in this regard distinguish different domains of automation: the mere acquisition and classification of input information (e.g., a sensor system); the analysis of this information to extract from it

further information (e.g. weather forecasts); the decision and action selection which may be the adoption of a plan of action (e.g., a flight route); and the implementation of the plan, which involves monitoring its execution (flying according to the established route). In these different domains, automation could reach a different level: merely supporting human information processing (as in the use of autonomous sensor-systems); being integrated with it (as when the system proposes options among which the human decides); or substituting human information processing, humans possibly having a residual monitoring function. It must be considered that even when the human remains in the loop, for instance having the task of pressing the button when prompted by the autonomous technology, or being able to monitor its behaviour and override it when necessary, the human contribution to the process may be very limited if the human does not have the information for taking a deliberate choice, or does not have the time and skill for processing it.

In a war scenario, the substitution of humans is most problematic when the adoption and execution of a plan of action is at issue. In fact the adoption of a plan involves not only means-end considerations, but also proportionality; an assessment of the expected collateral damage (on the morality of war, I refer to the seminal contribution of Walzer, 1977, 2006). Moreover, the execution of the plan involves the respect of norms pertaining to *jus in bello*, such as the principle of distinction, which requires the autonomous system to identify and follow norms of behaviour, having assessed the circumstances for their application.

In general, when a task is allocated to an autonomous artificial entity we need to analyse the cognitive functions that are involved in the task and examine whether the deployed automata possesses all the required epistemic and moral skills.

Many interesting issues pertain to engineering the norm-following behaviour of an autonomous robot (for various considerations, see Arkin (2007)). In particular we have to consider, taking into account the risks involved in the allocated tasks and the capacity of the robot, whether:

- it should be impossible for the robot to act against a norm (norms would be overriding exceptions in the robot's architecture, so that normative constrains would override means-end reasoning); or
- it should be possible for the robot to act against a norm, depending on the outcome of its deliberative process, norms providing only defeasible constraints.

RESPONSIBILITY FOR THE WORKING OF AUTONOMOUS SYSTEMS

In this last section I shall briefly address the issue of the responsibility for possible damage caused by the functioning of an artificial autonomous system. First of all, we need to ask ourselves what we mean by responsibility:

- mere causality in the production of the damage (e.g., causing harm to an unseen civilian);
- intentional causality in the production of the damage (e.g., causing harm to the targeted person);
- accountability for the effect of one's action (e.g., obligation to respond to question on why a civilian was harmed);
- blameworthiness for failing to act appropriately (e.g., for harming the civilian); or
- subjection to punishment or obligation to compensate, for violating a norm (e.g., violating distinction).

The first three notions of responsibility may well also apply to autonomous artificial entities. The first is applicable to any such system. The second presupposes a teleological system, if we limit the idea of intention to the cases where a plan is selected and stored in the agent. This third also seems applicable to an autonomous teleological system, to the extent that it keeps a record of its behaviour and is able to introspect its deliberative processes.

The last two responsibilities may apply to artificial agents only to a very limited extent, namely that to which the behaviour may be influenced by moral emotions or by the expectation or implementation of sanctions (see Sartor, 2009; Chopra and Which, 2011; Pagallo, 2013). Robots, having the appropriate motivational structure, may be the addressees of such responsibilities, but this cannot apply to all robots, and so cannot be made into a meaningful legal rule at the current state of the art. It also remains to be established what sanction against a robot would make sense, though their motivational structure may possibly be constructed in such a way as to be responsive to whatever may count as a sanction.

If autonomous agents are in general not going to be blameworthy or punishable, their users and controllers may be blameworthy or subject to sanction. The user may be responsible not only when the agent is assigned a forbidden task (an action which in itself, even when performed by a human, would violate the applicable rules), but also on the following grounds:

- negligent control over the performance of the autonomous agent (if control was possible);
- faulty design, resulting in a defective system which is below the state of the art;
- improper delegation in assigning to the autonomous agent a task that cannot be automated (taking into account side effects and norms governing it), at the state of the art; or
- strict liability for certain kinds of damages, such as those pertaining to accidents in air or land traffic.

CONCLUSION

Autonomous agents will soon become commonplace in the military domain. It is important to understand the qualities and skills of such agents when deploying them. The taxonomy proposed here (automatic, autonomous and teleological agents) may be useful for this purpose, though more refined classifications may also be needed.

In deploying such agents, the cognitive functions being delegated should be carefully analysed to determine whether the delegated agents possess all the required skills. An inappropriate delegation, negligent or intentional, may determine a responsibility on the side of the delegator.

REFERENCES

Arkin, R. (2009). *Governing Lethal Behavior in Autonomous Robots*. CRC Press.

Bratman, M. (1987). *Intentions, Plans and Practical Reasoning*. Harvard University Press, Cambridge, Mass.

Castelfranchi, C. and Falcone, R. (2003). From automaticity to autonomy: The frontier of artificial agents. In Hexmoor, H., Castelfranchi, C., and Falcone, R., eds, *Agent Autonomy*, pages 103–36. Springer.

Dennett, D. C. (1989). *The Intentional Stance*. MIT, Cambridge, Mass.

Castelfranchi, C. and Paglieri, F. (2007). The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intention. *Synthese*, 155:237–63.

Chopra, S. and White, L. F. (2011). *A Legal Theory for Autonomous Artificial Agents*. University of Michigan Press, Ann.

Rao, A. S. and Georgeff, M. P. (1991). Modelling rational agents within a bdi architecture. In Fikes, R. and Sandewall, E., eds, *Proceedings of Knowledge Representation and Reasoning (KRR)*, pages 473–84. Morgan Kaufmann, San Mateo, Cal.

Sartor, G. (2009). Cognitive automata and the law: Electronic contracting and the intentionality of software agents. *Artificial Intelligence and Law*, 17:253–90.

Truszkowski, W., Hallock, L., Rouff, C., Karlin, J., Rash, J., Hinchey, M. G., and Sterritt, R. (2009). *Autonomous and Autonomic Systems*. Springer.

Walzer, M. ([1977] 2006). *Just and Unjust Wars*. Basic Books.

Three Legal Challenges of Informational Warfare: On Force, Proportionality, and the Role of Sovereign States

Ugo Pagallo
Department of Law, University of Turin
Italy

ugo.pagallo@unito.it

The starting point of the analysis of whether, and to what extent, today's information revolution affects the current laws of war (LOW) and generally speaking, the framework of international humanitarian law (IHL), has to do with a basic fact. Whereas, over the past centuries, human societies have used information and communication technology (ICT), but have been mainly dependent on technologies that revolve around energy and basic resources, today's societies are increasingly dependent on ICT and, moreover, on information as a vital resource. In a nutshell, we are dealing with ICT-driven societies (Floridi, forthcoming); so, the more current societies are ICT-dependent, the more it is likely that both the causes legitimating war and the behaviour admitted in warfare will concern 'actions by a nation-state to penetrate another nation's computers or networks for the purposes of causing damage or disruption,' as the former U.S. counterterrorism chief Richard Clarke defines cyber warfare, or informational warfare (in Schmidt and Cohen, 2013: 103).

Such informational aggression, to be sure, can be an instrument for real world operations. For example, in the words of the U.S. Secretary of Defense, Leon Panetta, by penetrating another nation's networks and computers 'an aggressor nation... could derail passenger trains, contaminate the water supply in major cities, or shut down the power grid across large parts of the country' (in Schmidt and Cohen, 2013: 104). However, current LOW and IHL may fall short in coping with the new scenarios of informational warfare, once we take into account how the notion of force changes. Both LOW and IHL aim to regulate the use of physical violence as a last resort option, namely as one of the conditions rendering wars just *(jus ad bellum),* but informational attacks do not necessarily entail any physical violence, because they may aim to remove or annihilate informational resources from cyberspace, or delete them without backup. Hence, 'at what point does a cyber attack become an act of war?' (Schmidt and Cohen, 2013: 114). Do cyber attacks affect the causes legitimating war, such as the notion of reasonable success, or the right intention of the proper authority that enters the war? Do informational attacks impact on the right ways to behave on the battlefield, such as the proportionate use of force, discrimination, non-combatant immunity, and military necessity that makes collateral damage legal?

In order to provide a hopefully comprehensive view on these issues, the paper is presented in four parts. The next section will summarise the traditional framework of both LOW and IHL, namely the 1907 Hague Convention, the four Geneva Conventions from 1949, and the two 1977 additional Protocols, and the traditional distinction between causes *(jus ad bellum*) and conditions *(jus belli*) of just wars. Although lawmakers have added a third scenario, that is the provisions for the aftermath of warfare, or *jus post bellum*, the classical bifurcation suffices to describe basic tenets of today's legal framework that may be affected by a new generation of informational attacks.

In Section 3, attention is drawn to the (more than) ten year old-debate on the employment of robotics technology on the battlefield (Pagallo, 2011). Despite the differences between the field of military robotics and a new generation of cyber attacks, the former help us understand how the latter may upset basic pillars of LOW and IHL, to the extent that a new international agreement should govern both fields. This part of the paper focuses on what is specific to the new scenarios of informational warfare.

Section 4 examines the principle of proportionality, as both a cause of *jus ad bellum* and a condition of *jus in bello*, in order to determine whether the good achieved by an informational attack is proportionate to the evil of waging it, and how to link the level of virtual, as opposed to physical, force to the military ends that a nation aims to attain in the real world. Here, the 'four moral laws' of information ethics (Floridi, 2008, 2013) are particularly useful.

Section 5 considers how the role of sovereign States as the only war-declaring authorities in the international arena may change, since identifying the responsible party of an informational attack can be impossible and sovereign States will increasingly be unable to monopolise the use of force in cyberspace, the new domain of military operations. Although this trend of the privatisation of war is not new *(see, for example,* George W. Bush's 'war on terror'), it is likely that the new frontiers of informational warfare will exacerbate it (Pagallo, 2013).

The conclusions of the paper assess today's state of the legal art, in order to determine whether we should find a 'reasonable compromise' on the basis of legal expertise (Hart, 1961), or whether through the principles of the system that fit with the established law, a 'right answer' can be found for every case at hand (Dworkin, 1985).

THE POST-WESTPHALIAN FRAMEWORK

Two thousand years of debate on the characteristics of a just war were eclipsed three centuries ago in the modern Western world. Older Just War theories no longer made sense after the triumph of modern legal positivism and the 'paradigm of Westphalia' (1648). In the classical phrasing of Thomas Hobbes in *Leviathan,* '[it] is annexed to the sovereignty the right of making war and peace with other nations and Commonwealths; that is to say, of judging when it is for the public good, and how great forces are to be assembled, armed, and paid for that end' (Hobbes, 1999). By admitting that no one is set to judge the decisions of sovereign States, no room was left to ascertain the lawfulness of the causes of war, as the law is made up by a set of rules effectively established by national sovereigns. The immunity of sovereigns finally ended with the Nuremberg trials (1945-1946), and projects for a permanent International Criminal Court (ICC) culminated with the Treaty of Rome in October 1999 and the ICC's work in The Hague from 1 July 2002. Far from claiming that a Kantian cosmopolitan paradigm has replaced the old legal system within current international humanitarian law, it was only with the end of the Cold War (1989) and the first Gulf War (1991) that the topic of Just War again became a popular topic of debate among lawyers.

Legal scholars have increasingly debated in the past two decades the many conditions that make a war just: whether a legitimate claim exists, whether violence can be admitted as a last resort, whether there is a probability of success and proportionality in the use of force. Matters of proper authority have also been discussed, as has whether a declaration of war is always necessary. Without entering the philosophical debate on the just causes of wars *(e.g.* Walzer, 1977), the traditional distinction is between the causes legitimating war in the legal sense of *jus ad bellum*, and the behaviour admitted in warfare in the legal sense of *jus in bello.* As to the preconditions for war to be deemed just, we have to further distinguish between formal and substantial criteria: according to a basic tenet of the Westphalian paradigm, wars need to be declared by the competent authority of national sovereign States, so that such authority can be held responsible for operations occurring in the course of the war. In *The Better Angels of Our Nature*, Steven Pinker (2011) insists on this latter point: the Leviathan's monopoly on the legitimate use of force should be enlisted among the 'historical forces' that 'have driven the multiple declines in violence,' by defusing the temptation of exploitative attack, inhibiting the impulse for revenge and circumventing 'self-serving biases'. We return to this below in Section 5.

As to the substantial reasons of just war in the sense of *jus ad bellum*, the just causes of war traditionally comprise the good intention of the war-declaring authority, the reasonable success of war and especially in today's context, the use of force as the last option. In the phrasing of Article 51 of the UN Charter, 'nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack

occurs against a Member of the United Nations,' and thus force, save in self-defence, can only be used if authorised by the UN Security Council. Furthermore, the use of force should be proportional to the good achieved by war, so that, as a necessary condition for legal *jus ad bellum*, States cannot wage a massive war so as to remedy a trivial wrongdoing, as it occurred with the 1969 Soccer War between El Salvador and Honduras.

In addition to the causes of just war, a second set of legal provisions concerns principles of military conduct and rules of engagement as criteria of Just War in the sense of *jus in bello.* When analysing the conditions rendering conduct lawful on the battlefield, scholars usually distinguish between discrimination and non-combatant immunity, the doctrine of double effect and the principle of proportionality. The focus of the legal analysis is thus on the military necessity in fixing criteria for the target identified as a legitimate combatant: once political and military authorities have granted the use of force, such 'military necessity' may allow collateral damage, in accordance with the doctrine of double effect and the tactics for engagement, approach and standoff distance. The doctrine of double effect, however, should be further understood in connection with the principle of discrimination and non-combatant immunity that requires distinguishing between civilian and combatants and between friends and foes, so as to direct force only against enemy military objectives. The principle of proportionality also sets the necessary conditions for legal *jus in bello* that impose further restrictions on warfighting techniques: no unnecessary violence can be used in order to attain one's military ends; rather, the level of force should be proportioned to the goal of attaining such ends.

This legal framework may fall short in coping with the challenges of the information revolution and more particularly, in the field of military robotics with the use of unmanned aerial vehicles (UAVs), and systems (UAS). Research and development (R&D) in UAVs and UASs and their deployment have increased markedly over the last decade. Whilst more than forty countries are currently developing autonomous weapons and other types of robot soldiers, a 'Teal Group's 2013 market study estimates that UAV spending will more than double over the next decade from current worldwide UAV expenditures of $5.2 billion annually to $11.6 billion, totalling just over $89 billion in the next ten years,'[96] as the Teal sales representative in your area would be keen to illustrate. A 2011-2020 forecast by the HIS Industry Research and Analysis group claims that the U.S. will invest 56% of the global R&D in UAVs, China 12%, Israel 9%, Russia 8%, Pan- European research 3% and Britain, France and Italy 2% each. As a result, no Sci-Fi imagination is necessary to suspect that the massive employment of artificial soldiers will affect and is already impacting on a number of crucial fields in the legal domain.[97]

ROBOTIC WARNINGS

There are three reasons why robotics technology may affect both causes and conditions that make wars just, that is both *jus ad bellun* and *jus in bello.* First, in the opinion of several scholars, that which makes robot wars unjust hinges on the technical difficulty of designing robots which can distinguish between friend and foe, and between civilians and combatants, as a crucial condition of Just War in the sense of *jus in bello.* The failure to be able to do this was admitted by a 2008 research sponsored by the U.S. Department of the Navy, namely *Autonomous Military Robotics: Risks, Ethics, and Design.* In the words of Lin, Bekey and Abney, laws of war and rules of engagement 'leave much room for contradictory or vague imperatives, which may result in undesired and unexpected behaviour in robots'. Therefore, whilst 'it is morally unjustifiable to deploy military robots before we have any idea of their risk to non-combatants' and 'we may paradoxically need to use the first deaths to determine the level of risk,' Lin, Bekey and Abney acknowledge that 'whether or not robotic weaponry will soon be able to surmount the technical challenge of this moral imperative (at least as well as human soldiers) remains unknown.' Accordingly, following the suggestions of John S. Canning (2008) in

[96] See Teal Group's market study 2013, more information available at:
http://www.tealgroup.com/index.php/about-teal-group-corporation/press-releases/94-2013-uav-press-release

[97] In addition to LOW and IHL, think about constitutional law, tort law, administrative law, contracts, etc. (Pagallo, 2013).

Weaponised Unmanned Systems, a solution could be that robots target only weapons or, following the proposal of Noel Sharkey (2008) in *Grounds for Discrimination*, robot soldiers could be limited to operating only in particular regions or situations. Do such criteria make sense in the field of cyber wars?

The second reason why robotics technology would change the causes considered to make wars just concerns how the use of autonomous machines reduces the barriers to war and lowers the level of public awareness. While civilians targeted by AI attacks often consider those who send machines to fight for them 'cowards,' the reasons for sending robots to the battlefield may fade away, as shown by the new generation of drones that the U.S. CIA's civilian counsels authorise to attack almost every day. A fully automated military mission transforms war into a fairly technical and bureaucratic operation, risk-free so to speak, so that causes of war may also be trivial, once you imagine both armies engaging no humans but only robot soldiers. In the phrasing of Peter Asaro's (2008) *How Just Could a Robot War Be?*, 'this is the belief that these technologies will make it easier for leaders who wish to start a war to actually start one.' Does this warning apply to the new scenarios of informational warfare?

Third, the autonomy and unpredictability of the behaviour of AI machines would make robot-wars profoundly and irremediably unethical, because no human can ultimately be held responsible 'in relation to deaths caused by an autonomous weapon system.' This argument, illustrated by Robert Sparrow (2007) in *Killer Robots*, has obvious repercussions in the legal field, since the capacity of robots to operate in the real world without human control impacts on a very core principle of the laws of war. Wars need to be declared by a competent authority which can be held accountable for deaths occurring in the course of the war. If robots may cause serious harm by taking their own decisions, it is but a short step to envisaging robots that may provoke accidental wars. Does this risk reappear in the context of informational warfare?

Some other scholars suggest that the behaviour of robot soldiers on the battlefield does not fall within the loopholes of current legal systems, since the use of analogy, much as the principles of the law, would allow us to properly tackle the challenges of technology. Consider what Philip Alston stressed in the 2010 Report to the UN General Assembly on extrajudicial, summary or arbitrary executions, *vis-à-vis* the provisions of current international humanitarian law (IHL). In the wording of Alston, 'a missile fired from a drone is no different from any other commonly used weapon, including a gun fired by a soldier or a helicopter or gunship that fires missiles. The critical legal question is the same for each weapon: whether its specific use complies with IHL.' On this basis and by the use of legal analogy and the principles of the system, the conclusion is that most of the parameters determining when a war should be deemed just would not be affected by the employment of robot soldiers. This is the case of self-defence and right intention of the authority, as well as the hypothesis of reasonable success and war as a last resort option. Does the conclusion apply to the field of informational warfare as well?

All in all, taking into account issues of discrimination and non-combatant immunity, barriers to war, proper authority and the loopholes of current legal frameworks, no one-size-fits-all answer seems at hand. For example, going back to the principle of discrimination and non-combatant immunity, the first reason why some scholars believe that military robotics technology challenges the current laws of war, it seems fair to affirm that the aim to direct informational attacks against only military objectives, such as the enemy's networks and computers, does not raise any insurmountable technical problem. Still, as to the potential lowering of the threshold of entry into war, strict informational attacks present a crucial peculiarity, in that identifying the responsible party of such attacks is often impossible. This impossibility not only challenges the principle that wars need to be declared by competent authorities, but also reverberates on the traditional barriers to war, because anonymity may trigger new temptations for exploitative attacks while sheltering the anonymous informational offender from the reaction of others.

This latter scenario pinpoints a new parallel between informational warfare and military robotics technology, since the increasing complexity of network-centric operations, and the miniaturization of lethal machines, can make very difficult to detect the locus of political and military decisions (Pagallo, 2013).

Yet, current debate on military robotics technology and its impact on the laws of war suggest a further distinction between informational aggressions that represent a means for real world operations and cases of informational attacks that do not entail any physical force; strict informational attacks that aim to remove or annihilate informational resources from cyberspace, or delete such resources without backup. In this latter case, the use of analogy seems inadequate to determine whether force is applied in proportionate ways, as both a cause *(jus ad bellum)* and condition *(jus belli)* of Just Wars, because the type of weapon that an informational attack could be is still uncertain in the legal field, much as the ways in which such aggression should be interpreted in accordance with the principle of proportionality. How can we determine that the good achieved by a strict informational attack is proportionate to the evil of waging it? How can we compare the level of virtual, as opposed to physical, force with the military ends that a nation aims to attain in the real world? What is the level of abstraction that allows us to grasp the point of convergence between the traditional monopoly that sovereign States claim on the legitimate use of force and the new scenarios of informational warfare?

NEW SCENARIOS OF PROPORTIONALITY

Current laws of war are silent on the set of parameters and conditions that should regulate cases of informational aggression. Some provisions, such as those of the Budapest Convention on cybercrime and the definition of illegal access to networks and computer systems, illegal interception of non-public transmissions of computer data, system interference, and so forth, do not help in this context. That which is under scrutiny here concerns whether the actions of a sovereign State, which causes damage or disruption by penetrating another nation's computers, can be deemed 'just.'

We must distinguish between strict informational attacks and informational aggression as a means of real world operations. In this latter case, analogy helps us in tackling most of the challenges of informational attacks, since their impact on real world targets, such as power grids or water supplies, can conveniently be grasped with the traditional scenarios of just war theory, so as to determine whether such attacks abide by today's provisions of LOW and IHL. Yet things may appear tricky when the aim of the informational aggression is either to remove or annihilate some informational resources in cyberspace, or to delete them without backup. The difficulty concerns how we should evaluate a nation's aim to attain a military end in the real world by the means of virtual, as opposed to physical, force. The difficulty has to do with the ways in which the good achieved by an informational attack can be compared with the evil of waging it on, say, the internet. Since we lack specific rules on the subject matter, such as provisions that regulate the use of strict informational attacks in the laws of war, how can scholars address this set of legal issues?

A fruitful approach is given by Luciano Floridi's ethics of information as an 'ontocentric,' 'patient-oriented' and 'ecological macro ethics' (Floridi, 2008, 2013). By rejecting a rigid methodological anthropocentrism, this approach calls for a wider perspective than that based exclusively on the role of human agents. This informational outlook also suggests a different understanding of the interaction between agents and receivers or reagents, assuming the 'level of abstraction' which asserts that all entities should be represented in terms of information. In the phrasing of Floridi (2008: 21), 'all entities, *qua* informational objects, have an intrinsic moral value, although possibly quite minimal and overridable, and hence can count as moral patients, subject to some equally minimal degree of moral respect understood as a disinterested, appreciative and careful attention.' The aim is not only to explain how interacting agents communicate and share informational resources by means of positive or negative messages: in accordance with the ontocentric stance of this theory, the tenets of information ethics provide a unified perspective for varying statuses and regimes that concern the content of such resources, regardless of the specific technologies with which we are dealing and in an impartial

and universal way (Pagallo and Durante, 2009). What Floridi calls the ontological equality principle means that the resources of the system are deemed informational entities that should morally be treated as part of the environment or 'infosphere,' bringing to 'ultimate completion the process of enlargement of the concept of what may count as a center of moral claim' (Floridi 2008: 12). As a result, a universal normative framework should govern the life cycle of information within the infosphere in a field-independent way and in connection with the ontological equality principle, in an impartial manner. More specifically, this normative framework hinges on the concept of informational entropy, which is structured according to four moral laws. Whilst informational entropy 'refers to any kind of destruction or corruption of informational objects (mind, not of information), that is, any form of impoverishment of being' (Floridi 2008: 11), the four moral laws command that:

1. Entropy ought not to be caused in the infosphere (null law);
2. Entropy ought to be prevented in the infosphere;
3. Entropy ought to be removed from the infosphere; and
4. The flourishing of informational entities as well as the whole infosphere ought to be promoted by preserving, cultivating, enhancing and enriching their properties.

In light of this normative framework, let us go back to the open issues of informational warfare, so as to bridge the gap between the traditional framework of the laws of war and the set of parameters and conditions that should regulate cases of cyber war. First, canonical causes and conditions of just war can be reinterpreted in accordance with these four moral laws, that is as 'just exceptions' to them. For example, in light of the laws which prescribe that entropy shall not be caused, or it should be prevented in the informational environment, causes of just war as self-defence and reasonable success of the war-declaring authority *(jus ad bellum),* much as conditions of *jus in bello* like the doctrine of double effect, should be deemed exceptional cases in which the use of force is necessary. The reasons for this legitimate necessity are traditionally given as either the aim of removing entropy from the infosphere, for example by defeating another nation's unjust aggression, or preventing the creation of further entropy through the means of self-defence, a pre-emptive attack and so forth. The tenets of information ethics allow us to grasp the common ground between traditional targets of real-world operations and the new means of informational warfare, because all the entities that are at stake can properly be considered in terms of information.

Second, the ontological equality principle does not aim to equate, say, human resources with such informational tools as networks and computers. Rather, the informational outlook intends to provide a universal normative framework with which to govern the life cycle of information in an impartial manner. The lawfulness of virtual, as opposed to physical, force can thus be determined and compared with the legitimacy of the military goals that a nation aims to attain in the real world, by tracing them back to the first and second laws of information ethics. As previously stated, the causes legitimating war, much as the behaviour admitted in warfare, concern either the aim to prevent the creation of further entropy on the battlefield, or the goal of removing such entropy from the informational environment. Although these scenarios are closely related to the principle of proportionality, attention is drawn here to a different aspect of the problem, namely the proportion, and comparison, between real-world operations, tactics and ends of just war, and their informational counterparts in cyber warfare, as defined by the notion of entropy in the first and second laws of information ethics.

Third, the outlook of information ethics does not only offer the appropriate framework to examine the legal effects of strict informational attacks. This perspective allows us to deepen at what point a strict informational attack should be understood as an act of war (Schmidt and Cohen, 2013: 114). We should pay attention to the amount of entropy that is caused by such an attack aiming to remove or annihilate some informational resources in cyberspace. Then, in order to determine whether such strict a cyber attack has to be deemed as just, we have to evaluate the amount of entropy provoked by the attack with the first and second moral laws

mentioned above, namely the amount of entropy which was prevented by that cyber attack (first law), or the amount of entropy which was removed from the infosphere via the cyber attack (second law).

Fourth, the laws of information ethics help us to tackle the tricky scenarios of proportionality conceived as both a cause and a condition of just war. In the case of *jus ad bellum* and whether the good achieved by an informational attack is proportionate to the evil of waging it, we can determine that good in accordance with the first and second moral laws of information ethics so as to compare the military ends with the evil that is defined by the null law in terms of entropy. A similar ratio is at work in the case of *jus in bello* and whether the level of virtual, as opposed to physical, force is proportionate to attain a nation's military end: the evil of the null law, which is provoked by the cyber attack, should be grasped in accordance with the good that is illustrated by the first and second laws, the aim being to determine whether the level of that informational attack has to be deemed proportionate. Notwithstanding such a similar ratio, which is, after all, the reason why scholars usually refer to proportionality as both a cause and condition of just war, a crucial difference has to be stressed: whereas a proportionate cause to go to war may be ruined by a disproportionate use of violence, either virtual or real, a proportionate use of force cannot redeem a futile reason for fighting.

Admittedly, the devil is in the detail. From a legal point of view, this means that a number of issues are fated to remain open in this new context, including the exact moment at which a cyber attack becomes an act of war, and the grey zone affecting matters of discretion regarding how to interpret and apply the proportionality principle. Most of these uncertainties have affected the traditional laws of war as well: think of the good intention of the war-declaring authority and the reasonable success of war in the field of *jus ad bellum*, along with the doctrine of double effect *vis-à-vis* the principle of non-combatant immunity in the field of *jus in bello.* The realignment of these issues through the lens of information ethics does not mean that these issues are over, but rather that we can properly address them by taking into account both the ways in which these problems are reshaped by the means of the information revolution and how a normative framework, such as the moral laws of information ethics, can guide us throughout this huge transformation.

Still, we have to widen the spectrum of the analysis: so far, attention has been drawn to the content of current laws of war in terms of the causes and conditions that make wars just. However, the information revolution also impacts on the pillars of this traditional framework as a matter of procedure, rather than substance, in order to define the authority that may properly enter the informational wars. Is there any crucial difference between old and new scenarios?[98]

NEW SCENARIOS OF SOVEREIGNTY

The current legal framework of the laws of war can be grasped as a sort of compromise between a basic tenet of the Westphalian paradigm, the sovereignty of Nation States, and a post-Westphalian model of international law that restrains what Hobbes called the sovereign right of making war and peace with other nations and Commonwealths. Pursuant to the UN Charter and save in cases of self-defence, force can only be used if the UN Security Council authorises it, and yet States are deemed the only relevant actors in the field of international law and, more specifically, the only proper authority to declare and enter wars. Throughout this paper, we have noted the multiple ways in which the information revolution is affecting this traditional framework: the increasing dependence of societies on information as a vital resource challenges the aim of sovereign States to monopolise the use of force in cyberspace, the new domain of military operations. Authors

[98] By examining causes and conditions of Just War through the lens of the moral laws of information ethics, you may wonder what role the third law plays in this context, namely the aim to promote 'the flourishing of informational entities as well as the whole infosphere...by preserving, cultivating, enhancing and enriching their properties' (Floridi 2008). This moral law is very important for the laws of war, particularly in the field which scholars traditionally sum up as *jus post bellum. This* paper only
deals with the challenges of *jus ad bellum* and *jus in bello,* and so I have skipped this part of the analysis on *jus post bellum,* on which see chapter 7 of Schmidt and Cohen (2013).

of the new generation of informational attacks can be non-State actors, and identifying the party responsible for such attacks, whether non-State actors or traditional sovereign States, is often impossible. Whilst this very difficulty affects the principle that wars need to be declared by competent authorities, it also lowers the traditional barriers to enter into war, since anonymity may trigger new temptations for exploitative attacks and shelter the anonymous informational offender from the reaction of others.

The multiple ways in which the information revolution may thus impact on a crucial aspect of LOW and IHR, *such as* a procedural precondition for legitimating the war as the proper war-declaring authority, can be summed up with four different scenarios. The common ground is given by a Nation State that identifies or submits evidence as to the identity of the party responsible for a cyber attack. The first scenario can be dubbed as the constitutional one: the party responsible for a cyber attack is a non-State actor and in response to such aggression, a Nation State intends to prosecute pursuant to its own criminal laws and, eventually, in accordance with such international provisions as the rules of the Budapest Convention. The second scenario resembles the first with one crucial difference: here, in response to an informational attack, the State conceives the non-State actor as a combatant enemy under the current laws of war or, regardless of such rules, due to the evil nature of the aggression, for example derailing passenger trains by penetrating a nation's computer network. Whereas the national constitutional laws of the first scenario often provide for a stronger level of protection than terms and conditions of international laws of war, the second scenario can be summed up here with the George W. Bush doctrine of the 'war on terrorism.'

The third scenario is a Hobbesian one: an informational attack is carried out by a sovereign State and the targeted nation mulls over the counter-attack that may be appropriate in the absence of parameters and conditions which, in accordance with international law, govern such cases of informational aggression. A first option is given by the good will of the sovereign: the latter decides to constrain itself and abide by the 'precept, or general rule, found out by reason,' according to Hobbes's definition of the laws of nature in Chapter 14 of *Leviathan.* Either for moral reasons, or for simple matters of convenience under the pressure of the public opinion, the State may end up following, for example, the moral laws of information ethics illustrated in the previous section. Yet, in accordance with Hobbes's standpoint and the old Roman maxim ' salus populi suprema lex esto.' ' namely the health of the nation represents the supreme law, we should be prepared for the other way around; that is a Hobbesian state-of-nature between sovereign States. Here, it is up to the discretion and power of the State to determine forms and means of its response to a critical informational aggression.

The final scenario is closely related to the previous one, because it concerns the attempt to find a way out of the new international state-of-nature of the information era. After all, previous technological advancements have given rise to the drafting of international conventions and agreements to discipline and regulate the use of chemical, biological and nuclear weapons, land mines and the like. However, going back to the robotic warnings examined above, the field of military robotics technology offers a cautionary tale: in his 2010 Report to the UN General Assembly, the Special Rapporteur on extrajudicial executions, Christof Heyns, urged that Secretary-General Ban Ki-moon to convene a group of experts in order to address 'the fundamental question of whether lethal force should ever be permitted to be fully automated.'[99] Since both the UN General Assembly and its Secretary-General had been quiescent up to the date of this paper, it is but a short step to suspect that a new international agreement on some critical aspects of informational warfare may take a long time. This stalemate will continue as long as sovereigns think they can exploit the loopholes of the current legal framework due to their technological superiority or strategic advantage. Therefore, waiting for a new agreement in the long run, what should we expect next?

[99] See Christof Heyns, 2010 Report to the UN General Assembly, available at: http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G12/128/73/PDF/G1212873.pdf?OpenElement

All in all, the Hobbesian scenario aside, the lack of an international agreement does not mean that each man will be a threat to his neighbour in informational warfare. Going back to the difference between strict informational attacks and such attacks as a means for real-world operations, most of the latter can properly be addressed with the current provisions of the laws of war, either directly or by the use of analogy. By adapting Alston's remarks in the field of military robotics to the new scenarios of cyber warfare, the impact of an informational attack on real world targets such as power grids or water supplies, 'is no different from any other commonly used weapon.... The critical legal question is the same for each weapon: whether its specific use complies with IHL' (Alston, 2010). Even in the case of strict informational attacks, we have seen that some of these attacks seem to fall within the loopholes of current legal frameworks, and that some outlooks, such as the moral laws of information ethics, allow us to approach many of these cases conveniently.

Paradoxically, since we are dealing with a field which traditionally concerns problems of international law, that is the relationship between sovereigns, one of the main threats of informational warfare may pertain to the realm of national constitutional law. As Nation States are progressively unable to monopolise the use of force in cyberspace (Schmidt and Cohen, 2013), it is likely that non-State actors will have a crucial role in this new domain of military operations. This trend is not new, as we have seen from the privatisation of war and the role of 'corporate warriors' (Singer, 2008); and yet, the new frontiers of informational warfare will exacerbate it.

The more that non-State actors shape the new scenarios of cyber warfare, the more we should pay attention to the alternative between the first and second scenarios examined here. This trend is illustrated by a number of national programs concerning online security and defence, in that unconventional challenges of cyber attacks are increasingly testing the framework of legal safeguards that have represented, so far, the salient quality of Western democracies (Pagallo, 2013a). In light of the alternative between the constitutional scenario and the Bush doctrine of how to deal with a new generation of informational attacks, declaring 'war on cyber-terrorists,' what is at stake here regards some pillars of national law, such as the protection of basic individual rights.

At times, such programs for online security and defence look satisfactory: consider the new *Police and Criminal Justice Data Protection Directive* that the European Commission presented in January 2012. In the Seventh Considerandum of the proposed directive, for example, the EU Commission significantly refers to 'the level of protection of the rights and freedoms of individuals for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties'[100] that 'must be equivalent in all Member States.' Moreover, in the words of the Commission, 'the protection of the rights and freedoms of data subjects requires that appropriate technical and organizational measures be taken to ensure that the requirements of the Directive are met' *(op. cit.,* n. 38). Further examples illustrate that a Hobbesian approach to issues of national security and defence is still popular among Western scholars and more importantly, a number of national programs on online security confirm that, rather than a Hobbesian state-of-nature in the international affairs of informational warfare, the main threat may be to the fields of national and constitutional law. Some believe that providing basic security must be the first priority in policy considerations, at least in international affairs, because security drives democracy, and not the other way around (Etzioni, 2007). However, the 2013 scandal of the U.S. National Security Agency (NSA)'s Prism project, much as the UK's GCHQ files,[101] have shown that the threat to constitutional rights does materialise in the field of national law.

[100] See *European Commission's proposal is COM(2012) 10 final 2012/0010 (COD).*

[101] See John Lanchester, *'The Snowden files: why the British public should be worried about GCHQ', The Guardian*, 3 October 2013, at http://www.theguardian.com/world/2013/oct/03/edward-snowden-files-john-lanchester (last accessed 5 October 2013).

CONCLUSIONS

A classic topic of legal philosophy concerns the ways in which we should address the hard cases of the law, namely the class of legal issues where the disagreement among scholars revolves around the meaning of the terms that frame the legal question, the ways such terms are related to each other in legal reasoning or the role of the principles that are at stake in the case. This class of legal issues is recurrently provoked by advances in technology, because scholars often disagree about whether technological innovation may affect concepts and principles of legal reasoning, or whether it creates new concepts and principles, or does not challenge them at all, the latter being the view of traditional legal scholars. Some affirm 'there is no possibility of treating the question raised by the various cases as if there were one uniquely correct answer to be found, as distinct from an answer which is a reasonable compromise between many conflicting interests' (Hart, 1961: 128). Others, such as Ronald Dworkin and followers of the 'right answer' thesis, interpret the law in a morally coherent way so that, given the nature of the legal question and the history and background of the issue, lawyers could obtain the solution that best justifies or fits the integrity of the law.

By examining the new scenarios and challenges of informational warfare in connection with the tenets of the Westphalian paradigm and today's laws of war, we should admit that no one-size-fits-all answer can properly tackle the complexity of the subject matter. We have seen a number of cases in which a Dworkinian approach is fruitful: a morally coherent theory, such as Floridi's ethics of information, allows us to reinterpret traditional causes and conditions of just-war theory so as to fill some gaps of the current legal framework. In particular, the aim was to:

(i) compare canonical targets of real-world operations on the battlefield with the new means of informational warfare, for all the entities can properly be considered in terms of information;
(ii) examine at what point a strict cyber attack should be understood as an act of war, in light of the amount of entropy that is caused by the aim to remove, annihilate or delete such informational objects as the enemy's communication networks or computers; and,
(iii) clarify the scenarios of proportionality as a cause and condition of just war, by relating the amount of entropy provoked by an informational attack to the amount of entropy which is either prevented or removed by such an attack in the infosphere.

We considered a further set of issues that is however fated to remain open in the legal field: think about Chapter VII of the Charter of the United Nations on 'action with respect to threats to the peace, breaches of the peace, and acts of aggression,' such as strict informational attacks. Here, the devil again is in the detail: a 'reasonable compromise' on the basis of legal expertise *(*Hart, 1961), more than a right answer (Dworkin, 1985), seems necessary to redefine the locus and strict conditions of political and military responsibility. The field in which such agreements are traditionally reached and their absence mostly perceived is the field of international law. Contrary to previous conventions on the use of various weapons, a twofold peculiarity of informational warfare has nonetheless to be stressed once again: on one side, the lack of an international agreement does not entail a new Hobbesian state-of- nature of the information era. Although some of today's laws of war shall be reformulated so as to include specific provisions for strict informational attacks and their counterpart in real-world operations, it does not follow a condition in which 'all is permitted' among sovereign States. Rather, we should wonder about how the international community may react to a devastating series of strict informational attacks that cannot be ascribed to a responsible party.

The increasing incapacity of Nation States to monopolise the use of force in cyberspace which goes hand in hand with the difficulty and, at times, the impossibility of identifying the responsible parties, suggests a national, rather than international, legal threat. Whereas non-State actors will incrementally play a crucial role in cyberspace as the new domain of military operations, several programs on online security and national defence have been developed by sovereign States to tackle the menace of a new generation of cyber attacks carried out by other sovereign States or non-State actors. The endurance of Western democracies and their

aim to protect basic individual rights, such as privacy, freedom of speech and data protection, have already been tested by such national programs over the past year. Rather than a new Hobbesian state-of-nature in the international affairs of informational warfare, a Hobbesian approach to matters of security and defence may indeed be the main threat in the fields of national and constitutional law. The new scenarios of informational warfare do not only concern the field of international law, much as LOW and IHL, after all.

REFERENCES

Alston, Philip (2010) *Report of the Special Rapporteur on Extrajudicial, Summary and Arbitrary Executions,* UN General Assembly, Human Rights Council, A/HRC/14/24/Add.6, 28 May

Asaro, Peter (2008) How Just Could a Robot War Be?, *Frontiers in Artificial Intelligence and Applications,* 75: 50-64

Canning, John S. (2008) 'Weaponized Unmanned Systems: a Transformational Warfighting Opportunity, Government Roles in Making It Happens' In: American Society of Naval Engineers' (ASNE) Proceedings of Engineering the Total Ship (ETS) Symposium, Falls Church, VA.

Dworkin, Ronald (1985) *A Matter of Principle.* Oxford University Press, Oxford

Etzioni, Amitai (2007) *Security First: For a Muscular, Moral Foreign Policy.* Yale University Press, New Haven

Floridi, Luciano (2008) 'Information Ethics, its Nature and Scope', in van den Hoven J. and Weckert J. *(eds.)* Moral Philosophy and Information Technology, 40-65, Cambridge University Press, Cambridge

Floridi, Luciano (2013) The Ethics of Information: Volume II of Principia Philosophiae Informationis. Oxford University Press

Floridi, Luciano (forthcoming) *The Fourth Revolution — The Impact of Information* and Communication Technologies on Our Lives. *Oxford University Press*

Hart, Herbert L. A. (1961) The Concept of Law. Clarendon, Oxford (2nd edition 1994)

Hobbes, Thomas (1999) *Leviathan,* R. Tuck (ed.). Cambridge University Press, Cambridge

Lin, P. Bekey, G. and Abney, K. (2007), *Autonomous military robotics: risk, ethics, and design.* Report for US Department of Navy, Office of Naval Research. Ethics + Emerging Sciences Group at California Polytechnic State University, San Luis Obispo, CA.

Pagallo, U (2011c) Robots of Just War: A Legal Perspective, *Philosophy and Technology, 24(3): 307-323*

Pagallo, U (2013a) Online Security and the Protection of Civil Rights: A Legal Overview, *Philosophy and Technology,* forthcoming

Pagallo, U (2013b) The Laws of Robots: Crimes, Contracts, and Torts. *Springer, Dordrecht*

Pagallo, U and Massimo D (2009) Three Roads to P2P systems and their Impact on Business Ethics, *Journal of Business Ethics,* 90(4): 551-564

Pinker, S (2012) *The Better Angels of Our Nature.* Penguin, New York

Schmidt, E and Cohen J (2013) *The New Digital Age.* John Murray, London

Sharkey, N (2008) Grounds for Discrimination: Autonomous Robot Weapons*, RUSI Defence Systems*, *11(2): 86-89*

Singer, *PW*. (2008) Corporate Warriors: The Rise of the Privatized Military *Industry*. Cornell University Press, Ithaca and London

Sparrow, Robert (2007) Killer Robots, *Journal of Applied Philosophy,* 24(1): 6277

Walzer, M (1977) *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. Basic Books, New York

Biographies

Dr Edward T. Barrett is the Director of Research at the United States Naval Academy's Stockdale Center for Ethical Leadership. A graduate of the University of Notre Dame, he completed a Ph.D. in political theory at the University of Chicago, and is the author of *Persons and Liberal Democracy: The Ethical and Political Thought of Karol Wojtyla/John Paul II* (2010). While in graduate school, he served for two years as speechwriter to the Catholic Archbishop of Chicago. He joined the United States Air Force reserves after serving nine years as an active duty C-130 instructor pilot, and is currently a Colonel in the Air Staff's Directorate of Strategic Planning.

Dr Bill Boothby retired in July 2011 as Deputy Director of Legal Services (RAF) in the rank of Air Commodore (1 star). During a 30 year career in the RAF Legal Branch, he serviced in UK, Germany, Hong Kong, Cyprus and Croatia. In 2009 he took a Doctorate in International Law at the Europa Universität Viadrina, Frankfurt (Oder) in Germany. In the same year he published Weapons and the Law of Armed Conflict through OUP. He is currently working on The Law of Targeting to be published in 2012 by the same publisher. He was a member of the Group of Experts convened by the ICRC to discuss Direct Participation in Hostilities, was a member of the Group of Experts who produced the HCPR Manual of the Law of Air and Missile Warfare and is on the drafting committee of the CCD COE project to produce the Tallinn Manual of the Law of Cyber Warfare. He lectures and speaks widely on international law issues. He is currently Associate Fellow at the "Geneva Centre for Security Policy" in Switzerland.

Professor Sandra Braman's research on the macro-level effects of the use of digital technologies and their policy implications has been supported by the Ford Foundation, Rockefeller Foundation, Open Society Institute, and the US National Science Foundation. Her books include Change of State: Information, Policy, and Power (MIT Press, 2006/2009; second edition in progress), and the edited volumes The Emergent Global Information Policy Regime (Palgrave Macmillan, 2004), Biotechnology and Communication: The Meta-technologies of Information (Lawrence Erlbaum, 2004), and Communication Researchers and Policy-Making (MIT Press, 2003). She is the current Chair of the Law Section, International Association of Media and Communication Research in University of Wisconsin-Milwaukee, and former Chair of the Communication Law & Policy Division, International Communication Association.

Professor Selmer Bringsjord specializes in the logico-mathematical and philosophical foundations of artificial intelligence (AI) and cognitive science, and in collaboratively building AI systems on the basis of computational logic. Bringsjord received the bachelor's degree from the University of Pennsylvania, and the PhD from Brown University, where he studied under Roderick Chisholm. Bringsjord has long been on faculty at America's oldest technological university: Rensselaer Polytechnic Institute (RPI) in Troy; where he currently holds appointments in the Department of Cognitive Science, the Department of Computer Science, and the Lally School of Management & Technology, and where as a Full Professor he teaches AI, formal logic, human and machine reasoning, philosophy of AI, other topics relating to formal logic, and the intellectual history of New York City and the Hudson Valley.

Professor Randall R. Dipert has taught at the (SUNY) University at Buffalo since 2000, where he is the C.S. Peirce Professor of American Philosophy. Previously he taught at SUNY Fredonia (1977-1995) and the U.S. Military Academy at West Point (1995-2000). Prof Dipert's most active current interests are military ethics (the philosophy of war and peace) and applied ontology: he is a founding member of the National Center for Ontological Research (NCOR). his past research includes logic, the history and philosophy of logic, Peirce and Early American Pragmatism, the philosophies of logic and mathematics, the philosophy of artifacts, aesthetics, action theory, and metaphysics, especially the metaphysics and logic of relations. He also has interests in ethics and political philosophy and most recently, in the philosophy of war and peace (Just War theory) and value-theoretic implications of game theory, especially of the Iterated Prisoner's Dilemma. Recent publications and papers include the ethics of cyberwarfare and the ethics of preemptive/preventive war.

Dr Massimo Durante is Researcher in Philosophy of Law at the Department of Law of the University of Turin, and holds a Ph.D. in History of Philosophy at the Faculty of Philosophy of the University of Paris IV Sorbonne. His main fields of research concern Philosophy of Law, Legal Informatics and Information Ethics. Author of several books, he has widely published articles in Italian, English and French. He has recently edited, with Ugo Pagallo, the book: *Manuale di informatica giuridica e diritto delle nuove tecnologie*, Utet, Torino, 2012.

Lieutenant Colonel Maurizio D'Urso was born in Rome on 18th February 1963. Graduated in law *cum laude* at the University of Rome "La Sapienza", has become a lawyer. Attracted by the military career, he joined the Administrative and Legal Corp (Commissariat Corp) of Italian Air Force in 1994. He directed the Administrative Branch of Military Airport of Guidonia (near Rome) from 2001 to 2004 and since then he has worked as legal advisor for national and international legal affairs. Currently he has the rank of Lieutenant Colonel and is applied to the Italian Defence General Staff – Legal Affairs General Office as head of the European Union legal affairs Sector. He has been involved in Cyber Defence legal consulting since 2011. He speaks English, German and French.

Professor Luciano Floridi is Professor of Philosophy and Ethics of Information at the University of Oxford, Senior Research Fellow at the Oxford Internet Institute, and Fellow of St Cross College, Oxford. Among his recognitions, he was the UNESCO Chair in Information and Computer Ethics, Gauss Professor of the Academy of Sciences in Göttingen, and is recipient of the APA's Barwise Prize, the IACAP's Covey Award, and the INSEIT's Weizenbaum Award. He is an AISB and BCS Fellow, and Editor in Chief of Philosophy & Technology and of the Synthese Library. He was Chairman of EU Commission's "Onlife Initiative". Floridi's research concerns primarily the Philosophy of Information, Information and Computer Ethics, and the Philosophy of Technolgy. Other research interests include Epistemology, Philosophy of Logic, and the History and Philosophy of Scepticism. He has published over a 150 papers in these areas, in many anthologies and peer-reviewed journals. His works have been translated into Arabic, Chinese, French, Greek, Japanese, Italian, Hungarian, Persian, Polish, Portuguese, Russian, and Spanish.His most recent books are: The Ethics of Information (OUP, 2013), The Philosophy of Information (OUP, 2011), The Cambridge Handbook of Information and Computer Ethics (editor, CUP, 2010), and Information: A Very Short Introduction (OUP, 2010).

Lieutenant Ludovica Glorioso is an Italian Army Officer, working as a scientist at the NATO CCD COE Law & Policy branch. Since she joined the Army, Lt Glorioso served as a legal adviser in NATO Peacekeeping Operations in the Balkans and Afghanistan. She was assigned at the Italian Joint Operation HQ and at Army General Staff HQ. Lt Glorioso holds a Law Degree from University of Palermo, an LL.M. in European and Transnational Law from Trento University (Italy) and she is admitted to the Italian bar. Lecturer at the University RomaTRE, the Italian Center for High Defence Studies (CASD), Cyber Defence Symposium in Italy and World Summit on Counter-terrorism at the Interdisciplinary Center Herzliya (Israel).

Anna-Maria Osula works as a scientist at the NATO CCD COE Law & Policy branch. She holds an LLM degree in Information Technology Law from Stockholm University and is working towards a law PhD at Tartu University, Estonia. In 2011-2012 she was researching international legal cooperation on cyber security at GSAPS, Waseda University, Tokyo. Her areas of research include national cyber security strategies, international organisations and cyber crime. She is also giving lectures on "Legal Aspects of Cyber Security" at the Tallinn Technical University, and has presented at various events of NATO COE DAT, NATO School, OSCE, AusCERT and others.

Professor Ugo Pagallo is Professor of Jurisprudence at the Department of Law, University of Turin, since 2000, faculty at the Center for Transnational Legal Studies (CTLS) in London and faculty fellow at the NEXA Center for Internet and Society at the Politecnico of Turin. Member of the European RPAS Steering Group (2011-2012), and the Group of Experts for the Onlife Initiative set up by the European Commission (2011-2013), he is chief editor of the *Digitalica* series published by Giappichelli in Turin and co-editor of the AICOL series by Springer. Author of nine monographs and numerous essays in scholarly journals, his main interests are AI & law, network and legal theory, robotics, and information technology law (specially data protection law, copyright, and online

security). He is currently a member of the Ethical Committee of the CAPER project, supported by the European Commission through the Seventh Framework Programme for Research and Technological Development.

Professor Giovanni Sartor is part-time full professor in legal informatics at the University of Bologna and part-time professor in Legal informatics and Legal Theory at the European University Institute of Florence. He obtained a Ph.D. at the European University Institute (Florence), worked at the Court of Justice of the European Union (Luxembourg), was a researcher at the Italian National Council of Research (ITTIG, Florence), held the chair in Jurisprudence at Queen's University of Belfast (where he now is honorary professor), and was Marie-Curie professor at the European University of Florence. He has been President of the International Association for Artificial Intelligence and Law. He has published widely in legal philosophy, computational logic, legislation technique, and computer law.

Dr Mariarosaria Taddeo is Research Fellow in Cyber Security and Ethics, at the Politics and International Studies Department (PAIS), University of Warwick, and Research Associate at the Uehiro Centre for Practical Ethics, University of Oxford. Her work focuses mainly on the ethical analysis of cyber security practices and information conflicts. Her primary research interests are Information and Computer Ethics, Philosophy of Artificial Intelligence and Multi-Agents Systems. Until 2012, Dr. Taddeo held a Marie Curie Fellowship to work on the ethics of information warfare. She is the editor (with Luciano Floridi) of *Ethics of Information Warfare*, a volume forthcoming for Springer.