Conference on the Laws of War for the Cyber Era Background Guide

Head Chair: Parth Nobel

Contents

1 Letter from the Chair

Dear Delegates,

My name is Parth Nobel and I am honored to serve as Head Chair for *The Conference on the Laws of War for the Cyber Era* at UCBMUN XXIII. I am a sophomore at UC Berkeley studying Electrical Engineering and Computer Science. My research is centered on improving the mathematical models of circuits by designing new algorithms to generate such models. I serve as the Director of Technology for this UCBMUN, and have previously been an Editor for the Berkeley Student Journal of Asian Studies, and an intern at HP Inc. I joined UCBMUN as a freshman and quickly found a new family of friends and an unparalleled community. After staffing UCBMUN XXII as an ACD and Spring 2018 BayMUN as a Head Chair, I was motivated to run this committee in the hopes of facilitating and hearing debates by accomplished delegates on a topic that combines my two areas of interest, the disconnect between contemporary public international law and the modern cyber capabilities of war.

We live in a rapidly changing world. Google Duplex will call your Hair Salon for you. Waymo's driverless cars can drive you from SFO to the Hilton in the Financial District. Tensorflow allows me to teach a computer how to count the number of gavels in a picture of a UCBMUN delegation. And Google, Waymo, and Tensorflow are all part of, or made by, Alphabet, one private company. Our world is digital. Hospitals have been held for ransom by hackers. A single computer virus managed to create explosions in a classified nuclear facility deep underneath Iran. Russia has dismembered the Ukrainian power grid days before Christmas in retaliation for the Ukrainian people electing an anti-Russian government. The Chinese government has copies of the personnel files of nearly every single American with security clearance. Any of these facts could be construed as an act of war, and as justification for a firm, violent, and physical response. In contrast to technology, international law is slow to change. It can take decades, even centuries, for a new international custom to form. World-shaping treaties are rare; one could reasonably argue the last one signed was in this very city of San Francisco, in 1945. This conference aims to speed up the latter process. Delegates are expected to draft a treaty updating the laws governing the use of force and the laws of war for a digital world.

But be warned, good law should transcend the facts of the technology. A treaty which defines Stuxnet to be an illegal act, but fails to provide a framework by which to judge Flame is worthless. A treaty which will be outdated or whose intent will be broken by

further migrations of data into the cloud, as the US's Stored Communications Act was by the invention of webmail, could either be useless or dangerous, respectively. Delegates should draft tests of legality based off impact and intent rather than presenting tests based off technological details and nuances that could easily become outdated. I reserve the right to deny working papers that I feel will not generalize as technology changes. Also ensure that judges with minimal technical training, at the International Court of Justice or elsewhere, will have no difficulty applying the standards you propose.

Finally, I must express my utmost appreciation to UCBMUN XXIII's Secretary General Ruhee Wadhwania for being a perpetual source of inspiration, my Under Secretary-General Anya Mansoor, Dr. Karen Seif for introducing me to international law, and Rohan Shah and Nick Romanoff's European Conference on International Organization at ChoMUN XXI for inspiring me as I put together this committee.

Sincerely,

Parth T. Nobel
Head Chair
Director of Technology
cyberwar@ucbmun.org

2 Committee Overview

2.1 Procedure

The parliamentary procedure of this committee will be mostly identical to the procedure described in the Rules & Procedure of UCBMUN XXIII. There are a few areas where this Conference will diverge from those rules.

2.1.1 Passing Multiple Treaties

The Dais will not allow contradictory treaties to be passed. The Dais also expects draft treaties to be comprehensive, addressing all of the issues discussed in this document. This has led to the Dais anticipating only one treaty being passed at the end of the Conference.

2.1.2 Setting the Agenda

This Conference has only one topic and was created solely to address that issue. There will, therefore, be no debate on setting the agenda of the committee.

2.2 Position Papers

All delegates are expected to submit a 2 to 4 page double-spaced paper that presents their positions on the major questions of debate, and which should include at least one detailed solution to a problem. Delegates may include up to a half page of content on research they have conducted beyond this document. Failing to submit a position paper before **11:59PM Pacific Time on February 20th, 2019** will result in a significant penalty during awards consideration.

2.3 Technology in Committee

In an effort to reduce the ecological impact of the Conference, we will not be printing copies of committee materials. Delegates will be allowed laptops and tablets during committee beginning on Friday's second session. **Use of electronics is strictly limited to reading and marking up committee materials**. The Dias will have a small collection of USB sticks containing the committee materials to pass around. Delegates are encouraged to **bring a USB stick along with a laptop to committee beginning on Friday**.

2.4 Scoring

Scoring will be determined off the prevalence and quality of novel and substantive speeches during debate, leadership in blocs, and, especially, on unique ideas presented in a position paper reaching the final stages of debate.

2.5 Law v. Technology

This committee is first and foremost about creating law. This conference is not here to regulate technologies. Debates on blockchains or even copyright have a questionable place in this conference.[1] Technology moves faster than law. Delegates should be writing law that can survive next year's iPhone and Pixel release, the invention of IPv8, and any other new technological invention that doesn't fundamentally transform the world on the scale the internet and world-wide web did. Knowledge of technology is critical and helpful for being able to explore these questions, because it provides a sense of clarity on what is and is not possible, but an over reliance on technological expertise can be constraining.

2.6 External Research

Delegates are encouraged to research public international law beyond what is described below. This document focuses on customary public international law, leaving most treaties that are critically relevant to this topic unexplored. Delegates are encouraged to also do as much research as they require to come to the realization that every networked system is vulnerable.

3 Contemporary Public International Law

Limitations on war are as old as modern civilization. In the Mahabharata, a Hindu sacred text dating to at least 400 BCE, laws on proper combat and battle are agreed to by both sides of the Kurukshetra War. In Western traditions, the Torah, in Deuteronomy, defines limitations on who may serve in an army and how sieges may be conducted. St. Augustine authored a theory of when war was acceptable and proposed some limitations on actions taken during war.

[1] Don't interpret this as a ban on discussing those two words, just a warning that specific technologies and law regulating private parties are not relevant to the questions we seek to answer.

It is from this ancient history that contemporary public international law has formed, often based off of the same principles as those expressed by the ancient sources. This conference will focus on two issues: *jus ad bellum*, the determination of when engaging in war is legitimate, and *jus in bello*, the understanding of what actions are appropriate in a war.

3.1 Sources of Public International Law

The two most pertinent sources relating to the study of the laws of war are treaties and customary international law. While these sources are equally normative, making a doctrine normative via treaties is straightforward—each State that signs the treaty is bound thereby.[2] Making a doctrine normative via customary international law is significantly more complicated. There are two components to establishing a usage as customary international law: showing both that the usage is a general practice among States and that the usage has *opinio juris*.[3] General practice does not require a long history of practice, but does require that a vast number of States follow the practice strictly.[4] Proving *opinio juris*, or the belief that the usage is mandated by law, is a significant challenge. To prove *opinio juris*, it has to be shown that some States are complying with the usage and are not obligated by treaty to obey the usage.[5] Further, it is also critical that either all the relevant States repeatedly refer to the usage as being mandated by law[6] or that legal scholars describe it as mandated by law.[7]

Much of the relevant law on force and war is drawn from a mix of both treaties and customs. Customs are slow to form. In order to speed up the process of creating workable norms around cyberwar, it is critical that this conference author a workable treaty to address the questions that are associated with this new technology.

[2]Crawford, James. 2012. *Brownlie's Prinicples of Public International Law*. 8th ed. Oxford University Press. p. 22.

[3]*North Sea Continental Shelf. Judgement*. 1969. *ICJ Reports*. paras. 72, 74–75.

[4]*North Sea Continental Shelf. Judgement*. 1969. *ICJ Reports*. para. 75.

[5]*North Sea Continental Shelf. Judgement*. 1969. *ICJ Reports*. para. 76.

[6]Case Concerning Military and Paramilitary Activities in and Against Nicaragua. 1986. ICJ Reports. para. 189.

[7]"The Paquete Habana." 1900. In *International Law — Cases and Materials*, by Damrosch et. al., 4th ed. p. 66.

3.2 *Jus ad Bellum*

Sovereign States not only have incredible power over their domestic affairs and subjects, but possess established rights in the international sphere. To maintain these rights and powers, States must be shielded from some external influences, motivating the Principle of Non-Intervention. Article 51 of the UN Charter, however, provides a critical exception to the Principle of Non-Intervention—it grants Member States the right to respond to an armed attack with either collective or individual self-defense. Under the UN Charter, there are two legal uses of force: that taken in self-defense and that taken with the express authorization of the UN Security Council.

3.2.1 The Principle of Non-Intervention

In the 1986 *Case Concerning Military and Paramilitary Activities In and Against Nicaragua*, the International Court of Justice held that customary international law prohibits, in general, States from threatening and using force.[8] The ICJ held that since interventions undermine global stability, the very purpose of the UN Charter, they must be prohibited by customary international law.[9] The ICJ then found that non-intervention had been expounded upon by the UN General Assembly, and numerous other bodies, showing *opinio juris* for the Principle of Non-Intervention.[10] The Court summarized the Principle as forbidding "all States or groups of States to intervene directly or indirectly in internal or external affairs of other States … on matters in which each State is permitted … to decide freely."[11] It further concluded that states cannot intervene in support of a rebel group in another state with force.[12] International law prohibits intervention, and in general prohibits the use of force. However, there are exceptions to this prohibition. One such exception allows States to defend themselves from existential threats: the self-defense exception to the Principle of Non-Intervention. This self-defense exception takes two forms, individual and collective.

[8] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 191.

[9] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 202.

[10] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 203.

[11] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 205.

[12] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 209.

3.2.2 Self-Defense: A Legal Response to Armed Attacks

The standard for when a State has a right to individual self-defense is established in *Nicaragua*. In its opinion in *Nicaragua*, the ICJ drew a distinction between a grave use of force, such as an armed attack, and "less grave forms."[13] The ICJ then established that a State must be the victim of an armed attack, i.e. a grave use of force, in order to have a right to individual self-defense. The Court held that an armed attack can be executed by either regular uniformed soldiers or irregulars and mercenaries under the control of the attacking state. The Court explicitly excluded assisting rebels by providing weapons or logistics assistance, instead saying that such actions were a less grave use of force. The ICJ also placed a burden on the victim of an armed attack to both assert that they were the victim of an armed attack and to prove that such an attack occurred.[14]

Collective self-defense, in contrast to individual self-defense, allows a state to invoke self-defense on behalf of a different state. The ICJ, in *Nicaragua*, required that to claim collective self-defense the State who was the victim of an armed attack must request assistance from other States before the other States can intervene on the victim State's behalf.[15] Combined with the general requirements to claim self-defense, in order to allow other States to use force on its behalf, a State must prove it was the victim of an armed attack, assert such, and ask for assistance.

3.2.3 Self-Defense: Necessity and Proportionality

Being a victim of an armed attack, however, does not give a State a blank check to respond with unlimited force. There are limits on the force authorized by the self-defense exception. The details of these limits were established in the case of the *Caroline*, a dispute between Great Britain and the United States in the 1830s. The *Caroline* was a US merchant vessel that, acting independently of the US government, was transporting supplies to Canadian rebels from private American supporters. Britain, attempting to cut supply lines to the rebels, pursued the boat into American territory at night, lit the boat aflame, and set it

[13] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 191.

[14] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 195.

[15] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 199.

adrift down the river.[16] As part of a series of protests, Secretary of State Daniel Webster sent a communique to the British government espousing a standard for the limits on force used by a State under the self-defense exception to the Principle of Non-Intervention. Webster contended that actions taken in self-defense must be necessary and proportional to the attack justifying self-defense.[17] Great Britain accepted Webster's construction of self-defense, and argued that it had complied with his standard. Britain claimed the attack was necessary since the *Caroline* was rapidly increasing the number of arms held by Canadian rebels and that it was proportionate, since attacking at night minimized unneeded loss of life.[18] While Webster rejected the British argument, alleging that the attack was unnecessary and disproportionate to simply transporting supplies, he accepted the British acceptance of his principles as the final statement on the dispute.[19]

This standard has grown into a fundamental component of the self-defense exception. In *Case Concerning Oil Platforms*, the US alleged that Iran had launched a number of assaults on American ships traveling in the Persian Gulf, and responded by destroying multiple Iranian oil platforms.[20] The ICJ held that the US must "show that its actions were necessary and proportional" to Iran's action,[21] showing that the standard espoused by Daniel Webster in the *Caroline* is still customary international law. Nonetheless, the ICJ rejected the US's claim it was the victim of an armed attack, by finding none of the evidence of the alleged attack compelling, and held that the US's actions could not be justified.[22] Regardless, the Court proceeded in its analysis and rejected the US's assertion that the attack and destruction of two permanent oil platforms was needed to defend American security interests, since the Court found that there was questionable evidence that the platforms served any military purposes.[23] The Court also questioned whether it was possible for the attack, had it been justified or necessary, to have been proportional, noting that one of the attacks was part of a significantly larger military operation launched

[16]Noyes, John E. 2007. "The Caroline: International Law Limits on Resort to Force." In *International Law Stories*, by Noyes *et. al.*, 1st ed. p. 268.
[17]Noyes, John E. 2007. "The Caroline: International Law Limits on Resort to Force." In International Law Stories, by Noyes *et. al.*, 1st ed. p. 304–5
[18]Noyes, John E. 2007. "The Caroline: International Law Limits on Resort to Force." In *International Law Stories*, by Noyes *et. al.*, 1st ed. p. 291.
[19]Noyes, John E. 2007. "The Caroline: International Law Limits on Resort to Force." In *International Law Stories*, by Noyes *et. al.*, 1st ed. p. 292.
[20]*Case Concerning Oil Platforms*. 2003. ICJ Reports. para. 50.
[21]*Case Concerning Oil Platforms*. 2003. ICJ Reports. para. 51.
[22]*Case Concerning Oil Platforms*. 2003. ICJ Reports. para. 64.
[23]*Case Concerning Oil Platforms*. 2003. ICJ Reports. para. 76.

by the US allegedly in response to the mining of a single ship that was not sunk nor suffered any loss of a life.[24]

The Non-Intervention Principle is fundamental to State sovereignty; it ensures that States have the right to conduct their business without undue external influence. Yet, States must have the right to protect their existence when they become victims of an attack. Any action taken in self-defense must be taken in response to an armed attack by another state, and be both necessary for the defense of the State and proportional to the armed attack justifying the self-defense. In *Nicaragua*, it was clear that the US supplying rebels was not an armed attack, but was a less grave use of force. In *Caroline*, questions abound about whether Great Britain's actions were justified; however, both sides agreed that acts taken in self-defense must be necessary and proportional. In *Oil Platforms*, it was clear that missiles and underwater mines striking American warships are an armed attack (even if who made the attack is unclear), but that the American response to the attack could not be justified. But another question remains about *Nicaragua*: does the US, a major supplier of the Contras, bear criminal responsibility for the human rights violations committed by the Contras?

3.2.4 Identifying Groups Controlled by Foreign States

If a State commits a human rights violation in another State's territory, the offending State is criminally liable for its acts. *Nicaragua* and the criminal trial of Tadic each define a standard for a State's criminal liability in circumstances where it had some degree of control over a third party committing human rights violations in a separate State.

The United States interfered in Nicaragua by funding the rebel group known as the Contras, who allegedly committed some human rights violations.[25] Nicaragua contended that the US was criminally liable for the Contras' human rights violations, since the Contras were so dependent on the United States that they constituted an organ of the United States government.[26] The ICJ, in Nicaragua, declared that mercenaries "recruited, organized, paid and commanded" by a State qualify as organs of the State, but that the Contras were not such mercenaries.[27] This critically clarifies that mercenaries directly

[24] *Case Concerning Oil Platforms*. 2003. ICJ Reports. para. 77.

[25] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 15.

[26] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 109.

[27] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para.

and unambiguously controlled by a State qualify as organs of the controlling State. By not finding the Contras to be such mercenaries, the Court set a standard that selecting and controlling the leadership of an organization does not constitute effective control and, thus make the State responsible for the organization's human rights violations.[28] The Court concluded that for the United States to be liable for the Contras actions, it must have directed or enforced the perpetration of the human rights violations.[29]

The *Nicaragua* case sets a very high standard for a State to be liable for another group's actions, requiring effective control of the group. The height of this bar motivated the International Tribunal for the Persecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia Since 1991 (ICTY) to unambiguously reject the ICJ's *Nicaragua* test as flawed in *Tadic*. The ICTY argued that *Nicaragua*'s effective control test makes it too easy for States to escape international responsibility by having humanitarian crimes be committed by private groups rather than the State's organs.[30]

The ICTY proposed a more complicated test, one which treats individuals or unorganized groups of individuals differently from organized groups, such as a military or paramilitary group. The ICTY applies the effective control test defined in *Nicaragua* only on individuals.[31] The ICTY held that if a State had either given explicit instructions to the individual or expressed an *ex post facto* appreciation for the individuals actions, the State was liable for illegal acts committed by that individual.[32] Yet, a State is still responsible for an individual violating international law while carrying out a lawful activity on behalf of the State.[33]

114.

[28] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 112.

[29] *Case Concerning Military and Paramilitary Activities in and Against Nicaragua*. 1986. ICJ Reports. para. 115.

[30] Prosecutor v. Tadic. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 117.

[31] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 124.

[32] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 118.

[33] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 119.

The ICTY's test for organized groups is significantly weaker than *Nicaragua*'s effective control test, instead merely requiring overall control.[34] The weaker test is offered in part because organized groups impose inherent structure on the actions taken by their members, and that any stronger of a test allows States to escape international obligations by off-loading illegal work to partially controlled organizations.[35] The ICTY defends this standard by pointing to broad practice.[36] The ICTY also clarifies that overall control is more than just financial and military aid, but includes coordinating and planning with the group regarding military activity.[37]

3.2.5 Actions of the Security Council

Under Article 42 of the UN Charter, the Security Council is authorized to order the use of force if non-violent means of coercing a state into maintaining peace fails.[38] This authorization is clearly a much simpler means of justifying the use of force, as there is no complicated legal definition of an armed attack, no distinction between grave and non-grave uses of force, and no requirement for the victim state to definitively prove who was liable.

The challenge with this strategy is, of course, the P5. Asserting self-defense to justify force does not require getting many historically opposed factions of geopolitics to agree and, hence, is the most often used to justify force. There is no clear requirement to rewrite Article 42, or even reconsider it. However, a final treaty that does include a recommendation to the Security Council, or that proposes a standard of when cyber-activities taken by a state obligate the Security Council to act, could be justified.

[34] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 120.

[35] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. paras. 120–21.

[36] *Prosecutor v. Tadic*. 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. paras. 125–29.

[37] *Prosecutor v. Tadic.* 1999. International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991. para. 131.

[38] *Charter of the United Nations*. 1945. Article 42.

3.3 *Jus in Bello*

Jus in bello translates from Latin to the laws in war. Some view this as a contradiction. Isn't war the breakdown of rules? But the laws of war are as normative as the rest of international law: they lack a clear enforcement strategy, but States obey them so that other States are obligated to obey. This section will not be as extensive as the previous one, as the Dais considers *jus in bello* more straight-forward and views portions involving weapons of mass destruction not directly related to the committee's topic, and hence omits discussion thereof. This lesser emphasis should not be interpreted as an indication of lesser importance.

When do the Laws of War apply? Some portions of the Laws of War are customary international law, others are introduced as treaties, some treaties have become part of customary international law—notably the humanitarian protections in World War II—and some treaties have reciprocity clauses. The reciprocity clauses tend to state that obligations created by the treaty are only obligations with respect to other parties to the treaty.[39] Deciding which model to use for this treaty is one of the hopefully straight-forward obligations of this Conference.

Following a structure to define the Laws of War used by the United States of America's Department of Defense, we will explore the underlying principles of the Laws of War, with the expectation that all these principles will be reflected in the output of this Conference.

3.3.1 Military Necessity

Military necessity is fundamentally the idea that armed forces can take actions that further the goal of quickly and effectively defeating the enemy.[40] This is the justification of force, the justification of destroying property, of imprisoning combatants and potential combatants, and the justification of espionage in wartime.[41] It is important to note that military necessity cannot and does not remove other restrictions on force; it does not override the laws of war, but merely provides a justification for the necessary losses of war. One critical application of military necessity is that States have an obligation to ensure

[39]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 90.

[40]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA.

[41]Office of General Counsel. 2015. Department of Defense Law of War Manual. Department of Defense, USA. p. 52–53.

that they take active effort to minimize civilian casualties, but the extent of the obligation is limited by military necessity.[42]

3.3.2 Humanity

Humanity is the goal of the Laws of War that, fundamentally, aims to minimize human suffering and loss of life. This is what prohibits raping and pillaging captured civilian populations, prohibits killing enemy combatants who are too injured to be a danger, and provides civilian populations immunity from being targets of attacks. Humanity is closely tied to military necessity: humanity prohibits unneeded violence and destruction, while military necessity allows needed violence and destruction.[43]

3.3.3 Distinction

Distinction centers on ensuring that forces have the means of discriminating between civilian and military populations. This entails everything from ensuring that forces are uniformed in a way that discriminates between civilian populations and military forces(not necessarily discriminating between the different nation's forces), to minimizing the connectedness of civilian and military forces.[44] This is the responsibility of forces to facilitate their opponents compliance with the principle of humanity; it is about ensuring that civilian lives are not at risk when the enemy is selecting objectives to attack.

3.3.4 Honor

Honor serves as a good faith expectation. It encompasses the expectation that enemy States won't attempt to take advantage of the other army's compliance with the Laws of War e.g. they won't claim they are afforded some form of protection, such as claiming to be purely present as medical personnel while also conducting military activities. Honor motivates most rules regarding prisoners of war. Honor provides the main justification of only allowing captured members of forces with an organized structure to be treated as

[42]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 56.

[43]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 58–59.

[44]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 62–64

POWs.[45]

3.3.5 Protected Members of a Military Force

The Laws of War also protect certain classes of non-combatants who are interspersed with military forces. Protections for these forces are drawn from the principles described above. Protected classes include individuals whose only purpose and work is care for the wounded, those who have no other function but facilitating those caring for the wounded, and chaplains who serve in no military capacity.[46] It is the responsibility of parties to minimize and avoid injuring these classes of people.

3.3.6 Laws of War and the Cyber Era

This conference should labor to design versions of all of these principles, including both the laws described here and the principles and laws of war not described here, for cyberwarfare. The Dais, however, expects that the largest sticking points will lie in deciding what it takes to provide a distinction between military cyber resources and civilian cyber resources. Standard and trivial filtering would allow easy disabling of a cyber-units effectiveness, if it is required to clearly and automatically label all of its network addresses as having military purposes. Coming up with a standard here is critical, and almost certainly difficult.

4 Contemporary Cyberwarfare Threats

Cyberwarfare *(noun)* The use of computer technology to disrupt the activities of a state or organization, especially the deliberate attacking of information systems for strategic or military purposes.[47]

In contrast to The Oxford English Dictionary, Merriam-Webster does not consider "cyberwarfare" a word. So, depending on which side of the Atlantic the reader is on, this section of the Background Guide is either on a entirely undefined topic or a vaguely

[45]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 66–69

[46]Office of General Counsel. 2015. *Department of Defense Law of War Manual*. Department of Defense, USA. p. 129–130.

[47]"English: Oxford Living Dictionaries." n.d. Oxford University Press. Accessed July 15, 2018. https://en.oxforddictionaries.com/definition/cyberwarfare.

defined topic. This Conference should write a new definition of Cyberwarfare, clarifying what constitutes war, and what does not.

To facilitate this debate, in this section, I will review current known cyberattacks. Some of these were not perpetrated by state actors and are ostensibly off-topic; I include them since everything described is well within the reach of a well-funded and trained forces of a sovereign State and could be conceivably used during a cyberwar or a mixed war, one combining cyber and conventional attacks.

4.1 Stuxnet

When discussing physical damage caused by a cyberattack, the most famous, successful, and publicized is Stuxnet. Stuxnet was a joint initiative by the Israeli and American government intended to slow the Iranian nuclear program.[48] It was discovered in Belarus and largely disassembled by Kaspersky Labs, who, based at first off the sheer number of zero-day vulnerabilities,[49] suspected it was a nation-state-developed attack.

Stuxnet was innovative in a number of ways; its sophisticated and highly targeted code caused it to spread widely, while remaining completely dormant, until it crossed an air gap[50] into the control servers of the Iranian nuclear facility.[51]

Once inside the network, Stuxnet started to record what normal status updates were

[48]Sanger, David E. 2012. "The New York Times." The New York Times Company. June 1, 2012. https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html.

[49]A **zero-day vulnerability** is a vulnerability that is not known publicly. For example if a hacker discovers that holding down the enter key allows them to bypass a login screen, and no one else knows about this vulnerability, then that flaw has been public for zero days, and is considered a zero-day.

People often refer to the number of zero-days as a metric of how sophisticated malware is, since discovering vulnerabilities is a difficult and time-consuming task and it is often cheaper to just build malware using previously known vulnerabilities and just target computers that haven't been updated.

[50]An **air gap** is when a network of computers is completely disconnected from the internet and other networks, so that the only communication between the air gapped network and the outside world requires a human physically moving a USB stick, a CD, or some other physical medium from the outside world into the secure area that houses the air gapped network.

For malware to cross an air gap it is necessary to either have something travel via USB stick, or to have a spy deliver the malware. Considering that Stuxnet's most interesting innovations included a valid certificate authenticating the software as malware free and 4 zero-day vulnerabilities that facilitated its spread, it is believed that it crossed the air gap via a USB stick. In other words, Stuxnet spread as far as it could throughout Iran, and eventually someone put a USB stick into an infected computer outside the air gap, carried the USB stick into the secure network and plugged it in, unleashing Stuxnet on a previously safe network.

[51]Kushner, David. 2013. "IEEE Spectrum." IEEE. February 26, 2013. https://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet.

and monitors showed. It then took control of the centrifuges and began to increase their speed to dangerous levels, triggering explosions and destroying equipment. While doing this it played its recordings of a normally operating facility on repeat to make the centrifuge explosions seem like strange inexplicable accidents, possibly just Iranian mistakes.[52] It succeeded. The Iranians had fired engineers, thrown out perfectly fine centrifuges, and did not seem to be aware why everything was falling apart.[53]

But what makes Stuxnet interesting? It was one of the first recorded politically motivated cyberattacks. It was a sovereign State attacking another sovereign State. It caused the physical destruction of equipment. Had the US and Israel been sneaking bombs into the nuclear facility and blowing up centrifuges, that would be easily construed as an act of war. Was Stuxnet an act of war? And what if American and Israeli officials hadn't leaked to the press which countries were responsible? If Iran had discovered the malware, hardly a certainty considering the malware's years of successfully secret operation, how do they discover, let alone prove to a skeptical world, who was behind the attack in order to respond in self-defense?

4.2 Ukrainian Power Grid

Two days before Christmas in 2015, Ukraine received a Christmas present allegedly from its neighbor, the Russian Federation: nearly a quarter million Ukrainians losing electrical power. The power stayed out for a less than a half-dozen hours in any given region, but the damage to the underlying infrastructure was significant. The underlying attack

[52]Sanger, David E. 2012. "The New York Times." The New York Times Company. June 1, 2012. https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html.

[53]Sanger, David E. 2012. "The New York Times." The New York Times Company. June 1, 2012. https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html.

was a phishing campaign[54] that was then used to spread malware within the network and destroy firmware[56] on devices critical to the power grid's operation. From a legal perspective, the attack could be construed as an act of war, destruction of infrastructure and so forth. However, there is a fundamental problem, we don't have strong evidence that the attack was Russia.

This attack highlights one of the most important questions this Conference must address: what is the standard of evidence that a State is responsible for an attack? Most of the usual tells to pin an attack against a specific actor are actually non-technical, but information gained from traditional espionage, something States are loathe to publish publicly. IP addresses can easily be re-routed, masked, and if an attacker is trying to pin the blame on someone else, it is possible to include fake op-sec[57] mistakes to lead investigators astray. Efforts to analyze malware often center on unreliable measures like the time-zone it was made in, the name of the user account that made the malware, and the language of the computer used to make the malware. None of these are reliable, they rarely stand up in a court of law, and they make the US government's (rejected) arguments in *Case Concerning Oil Platforms* look ironclad.

This conference needs to address this critical question: what should the standards of evidence be for assigning blame for an attack? Further, the output of this conference should also indicate whether this qualifies as an armed attack and/or as a grave use of force justifying self-defense.

[54]**Phishing** (pronounced "fishing") is sending a misleading communication to people to trick them into giving up their usernames and passwords (or some other form of credential). This is one of the most common types of attacks as it requires minimal technical expertise and takes advantage of the weakest part of most computer system, its humans.

A scammer calling random phone numbers pretending to be a bank to get people to reveal their bank account information is an example of phishing against laypeople. Among the most famous phishing attacks in the United States was the email sent to John Podesta by the GRU appearing to be Google telling him to reset his password.[55] In the Ukraine, the attackers convinced operators of the power plant to give up the usernames and passwords needed to gain access to the system.

[55]Lipton, Eric, David E. Sanger, and Scott Shane. 2016. "The Perfect Weapon: How Russian Cyberpower Invaded the U.S." The New York Times. December 13, 2016. https://www.nytimes.com/2016/12/13/us/politics/russia-hack-election-dnc.html?mtrref=t.co&_r=0.

[56]**Firmware** is software embedded inside of hardware. It is normally possible to update and make changes to it, but if broken, can make the hardware useless.

In some devices, broken firmware is irreparable. Destroying it effectively destroys the electronics inside the device. Also, firmware for routers and many internet of things devices is infamously insecure and filled with unfixed and widely known exploits.

[57]In a cyberattack context, **Operational Security** (op-sec) is the effort to ensure that a cyberattack is untraceable. It includes everything from securing your servers so they aren't counter-hacked to ensuring that money made and spent is untraceable.

4.3 Attacking Hospitals with Ransomware

The next attack this document will address was most likely not carried out by a State, but by some form of organized crime organization.[58] A hospital in Kentucky, another in Maryland, and a few more in California were all hit by crippling malware, specifically ransomware,[59] that knocked out operations and/or destroyed medical records.[60]

The technical details of this attack are especially uninteresting, as the attack was essentially using older known attacks (some of it stolen from the NSA) and thus should have been preventable. If, however, a nation-state had decided to devote Stuxnet level resources to the attack, and aimed for Ukrainian Power Grid level destruction, they could have caused severe and significant damage to the chronically insecure medical infrastructure in digital societies. Such an attack is trivially comparable to bombing a hospital, if it was targeted, as these attacks were.

In summary, consider this attack as a reminder that hospitals are critical single points of failure in societies. A targeted attack could severely damage them.

4.4 Untargeted Ransomware hitting Hospitals

Let us now turn from targeted attacks against hospitals to a much more likely attack to be carried out by a nation-state. Computer worms have been seen traveling around the internet for a significant portion of the internet's history. It is not unreasonable to imagine a nation-state unleashing such malware in an effort to cripple the industrial sectors of an enemy's economy, or to undermine morale as the despicable Nazis attempted with the London Blitz. The question then rises: what happens when the malware strikes and cripples a hospital? With conventional bombs, there is active effort in targeting each bomb and it is entirely reasonable to expect states to just avoid bombing hospitals. Malware is most effective when there is no easy way for the unleashing state to control it. Should

[58]The fact we know so little about who carried it out emphasizes the need to ensure a reasonable standard for identifying an attacker.

[59]**Ransomeware** is malware that cripples a computer system, and often takes information hostage in the process. Then, as one might expect, the ransomware demands an exorbitant ransom in exchange for the computer being restored and the information returned. Information is often either taken hostage by using encryption to destroy the file in-place, or by uploading sensitive documents or pictures and then threatening to publish them.

[60]Gallagher, Sean. 2016. "Two More Healthcare Networks Caught up in Outbreak of Hospital Ransomware." Ars Technica. March 29, 2016. https://arstechnica.com/information-technology/2016/03/two-more-healthcare-networks-caught-up-in-outbreak-of-hospital-ransomware/.

states be expected to take precautions when designing generally destructive attacks to avoid hospitals and other protected institutions? Should indiscriminate attacks be fully prohibited?

4.5 Hacking the Office of Personnel Management

The US Government's Office of Personnel Management keeps extensive information on every American with security clearance. Hence, their computer system is a trove of information on almost every official in the upper echelons of the US government. It is believed that the Chinese government breached the databases of the OPM between 2014 and 2015, and siphoned out over 22 million records.[61]

The US Government has viewed this as an act of government-on-government espionage, in effect no different than traditional spy work that has occurred since the beginning of time. The question some have raised since then is whether the scale of the attack is grounds to consider it more than espionage, and if not a grave use of force then a lesser use of force.[62] This Conference should provide clarity about when espionage becomes an act of war and when it is not.

4.6 Hacking Sony

The final attack we will address raises questions about how states are obligated to protect their citizens. Think back to late November of 2014 in the United States: the Republicans have regained control of the Senate and Sony is advertising the heck out of its next movie, *The Interview*. The Democratic People's Republic of Korea took offense to the content of this movie—whose plot centers on a plot to assassinate Kim Jong-Un—and, allegedly, unleashed a massive cyberattack against Sony Pictures. They destroyed property (intellectual property and overwriting firmware to make hardware worthless), and stole and published all sorts of damaging documents (causing serious economic damage to Sony and, arguably by extension, to the US).[63] An analogous non-cyber attack could be an aggressor State destroying a highly productive factory that was owned and operated

[61]Adams, Michael. 2016. "Why the Opm Hack Is Far Worse Than You Imagine." Lawfare. March 11, 2016. https://www.lawfareblog.com/why-opm-hack-far-worse-you-imagine.

[62]The Dais apologies for the weasel words. This exact question was raised by a speaker, whose name the Dais has forgotten, at a lecture at Berkeley's Center for Long-Term Cybersecurity.

[63]Elkind, Peter. 2015. "Part 1: Who Was Manning the Ramparts at Sony Pictures?" Fortune Magazine. June 25, 2015. http://fortune.com/sony-hack-part-1/.

by a non-state organization in the victim State. The traditional clear discouragement of this arguable act of war, was that the victim State could scramble jets to defend itself, or otherwise provide a defense to its citizens and its citizens' property. The cyber realm offers no such method for States to be able to interfere during or in advance of an attack, the world-wide-web is currently a set of servers and actors in borders able to communicate and interact without any regard for those borders. Should international law be updated to create new incentives to not attack holdings in other States? And more importantly, was the Sony hack an armed attack and would an American response with conventional forces be legal? How about an American response with cyber forces targeting the DPRK's government rather than economic forces would that still be proportionate? Clarity is needed.

The Sony hack also serves as a reminder of the critical problem around all cyberattacks. Notice that the hack was "allegedly" committed by the DPRK; the decision to point a finger at the DPRK is motivated by American government accusations against the DPRK, which are most likely based off a mix of forensics, communication intelligence, and signals intelligence. Revealing forensics findings is dangerous for future investigations; highlighting op-sec mistakes that attackers made gives attackers an improved checklist of things to check for when unleashing their next attack. But some standard of evidence is necessary to ensure that force used in self-defense is used against the correct party. What should that standard be? And how does this conference write the standard so that it will survive even as technology changes?

5 Questions to Consider

1. What is the standard for attribution? What degree of evidence is considered sufficient for a state to take action in self-defense? If it is weaker than the current kinetic rules, how do we ensure that innocent states are not attacked?
2. What is a proportional cyber attack? Can a kinetic attack be responded to with a proportional cyber attack? Can a cyber attack be responded to with a proportional kinetic attack?
3. Should states be expected to take precautions when designing generally destructive attacks to avoid hospitals and other protected institutions? Should indiscriminate attacks be fully prohibited? Should targeted attacks that cause impact to other

systems be considered the same as an indiscriminate attack?

4. What is the Cyber equivalent of Distinction and Honor?
5. Do the other components of *jus in bello* require a restatement for the Cyber Era?
6. What is the line between sophisticated espionage and an Act of War?
7. To what degree does a State bear responsibility for the actions of its citizens? Its representatives? And what if a State's representatives or citizens are unknowingly spreading malware?