Cyber Defense from "Reduction in Asymmetrical Information" Strategies

Guy-Philippe Goldstein

This essay confronts two main problems in cyber defense: the attribution issue (who is attacking?) and the threshold issue (is it worth all-out war?). Starting with a war-game scenario, an analytical framework based on the *Tallinn Manual* is suggested to delineate cases for wars and areas of crises. The prosecution of cyber crises is then proposed through two "reduction in asymmetrical information" strategies. The threshold issue can be alleviated with a better understanding of observable and simulated effects on the defending networked nation modeled as a system, drawing on the initial concept proposed by Col. John Warden. The attribution issue must be solved through excellence in elucidation methods and internationally supported coercive investigation, inspired by Thomas Schelling's compellence. The growing preeminence of the digital domain in our modern societies could make these strategies among the building blocks of a new doctrine for military and political stability in the twenty-first century.

Keywords: cyber weapon, cyber defense, deterrence, doctrine, compellence, attribution, thresholds, escalation, *Tallinn Manual*

Hence the saying: If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle.

Sun Tzu, *The Art of War*[1]

Guy-Philippe Goldstei, MBA, HEC (France) is the author of Babel Minute Zero, a bestseller about international cyber warfare.

Introduction: A Regional Scenario

It is 9:00 in Country X. In the capital state, bank ATMs have stopped working. Some online customers cannot access their bank accounts at the top three national banks. In some cases, the balance in online accounts has been wiped to zero. Cell phones are barely functioning. The attack seems to be of a new kind. The effects are the same as with the Estonia cyber attacks of 2007. However, technically, it does not look like a distributed denial of service attack: no massive amount of IP-packets clogging servers has been detected. No immediate remedy is at hand. How long will this last? Can data be recovered? Is this a first wave announcing further attacks? On the streets of Country X, anxiety is quickly ramping up.

Country X is not alone. A week earlier, a prominent software security firm from Country B identified a new malware: GlobalWorm. Though its mode of action was unknown at the time of discovery, GlobalWorm seems to have infected many systems across various countries. In an alert bulletin, the software security firm is now linking the current attack against Country X to GlobalWorm. Furthermore, other countries infected by GlobalWorm are experiencing difficulties, including friends as well as foes of Country X. However, only Country X is suffering severely harmful effects.[2]

Who is responsible for the attacks on Country X with GlobalWorm? What type of threats does GlobalWorm pose to Country X? How should Country X retaliate?

The first two questions frame the third one. To further complicate matters, the security software company that knows GlobalWorm best has tight links to the military apparatus of Country B – and Country B is not a close ally of Country X. As the National Security Council of Country X convenes, the questions around the table coalesce: Is this another blow from Country Y, the proverbial enemy of Country X? Did Country Y not just increase investments in cyber weaponry?...Or is this coming from Country Z, a country whose relationship with Country X has dramatically soured over the last five years?

The head of state of Country X asks the three questions that are foremost on his mind:

a. Can you prove to me that this is related neither to Country Y nor to Country Z?
b. How much time do I have left before I am forced into retaliation?

c. How can I retaliate if I do not know the answers to my first and my second questions?

The head of intelligence for Country X confirms that at this stage, there is no clear indication that Country Y or Country Z is behind the attacks – though it is possible, he emphasizes. However, the possibility of manipulation by Country B cannot be dismissed either. Additionally, although the attacks have shocked the population, they have not escalated in kind over the last eight hours. It is not possible to say how the threat will evolve – if indeed it does evolve. What is clear is that Country X has been weakened. Without some form of elucidation, restoration, and retribution, its status as a cyber power will be contested. This does matter. In this day and age, it is understood that there will be major combat operations in cyberspace. So the domination of cyberspace becomes a test of overall military power.

The minister for foreign affairs says Country A, one of the closest allies of Country X on the international scene, does not possess clear indications about the origin of GlobalWorm's infection. However, as Country A considers it a global problem, Country A will not allow Country X to retaliate without evidence being put to the fore. To top that, Country A says that retaliation needs to be closely coordinated in case of cyber reprisals. After all, neither Country X nor Country A understands what tricks lie inside GlobalWorm. The situation is different from scenarios in which Country X is the attacker: Country X controls neither the test nor the environment. A wrong maneuver could be perilous for Country X, perhaps for everyone else too. All sorts of manipulations can be envisioned. There are just too many unknowns.

This state of strategic confusion is perhaps what the offender had in mind when designing the attack. Country X does not know yet what bargain is at work, nor with whom. The only clear offer comes from Country B: via its software security firm, it could bring unique expertise and support of GlobalWorm. But this help would probably come at a price. Additionally, Country A and Country B are global peer competitors. Country A may object to Country B helping Country X. Relationships between Country A and Country X could be damaged.

In this scenario, conventional or strategic deterrence tools are not operative. Country X is actually faced with strategic paralysis.

Perfect deterrence theory posits that "response in kind" is an optimal strategy.[3] It demonstrates that the defender has a credible retaliatory threat. At the same time, it signals that Country X is not necessarily seeking escalation – what Huth describes as a "firm-but-flexible" negotiation style.[4] Additionally, not to commit to full-fledged escalation but to engage in firm response allows opening up options without exercising them. This is the position most favored by politicians as well as financiers. It is also an optimal situation with regard to the decision laws of cybernetics. But in the current predicament for Country X, response in kind is not possible. First, there is a major obstacle: Country X does not know against whom to respond in kind. It is faced with an attribution issue.[5] But even if it knew with certainty, Country X would still face a second major obstacle: it may not know exactly how to respond in kind.

Let us assume for a moment that Country X has established that Country Y is the aggressor. Since bank ATMs, online banking accounts, and some cell phone networks have been breached, Country X tries to respond in kind. Let us also assume that Country Y has not hardened the cyber security in advance around what it would know to be the respond-in-kind targets of Country X's reprisals. An in-depth examination is still needed as to whether Country X would be able to inflict a level of degradation at least equal to what Country X suffered. If Country X tries but cannot equal the first blow, then its threat credibility will be further diminished. Yet if it retaliates too hard, it could trigger unexpected consequences and the conflict's spiraling. Unfortunately, at the current stage of technical advancement, cyber weapons' effects are hard to predict precisely – even more so if improvised for battle in the context of rapid retaliation. Country X is faced with a second problem: a thresholds issue.[6] Country X does not have a response-in-kind solution, that is, a credible retaliatory threat. A doctrine of "massive retaliation" policy in cyberspace may be subject to the same critiques as the one formulated by Will Kaufman against Eisenhower's NSC-162/2 in 1954[7] – with the added caveat that "massive" is hard to define, unless it applies to assured mass civilian casualty. At the same time, the absence of retaliation evidently goes against the principles of response in kind. It would invite further aggression.

At this stage, there are no good retaliatory options for Country X. If attacks have reached certain damage thresholds and Country X feels otherwise threatened by its geopolitical situation, then it may want to

intimate to neighboring countries that attacks will have consequences. It will then try to respond in kind imperfectly by highlighting its most capable and credible non-cyber, kinetic threat, for example by flexing muscles through a show of air or ground forces. This measure will have adverse diplomatic consequences if attribution is not well established, and it could backfire if cyber attacks continue, actually raising the credibility stakes for Country X now that it has exposed its conventional forces. However, if a cyber attack does not seem to exact too high a price and if its origins remain efficiently obfuscated, then Country X may want to defuse tensions and lower the stakes. Difficulties could be attributed to non-state or technical origins. Then Country X could accept the help from Country B via the software security firm. Of course, as noted, this help would come at a price.

A First Strategy of "Reduction in Asymmetrical Information": Elucidation of Thresholds

An Evaluation Framework

An optimal course of action may exist for Country X. First it must understand what types of attacks it is facing in order to devise the best response. In particular, two main informational issues, mentioned above, must be solved: attribution and thresholds.

Attribution must be strictly linked with the issue of "plausible deniability" because at stake are the political and diplomatic consequences of lack of attribution. Threshold definition is an even more complex problem: there is an inherent difficulty in defining "simple, recognizable, thresholds" in cyber-attacks.[8] Actions leading to thresholds can be split into two types: (i) those with direct effects on a nation (such as industrial disruption or loss of life) and (ii) military preparations that precede these effects (such as military mobilization or reconnaissance operations). Does the setting of logical trap doors in an opponent's electrical grid constitute an act of war? Is there an equivalent in cyber warfare for enemy mobilization and massing at the borders? These questions cannot be easily answered, especially as they refer to issues such as the thresholds for retaliation along the "curve of credibility."[9] The *Tallinn Manual*, written at the invitation of the Tallinn-based NATO Cooperative Cyber Defence Centre of Excellence, is a necessary starting place but does not at this time authoritatively answer all of these questions.[10] In a more general and historical sense, these are issues at the heart of the strategic conduct of nations, answered on a case-

by-case basis and grounded in practical reality, but they have not been comprehensively formalized. Cyber strategy may necessitate an additional effort at conceptualization. Though the task is beyond the scope of this article, some initial shortcuts may be noted.

A starting place, cited in the growing literature on cyber warfare studies as well as the *Tallinn Manual*, is direct effects.[11] This is an approach that can be understood by many militaries around the world, starting with the US Air Force, still a proponent of Effect-Based Operations, linking actions, effects, and objectives.[12] As highlighted by the *Tallinn Manual*, it also has legal precedents, especially around the term of "scale and effects."[13] Yet what effects constitute crossing a red line for the defender? It is easiest to start with what is benign or tolerable, then explicate what can never be tolerable and would automatically elicit military retaliations. In between lies the territory of the crises.

For example, espionage is tolerable (albeit not officially). It enjoys international tolerance because it is "an extension of monitoring regimes" that thereby enables functional cooperation.[14] This tolerance seems to have extended to some cyber applications of espionage.[15]

What is never tolerable, what would automatically elicit military reprisals, is action leading directly to significant loss of life among non-combatants. In general, this action would be interpreted as a voluntary breach of the laws of armed conflict with regard to *jus ad bellum* as expressed in the 1949 Geneva Convention and clearly restated by the *Tallinn Manual*.[16] In strategic terms, what is never tolerable, what means war, is also initially obvious: destruction of a part or the totality of the sanctuary. This extends to any significant attempt at suppression of the protective institutions of the sanctuary. Because the state holds the monopoly on large-scale violence,[17] both the capabilities for large-scale violence and the monopoly-holding decision center commanding their use must be protected. In practical terms, preserving the sanctuary means first and foremost protecting the life of civilians. War then becomes inescapable if the nation suffers a significant loss of life.

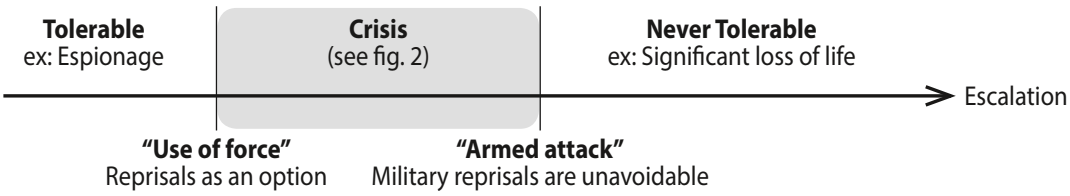
With regard to large-scale violence capabilities, some weapons are essential: first and foremost the nation's survivable second strike force, but also any weapon systems deployed so widely that malfunctions would significantly hamper the defense of the sanctuary. These include the specific networked communication systems and sensors required for

the proper use of those weapons. They also include the intergovernmental communication systems necessary for the head of state and staff to command and control these capabilities, as well as for heads of states to communicate. Such provisions were agreed upon by the two superpowers during the Cold War. The 1971 Accident Measures Agreement and Hotline Modernization Agreement established protection of satellite communications essential to US-USSR communications in times of crisis, as well as the communication facilities for missile warning systems.[18] In addition, attempts at first responder forces and at medical assets that limit significant loss of life constitute red lines. Elements reflecting this understanding were agreed upon by Russian and American diplomats in 2011 and were included in the *Tallinn Manual*, as a way to more generally align the conduct of cyber operations with the current laws of armed conflict.[19] These measures include assets and communication systems for command and control for medical and first responder forces, including with the head of state. Protecting the communication systems does mean preserving data from external corruption: if data cannot be protected then, de facto, the communication systems as means of sending the right instructions are being sabotaged.

Finally, there is the question of economic protection of the sanctuary. At what point do economic damages become so harsh that war is inescapable? Political literature hints that economic hardship can bring about political change: recessions can lead to changes of the ruling party in democracies[20]; depression can bring about regime change in the form of the rise of extremist movements, as shown in the interwar period.[21] If such economic upheavals are brought about by cyber sabotage, they constitute a coercive action intended to destroy the political integrity of the State.[22] This political result would come on top of the resource constraints imposed on the military by economic hardships, which in themselves constitute a threshold if there is significant reduction in military preparedness. Other scenarios could also hint at direct manipulation of the political control organs of the state (for example, electronic corruption of voting systems or mass electronic blackmailing of elected officials). If political majorities could be defeated by such cyber sabotage, it would constitute a significant attempt to weaken the integrity of the state, and thus the crossing of a red line.

In this framework, those effects that are never tolerable hurt so severely that they are easily and blatantly recognizable as such. In the *Tallinn Manual*, attacks yielding such effects are construed as "armed attacks."[23] At this threshold, military reprisals are a certainty. If the identity of the attacker is known, then it is subject to the idiom of military action established among states. The rules of this idiom apply, ensuring what Thomas Schelling has called the diplomacy of violence.[24] States are entering a game of escalation, from conventional retaliation to potentially strategic reprisals. Cyber weaponry becomes an adjunct to other weapon systems.[25] States can credibly respond in kind with non-cyber weaponry. This will bring clarity and recognizable accents to this dialogue, as illustrated in figure 1.

Figure 1. Decision Framework with Tolerance for Effects



If the effects are recognizable and have an impact on civilian populations or assets although the identity of the attacker is unknown, then the action can be construed as terrorism. Hackers enabling these attacks without a recognized national attribution are acting as unlawful combatants[26] or unprivileged combatants,[27] that is, civilians who directly engage in an armed conflict in violation of the laws of war. Because they cannot be linked with a state bound by the limitations of the 1949 Geneva Convention while conducting military operations against military targets, they pose a de facto threat to any civilian targets the moment their attack causes harm that is never tolerable. The response to such a terror campaign must lead to the arrest of the hackers, or at a minimum to punishment of the state harboring them, as per the evolving legal standard applied in the attack against the Islamic Emirate of Afghanistan after the events of September 11, 2011, and in particular in light of UN Security Council Resolutions 1368 (2001) and 1373 (2001).[28] As in the case of nuclear terrorism with lack of attribution, the collection of intelligence becomes central for any retaliatory measures.[29] This issue is explored below in the section on joint compellence.

In the area between the tolerable and the never tolerable exists the territory of crises and its many shades of gray. The harm is conspicuous

enough to be construed as a use of force but its severity is not elevated enough to identify it with certainty as an armed attack.[30] According to the International Court of Justice, as cited in the *Tallinn Manual*, "not every use of force rises to the level of an armed attack."[31] The crisis can be kept outside of the public eye – a default option to avoid tying one's hands too much within the unchartered waters of cyberspace. Still, the crisis will be real. Uncertainty here has many sources. The never-tolerable effects may not be observable yet, but they could be perceived as an imminent outcome: if online banking problems spread and last a few weeks, would they lead to financial panic? Could losses be easily recovered? The same questions apply if the energy grid is breached. On Day 2, it might be hard to tell. Additionally, not only might direct effects be hard to assess; the meaning of the enemy's military actions in cyberspace, its "virtual mobilization," might also be difficult to evaluate. The last point is critical because, following the rules of warfare first described by Sun Tzu, surprise is the key to victory[32]: the better warrior will not create patterns or precedents. His or her moves will be difficult to evaluate.

Nonetheless, this grey area must be addressed and charted. The escalation categories delineated by Herman Kahn in *On Escalation*[33] are useful here. What is the intensity of the attack, as a probability of reaching the never-tolerable level? How many different components of the nation seen as a system are being attacked? What is its evolution and tempo – especially as intense acceleration could be indicative of impending physical military actions? Using Herman Kahn's delineation, a simple distinction can be drawn between:

a. What is not benign, but reflects self-limitation in escalation: the attack is limited in intensity and cannot be construed as threatening non-combatants; it is limited in scope: only one type of targets is being attacked; it is limited in its temporal dimension: it happens only once or a few times, or has a date of termination. These attacks can be labeled as limited.
b. What is not benign and can be construed as potentially escalating: the intensity or scope of the attack seems not to be self-constrained and could be escalating; or there is repetition and acceleration along the temporal dimension, without a distinct termination date. These attacks can be labeled as escalating attacks.

For example, if GlobalWorm was recognizably set to alter the functioning of only very specific software or equipment, if the software or equipment specifically targeted by GlobalWorm was only for military use or dual-activities, if the effects did not lead to significant collateral damage among civilian personnel or civilian life, and if GlobalWorm had a recognizable date of expiration – for example with digital certificates protecting it and it was due to expire at a certain time – then the GlobalWorm attack against Country X would be a limited attack. This does not seem to be the situation in the Country X case. Effects are not limited and circumscribed to specific equipment, but are escalating. They are also hard to recognize: what may be the secondary effects of 48 hours without online banking?

In simplified terms, effects that are recognizable (that is, they can be acknowledged with all immediate consequences fully understood)[34] but escalating ,and effects that are hard to recognize (that is, not all immediate consequences are fully understood) can be grouped together: both pose a high risk of surprise, miscalculation, and escalation (figure 2).

Figure 2. Decision Framework for "Crisis" (Detail)

		Discerning Effects	
		Recognizable & Limited	**Hard to Recognize/ Recognizable & Ascalating**
Discerning Identity	Known	Special Ops/Limited Strike Warning shot	Attacks against some tactical weapon systems Low intensity attacks against civilian
	Unknown	Convert Ops Espionage Operation (uncovered)	Sabotage campaign Low intensity terror Reconaissance Operation

An Evaluation Process

The "hard to recognize" category of effects remains highly problematic. A sufficient level of prediction for these effects is difficult to achieve: these are not what the *Tallinn Manual* terms "reasonably foreseeable" harms.[35] To rely on observation of effects as comprehensively as possible with centralization of intelligence, or to develop an analysis of the mode of action of the malware in its software environment is not sufficient. The

impacts on a "nation seen as a system," to use the concepts of Col. John Warden,[36] cannot be understood through these necessary but insufficient first steps. Such an evaluation is the purview of modeling, simulation, and analysis of system of systems, including economic and social components. The objective of this evaluation is to determine the expected political harm against the defending state.

In a defense context, the further analytical step will naturally lead to a reverse-engineered "Effect Based Operations" (EBO) analysis. The point here is not to achieve the required precision necessary for an offensive use of EBO that has been elusive so far with current software tools.[37] The objective is different: it is, in a defensive use, to deploy an idiom for cyber warfare made of internationally recognized thresholds. This baseline would link cyber actions with direct effects and intended objectives. It would also serve to legitimate all options reactions, including diplomatic or kinetic actions. Here, "simple, recognizable, and conspicuous" will trump "most precise." To be trusted, this idiom can only be enunciated by the most preeminent cyber powers.

However, international participation in its development by other nations, perhaps along the logic of concentric circles, will ensure that it is recognized by many and thus becomes conspicuous. To be credible, it will have to reflect the real impacts on a nation's curve of credibility. To that effect, it may follow the path laid down by Col. John Warden, and pursue a robust course of studies and simulations to understand the networked nation as a system. Not only could the internet be tested in virtual "cyber ranges;" sub-components of the nation could also be simulated. All sorts of organizations and infrastructures take part today in the release of big data sets, from open data projects in public sectors to application programming interfaces (APIs) in internal corporate and industrial processes,[38] and to social and political sentiments as expressed in social networks. This approach, in turn, promises to help develop a better and much finer baseline modeling of the networked nation as a system. These dynamic data models can then be tested against simulated shocks. Here too, exactitude is not as important as agreed-upon, credible, ballpark estimates. However, this development will be an ongoing effort, as cyberspace is consistently evolving.

Understanding thresholds does not resolve the second main informational issue: attribution. The latter will require a specific intelligence, diplomatic, and coercive effort.

A Second Strategy of "Reduction in Asymmetrical Information": Elucidation of Attribution with "Joint Compellence"

Attribution

Because cyberspace consists of three pillars − hardware (calculation, memory, or communication devices), software, and brainware[39] − intelligence work must investigate and develop hypotheses for each of these three sources. Clues as different as IP traffic patterns, styles of coding, and methods of actions should feed an attribution matrix. It should also include classical human intelligence on hackers themselves and their political sponsors. These investigative activities should adhere to the best practices in elucidation, with emphasis on deductive methods applied to intelligence as suggested by Ben-Israel.[40] As one methodology in the context of general intelligence works suggests,[41] attribution hypotheses could be laid out in different buckets (for example, "Hypothesis #1: Country Y is the aggressor"; "Hypothesis #2: Country Z..."). Then, empirical data refuting each hypothesis could be set against each bucket. Stacking data against attribution hypotheses would be a first step toward identifying which country is most liable to be the originator.[42] This would require advance identification and simulation of the multiple models of necessary preparations required to launch a massive cyber attack for each country. These models of preparation would of course include additional defensive hardening efforts and obfuscation efforts. Ideally, then, deductive A/B tests in the manner of controlled experiments launched against possible culprits could be set to confirm or infirm attribution hypotheses. For example, taking a page from the strategies used by fictional character George Smiley, by simulating unexpected effects of the malware, the true place of origination could inadvertently reveal a surge in unease and embarrassment.[43] The detection of this unease would help with attribution.

Excellence in truth seeking is critical for establishing defense. It is instrumental in convincing allied countries that one is not trying to manipulate them. In return, once genuinely convinced, these countries can then serve as the equivalent of character witnesses toward the greater world audience, and can increase diplomatic acceptance of retaliatory

options. Excellence in truth seeking also ensures that the political echelon of the defending country is not making a grave attribution mistake. The government has confidence in its own decision. At this point, the government becomes more at ease than before the elucidation phase to explore non-public, non-retaliatory measures if need be. As in any counter-intelligence work, it is perhaps best to temporarily maintain the illusion for the enemy that his stratagem has not been uncovered.

In cyberspace, truth is power, as it is for any other information domain, such as traditional intelligence.[44] The means and methods of establishing a quasi-incontrovertible truth are key instruments of power. As such, they can become instruments of influence. One day, the cyber-diplomatic scene could resemble the civilian internet mainstream scene, where some of the largest search engines or reference content providers (such as Wikipedia) are already vying for the highest relevance in terms of content. After all, the most important feature of any information system is the ability to distinguish the right signal.

However, it may be difficult to share the attribution techniques and data described above with a large audience of countries, as is often the case in intelligence sharing. In an increasingly multipolar world, this difficulty could lead to further defense paralysis or diminished deterrence credibility if no method to jointly carry out attribution elucidation is established. However, such a method may exist by way of a large-scale deductive test carried out publicly, especially as deduction is a superior method for truth elucidation in intelligence analysis.[45] In *Cyberwar*, Richard Clarke and Robert Knake highlight the "arsonist principle": the burden of the investigation should be shifted from the investigators to the nation in which the attack was launched.[46] If the suspected nation refuses to cooperate, it would be held responsible. Then an international body – what Clarke and Knake term an "International Cyber Forensics and Compliance Staff" – could suggest cyber sanctions, from shutting down certain ISPs to even blockading the nation from cyberspace.[47]

Building and expanding on this approach, there is actually the possibility to defend against some of the potentially most severe cases of cyber warfare offensive and reestablish cyber-deterrence.

A crucial initial observation is appropriate here: in addition to forcing attribution via the arsonist principle, this approach can actually establish it formally. In diplomatic terms, it can deny the offender the option of

plausible deniability. Establishing attribution is as much an intelligence investigation as a diplomatic process. Other nations must be convinced. First, the credibility of the truth is best established when other observers (or testers) can confirm or infirm the attribution hypothesis. This social process is well established, from the two-witness rule governing the trials of treason as early as the Elizabethan era in England,[48] prefiguring Hooper's rule on concurrent testimony[49] to modern statistics where confidence in predictions is increased by the number of observations. To create a public test is to force other nations and their people to become observers. Second, a diplomatic process ensures higher coordination and thus strengthens the cyber blockade required to pressure suspicious states. The strength of the blockade is vital for the threat to be capable. If it can be significantly evaded, as Western powers managed to do during the Berlin Crisis of 1948 against the Russian blockade, then the threatening country fails.[50] If the blockade cannot be evaded, then the threatened country is forced to decide between escalation and backing down – and if the stakes are too high, it may back down as Russia did during the Cuban Missile Crisis. In addition, carrying out the attribution process first with close allies, then with a wider group of nations, might foster goodwill, rapprochement, and greater understanding toward the defending state. That, in turn, frees up political margins of maneuverability if the defending state is to move toward additional diplomatic, economic, or military sanctions beyond cyberspace and a cyber blockade. It lends further credibility to what is essentially a compellence strategy, as described by Schelling: "a threat intended to make an adversary do something."[51] Suspected states are compelled to collaborate or else they will continue not only to suffer from the cyber blockade, but also to single themselves out. In that new context, countries wanting to prove their goodwill will genuinely cooperate. Perhaps they may even share their own intelligence with regard to attribution, as a further proof of goodwill. Countries that do not cooperate will de facto reveal their true intent.

In addition, cooperation is all the more easily compelled when it means that cooperating countries do not have to lose face. Taking a page from Rattray and Healey's model of public health for cyber security,[52] the metaphor of World Health Organization (WHO) investigation teams at times of pandemics can be used. National governments do not have to be nominally accused – they do not have to be held initially responsible for the pandemics. Officially, the blame is placed on the malware or the nefarious

teams of hackers behind it. Using the public lack of attribution for the sake of the compellence action, the coalition of defenders can then request the heads of the suspected states to cooperate. A cyber blockade can still be implemented, analogous to WHO quarantining regions or countries during pandemics. Thus the cost of not cooperating still weighs on the offenders – and it will grow as other states cooperate and the offending state becomes ever more isolated. Conversely, the cost of cooperating is lessened because there is no loss of face. And still, there is a genuine threat, that is, a cost for having launched the operation in the first place: finally accepting cooperation, the offending capabilities (servers, codes, hackers) will be publicly branded. They will be rendered inoperative. Ongoing cooperation – and the additional intelligence it will provide – will help maintain this calculus. This is the end game. Defecting nations are forced to cooperate again. Their investment in defection capabilities is nullified. But there is not necessarily the audience cost attached to backing down. This makes renewed cooperation acceptable, and thus potentially stable. Additionally, the difficult task of a formal, public attribution, requiring a very high degree of certainty because of its public format, is rendered unnecessary.

Strategies and Requirements for Joint Compellence

To be successful, this strategy must leverage the attribution efforts already mentioned. The quality of intelligence is critical in conducting this compellence approach. Heads of state are at the heart of this strategic conflict. Their methods and manners of communicating threats affect the credibility of their retaliatory threats. The defending head of state, assisted by a coalition of friendly countries, behaves like a police investigator interrogating suspects: "Give us access and information. Cooperate with us – or we keep you locked down." This is bargaining, comparable to an actual police interrogation.[53] The better the intelligence, the better the design of the interrogation and the more efficient the process: "Information power may be the most important source of power" in interrogation.[54] Used as an argument in the interrogation process, it demonstrates the deep knowledge of the interrogator, thereby reaffirming his credibility because he cannot be deceived. The interrogated will then hesitate to misinform; at the same time, the interrogator demonstrates that he can be a knowledgeable partner. A cooperation deal will be solid. Finally, as

mentioned above, the interrogator can run tests to check the reaction of suspected states. These tests could simulate unexpected consequences for the defending state. By counter-manipulating, the defending state can instill doubts in the aggressor: cyber weapons are not reliable and could trigger an undesired escalation. The defending state could more easily mobilize external sympathy and support as its vital domestic interests are made more vulnerable to the malware. Solidarity from other countries is all the more extended as the malware has no defined origins: anyone could be its target. The diplomatic aspect of the compellence process helps turn the strength of the attack against the attacker, as in Judo. The harsher the cyber attack, the stronger the solidarity between the defending state and its ally – and the tighter the cyber blockade against suspected states. Defense retakes the initiative. It can dictate the tempo in escalation control.

This compellence strategy to resolve attribution is feasible because behind a sophisticated attack, there must be a nation-state. Non-state actors are necessarily harbored by advanced developed states. Terrorist organizations based in under-developed, failed states do not currently have the technical capabilities to wage strategic, sophisticated cyber attacks. For example, Stuxnet was a piece of coding developed by very talented IT engineers; it used digital certificates perhaps stolen from two legitimate Taiwanese companies,[55] and it had been tested on a full cyber-physical model that included replicas of the P-1 centrifuges.[56] However, all this requires deep pockets to recruit and retain talent, actual local access to a multidisciplinary pool of talent (especially if cyber-physical models are necessary), and constant training and development as cyberspace is upgrading constantly, not to mention secret services to infiltrate or enable access to privileged software information. These are development capabilities that currently cannot be acquired in tribal areas. In all probability, behind any ad hoc group launching a sophisticated cyber attack, there will be the active sponsorship of an advanced developed nation. Advanced developed nations are to become ever more dependent on access and development in cyberspace for data, instructions, and actual processing. A large portion of business-to-business communication and data processing is shifting to the so-called cloud, that is, servers often situated in foreign locations. In that context, the crippling effects of a cyber blockade may be particularly acute for advanced developed nations that come under suspicion.

This strategy will work if allies of the defending country are also compelled or incentivized to act. Ongoing coordination, agreement on norms, and sharing of processes are prerequisites, before a crisis starts. In practice, cooperation levels might correlate with existing circles, from the closest allies to the most distant – embracing in cyberspace what is currently the cooperative arrangement at the overall political level.[57] Additionally, in order to give credence to the whole process, there can be a move toward greater cooperation within circles, and greater rapprochement between adjacent circle levels. Gently pointing the way forward has the advantage of solidifying the current level of international cooperation. Even more importantly, the ties that bind these cooperative links should find a credible translation in practical terms. For example, friendly countries can employ additional layers of software used by other friendly countries. Joint use of the same software or standards increases the risks of unexpected consequences for the attacker. It credibly conveys the possibility that to attack one country is to attack all of its allies. Shared use of the same software in cyberspace may play the same role as the US garrison in Berlin during the Cold War[58]: it would create automatic involvement and leave no doubt that the compellence process would be carried out jointly by a coalition of friends.

Finally, defending countries must acquire redundant cyber capabilities to absorb the first shock. Redundant communication and computing capabilities temporarily alleviate bottlenecks. Semantic manipulation could be partly offset by periodically saving critical data in write-only, non-volatile data storage in order to retrieve true pre-attack values. But defensive measures alone are largely insufficient. Without confronting the will of the enemy to learn new attack techniques, the attacker will continue to learn and adapt, mimicking the coevolution (Red Queen) dynamics found in nature.[59] Deterrence will not be achieved. What must be confronted is the attacker's will to learn and not share new offensive techniques: a cost must be imposed on this will to learn and not share. Nevertheless, to absorb the first shock is elemental. Conventional deterrence models posit that short-term weaknesses on the part of the defender can invite attacks[60]: for example, a first blow might be so hard that the defender would not have time to respond properly and mobilize a coalition of allies. Additionally, the attribution process should ideally entail an alternate international team of inspectors. This would ensure that the long "shadow of future"[61]

is preserved: whatever happens, the truth will survive. Attribution will be made. Responsibilities will not be evaded.

To summarize, once attribution is made, and once effects can be recognized and evaluated within a defending nation's curve of credibility, informational asymmetries in favor of the offense cease. The idiom of military action is restored to the benefit of the defender. The defender can make credible retaliatory threats. In particular, after effects are properly recognized, the defender can credibly retaliate in kind by using non-cyber means – diplomatic, economic, kinetic, or strategic. All options are made available anew, thereby giving more weight to the hand of the defender. Non-cyber retaliatory threats may even be superior if proven non-vulnerable to cyber attacks: their resilience will render them highly capable. By setting a limit to the potentially confusing game induced by cyber-only retaliatory means, the defender will signal the translation from cyber attacks to real effects, thus providing a clarity that will force the attacker either to back down or to escalate. In particular, the restrictive environment created by joint compellence will become a difficult situation for the attacker. Again, as the Cuban Missile Crisis demonstrated, in such a situation the non-status-quo power may prefer to back down rather than escalate.

Conclusion: Toward a New Political and Military Doctrine for the Digital Age

The necessity of establishing equivalence between cyber and non-cyber weapons by means of equivalent effects – and the need to switch from cyber to non-cyber retaliatory means – demonstrates the criticality of reframing cyber warfare operations in the context of other weapon systems. Following Edward Luttwak,[62] one-force cyber strategies may at this stage be as confusing and minimally operative as what Luttwak dismissively termed "nonstrategies" – namely, other one-force strategies claiming strategic autonomy such as "naval strategy," "air strategy," and "nuclear strategy."

However, centers of gravity have always shifted as technological disruption changes warfare. The centers of gravity during Cold War fighting were quite different from the ones at the time of Gunderian's blitzkrieg or that of Vauban and its massive fortresses. In the naval domain, strategist Julian Corbett determined that gaining sea control was ensured not by conquering areas of water, which are impossible to hold, but by

ensuring the act of passage on the sea.[63] As conflicts move into the digital domain or digital *logos*,[64] centers of gravity are going to shift. The higher criticality of the semantic domain over the physical support reduces the relative importance of communication lines: the internet was built to send information despite the unavailability of hardware. What becomes critical is to ensure that true meaning is protected: Who is attacking? What is being attacked? To know attribution and to recognize and predict effects become the higher grounds. These are cognitive centers of gravity. In strategic terms, this is knowledge supremacy: to control and to preserve the nation and its sub-systems from information manipulation. To put it differently, in an information domain, truth is the highest ground.

The importance of the digital information domain relative to other components of the networked nation as a system may alter strategic priorities. Additional industrial shifts could further strengthen this new order of priorities. As software continues to "eat the world"[65] and the value of data and data-based applications becomes ever more important, the preserve of the digital *logos* could become as valuable as the physical assets it reflects and partly controls today. In some vital areas, this is already the case: today, wealth is measured and exchanged by means of electronic bits identifying monetary value. So while cyber warfare today is a non-strategy in Luttwak's definition, there is a possibility, small and remote but not nil, that strategy in the digital *logos* claims its autonomy, that it represents both means and ends. Information systems, from DNA to spoken language, are critical to the management of any organism. Therefore such preeminence for the digital *logos* should not be surprising in theory.

This ongoing transformation will mark a profound change in the role of the state defending the nation. The state must maintain the monopoly over large-scale violence, which can be construed as protecting physical assets from corruption by kinetic force. It will also have to protect the reliability of data in use by strategic military and civilian systems, and at a higher level, maintain accuracy of strategic information for the situational awareness of the nation as a system. The state will be the custodian of last resort for the truth.

All these remote possibilities are portended by the ever-increasing acceleration of IT calculation and storage capabilities. As an example, the calculation power of top supercomputers will increase by a factor of at least 10^3 Floating-point Operations per second (FLOPs) over the next

ten years.[66] As the scale of calculating power continues to increase, major changes in machine learning and simulation cannot be discarded.[67] The limitations found today in analysis of EBO and the nation as a system may be as temporary as the difficulties in the field of artificial intelligence. For decades, artificial intelligence has been defined as a difficult field of research.[68] Today, it is proving promising again.[69] In this context, advanced EBO capabilities for further simulation and analysis of effects could also change the calculations regarding national powers.

However, an increase in simulation means further predictability: a longer, more predictable view of the game is then possible. The better the information is regarding each party's true capability, the lesser the risk of war. Additionally, both Zagare[70] and Axelrod[71] demonstrate in their respective works that the longer the perceived game, the higher the chances that cooperative (or status-quo) strategies dominate.[72] Finally, successful enforcement of a joint compellence strategy would also, in the long term, favor the status quo: if the fruits of defection are being denied and the end game of joint compellence is further cooperation, then defection becomes an unnecessary cost. This automatically increases the relative value of the status-quo choice (namely, continued cooperation). As Perfect Deterrence Theory posits, the overall increase in the value of the status-quo choice over any defection strategies is also one of the most important factors to ensure stability.[73]

In this context, the complementary approaches of advanced nation-as-a-system simulations and joint compellence suggest that the accelerated immersion of our human civilization into the digital *logos* could become an additional force for peace and stability. These strategies of reduction in asymmetrical information could serve as key building blocks toward a new doctrinal framework for the societies of the digital *logos*. This doctrinal framework will continue to promote peace and stability and will have to integrate current nuclear and conventional deterrence doctrines. It will also recognize the new preeminence of digital information systems in civilian affairs and therefore in military affairs. Ultimately, it will lead to a refined definition of what is a conflict. The doctrine of mutually assured destruction has transformed wars between global peer-competitors into a futile exercise in conspicuous, immensely negative sum games, thanks in large part to survivable second-strike forces. A doctrine of enforced digital cooperation, supported by the elimination of any asymmetrical

information advantages of a challenging country, will further suppress spiraling escalation risks during international crises in our twenty-first-century digital civilization.

Notes

1 Sun Tzu, *The Art of War*, transl. Lionel Giles (1910), ch. 3, http://www.gutenberg.org/cache/epub/132/pg132.html.

2 This article can be viewed as a follow-up to the issues of destabilization in cyberspace discussed in Guy-Philippe Goldstein, "Cyber Weapons and International Stability," *Military and Strategic Affairs* 5, no. 2 (2013): 121-39.

3 See Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence* (Cambridge: Cambridge Studies in International Relations, 2000), pp. 296-301.

4 Paul K. Huth, *Extended Deterrence and the Prevention of War* (New Haven: Yale University Press, 1988), cited in Zagare and Kilgour, *Perfect Deterrence*, pp. 296-301.

5 For further details on this issue and introductory literature, see for example Goldstein, "Cyber Weapons and International Stability."

6 For further details on this issue and introductory literature, see for example Goldstein, "Cyber Weapons and International Stability."

7 William W. Kaufmann, *The Requirements of Deterrence* (Princeton: Center of International Studies, Princeton University, 1954). See also the discussion in Fred Kaplan, *The Wizards of Armageddon* (Stanford: Stanford University Press, 1983), pp. 193-200.

8 See discussion in Goldstein, "Cyber Weapons and International Stability" with reference to Schelling's definitions of red lines in Thomas C. Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966), p. 137.

9 See discussion in Goldstein, "Cyber Weapons and International Stability," with reference to the concept of "curve of credibility" in Carey B. Joynt and Percey E. Corbett, *Theory and Reality in World Politics* (Pittsburgh: University of Pittsburgh Press, 1978), p. 94-95.

10 See Michael N. Schmitt, ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare* (New York: Cambridge University Press, 2013), p. 88: "The international Group of Experts achieved no consensus as to whether non-destructive but severe cyber operations satisfy the intensity criterion." See also pp. 82-83, comments #14 and #15.

11 See Amit Sharma, "Cyber Wars: A Paradigm Shift from Means to Ends," in *The Virtual Battlefield: Perspective on Cyber Warfare*, eds. Christian Czosseck and Kenneth Geers (Amsterdam: IOS Press, 2009) for a discussion of cyber warfare in the context of effect-based warfare. More explicitly, the *Tallinn Manual* states that "a cyber operation constitutes a use of force when its scale and effects are comparable to non-cyber operations rising to the level of a use of force" (Rule 11) and that "a cyber attack is a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death

to persons or damage or destruction to objects" (Rule 30). The ensuing discussion does highlight that "'acts of violence' should not be understood as limited to activities that release kinetic force. This is well settled in the law of armed conflict. In this regard, note that chemical, biological or radiological attacks do not usually have a kinetic effect on their designated target, but it is universally agreed that they constitute attacks as a matter of law." See Michael N. Schmitt, ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare* (New York: Cambridge University Press, 2013). What matters is the direct effect on civilian populations or on properties, whatever the way – kinetic or not – these direct effects have been caused.

12 Paul M. Carpenter and William F. Andrews, "Effects-based Operations – Combat Proven," *Joint Force Quarterly* 52 (First Quarter, 2009): 78-81.

13 The international group of experts of the *Tallinn Manual* mentions the notion of "scale and effects" posited in the *Nicaragua* judgement of the International Court of Justice, "Case Concerning Military and Paramilitary Activities in and against Nicaragua" (Nicaragua v. United States of America), Judgement, *I.C.J. Reports* (1986), p. 14. See Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*, p. 45.

14 Christopher D. Baker, "Tolerance of International Espionage: A Functional Approach," *American University International Law Review* 19, no. 5 (2003): 1091-1113.

15 See for example Thomas C. Wingfield, "Legal Aspects of Offensive Information Operations in Space," *USAF Academy Journal of Legal Studies* 9 (1999): 140: "The lack of an international prohibition of espionage leaves decisionmakers with the usually acceptable liability of merely violating the target nation's domestic espionage law." See also Martin C. Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica: Rand Corporation, 2009), pp. 23-24. In the *Tallinn Manual*, the discussion of Rule 10 ("prohibition of threat or use of force") states that "not all cyber interference automatically violates the international law prohibition on intervention.... As noted by the Court in *Nicaragua*, 'intervention' is wrongful when it uses methods of coercion. It follows that cyber espionage and cyber exploitation lacking a coercive element do not *per se* violate the non-intervention principle." Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*.

16 See Part I, Chapter 2, Section 2 ("Self-defence") and Rule 32 ("Prohibition on attacking civilians") in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*.

17 See Charles Tilly, "War Making and State Making as Organized Crime," in *Bringing the State Back*, eds. Peter Evans, Dietrich Rueschemeyer, and Theda Skocpol (Cambridge: Cambridge University Press, 1985); see also Antonio Giustozzi, *The Art of Coercion: Armed Force in the Context of State Building* (CSRC Seminar, 2008).

18 See Agreement on Measures to Reduce the Risk of Outbreak of Nuclear War Between The United States of America and The Union of Soviet

Socialist Republics, September 30, 1971, http://www.state.gov/t/isn/4692.htm; Agreement Between The United States of America and The Union of Soviet Socialist Republics on Measures to Improve the U.S.A.-USSR Direct Communications Link, September 30, 1971, http://www.state.gov/t/isn/4787.htm, cited in Laura Grego, *A History of Anti-Satellite Programs* (UCS Global Security Programs, 2012).

19 See Karl Frederick Rauscher and Andrey Korotkov, *The Russia-US Bilateral on Critical Infrastructure Protection: Working Towards Rules for Governing Cyber Conflict* (New York: East-West Institute, 2011). See also Part II, Chapter 3 ("The law of armed conflict generally"), in particular Rule 20 ("Applicability of the law of armed conflict"), and Chapter 4 ("Conduct of hostilities"), in particular Rule 29 ("Civilians") and Section 3 ("Attacks against persons"), in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*.

20 Michael S. Lewis-Beck and Mary Stegmaier, "Economic Determinants of Electoral Outcomes," *Annual Review of Political Science* 3 (2000): 183-219.

21 Alan de Bromhead, Barry Eichengreen, and Kevin Hjortshøj O'Rourke, *Right Wing Political Extremism in the Great Depression*, Discussion Papers in Economic and Social History, No. 95 (Oxford: University of Oxford, 2012).

22 The *Tallinn Manual* defines as unlawful a cyber operation against the political independence of any state (Rule 10), Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare* (New York: Cambridge University Press, 2013).

23 See Rule 11 in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*, p. 51.

24 See Thomas C. Schelling, *Arms and Influence* (Yale University Press, 1966), pp.1-34 & pp.126-189

25 See Martin C. Libicki, "Cyberspace Is Not a Warfighting Domain," *I/S: A Journal of Law and Policy for the Information Society* 8, no. 2 (2012): 330.

26 See the *Quirin* case of 1942 on German saboteurs, with particular emphasis on saboteurs not wearing national emblems: "The spy who secretly and without uniform passes the military lines of a belligerent in time of war, seeking to gather military information and communicate it to the enemy, or an enemy combatant who without uniform comes secretly through the lines for the purpose of waging war by destruction of life or property, are familiar examples of belligerents who are generally deemed not to be entitled to the status of prisoners of war, but to be offenders against the law of war subject to trial and punishment by military tribunals." U.S. Supreme Court, *Ex Parte Quirin*, 317 U.S. 1 (1942). Unlawful combatants are nonetheless entitled to "to be treated with humanity and, in case of trial, shall not be deprived of the rights of fair and regular trial prescribed by the present Convention." See Geneva Convention Relative to the Protection of Civilian Persons in Time of War, August 12, 1949 (GCIV).

27 On "unprivileged belligerents" see comment #17 in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare,* p. 100.

28 See Ben Smith and Arabella Torp, "The Legal Basis for the Invasion of Afghanistan," *House of Commons, International Affairs and Defence Section*, February 26, 2010, pp. 4-5.

29 See for example Ashton B. Carter, Michael M. May, and William J. Perry, *The Day After – Action in the 24 Hours Following a Nuclear Blast in an American City*, Report based on Workshop (The Preventive Defense Project, Harvard and Stanford Universities, 2007), in particular "6. Retaliation and deterrence," pp. 15-17.

30 See Rule 11 in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*.

31 See Rule 13, comment #5, citing the *Nicaragua* judgment, para. 191, in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*, p. 55.

32 "All warfare is based on deception" quoted in Sun Tzu, *The Art of War*, transl. Samuel B. Griffith (New York and Oxford: Oxford University Press, 1963), p. 66.

33 Herman Kahn, *On Escalation* (London: Pall Mall Press Ltd., 1965).

34 The observation of effects should be complemented by a technical analysis of the malware itself. However, this could take too much time. For example, Stuxnet was identified by Virusblokada in June 2010 but only significantly analyzed by November 2010. See Nicolas Falliere, Liam O Murchu, and Eric Chien, *W32. Stuxnet Dossier* (Symantec, 2010). Hence the requirement for up-to-date information alerts from all military and civilian activity centers to a cyber intelligence collection point will make it possible to reinterpret cyber incident data points to form a coherent national picture for use by national security institutions.

35 See Rule 30, comment #5, in Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*, p. 106.

36 See John A. Warden III, "The Enemy as a System," *Airpower Journal* 9, no. 1 (1995).

37 See James N. Mattis, "USJFCOM Commander's Guidance for Effects-based Operations," *Joint Force Quarterly,* no. 51 (2008); see also criticism of USJFCOM decision by USAF officers in Paul M. Carpenter and William F. Andrews, "Effects-based Operations Combat Proven," *Joint Force Quarterly*, no. 52 (2009).

38 On the rise of corporate APIs, see Robin Vasan, "Business Process API-ification: The LEGO Promise Fulfilled," *GigaOm*, October 6, 2012, http://gigaom.com/2012/10/06/business-process-api-ification-the-lego-promise-fulfilled/ and Mark Boyd, "Getting C-Level Buy-In: Demonstrating the Business Value of APIs," *ProgrammableWeb*, September 11, 2013, http://blog.programmableweb.com/2013/09/11/getting-c-level-buy-in-demonstrating-the-business-value-of-apis/.

39 See discussion in Goldstein, "Cyber Weapons and International Stability," for main components of cyberspace.

40 Isaac Ben-Israel, *Philosophie du renseignement* (Paris : Editions de l'Eclat, 2004).
41 Ibid.
42 This example is directly inspired by the 1973 Yom Kippur War post-mortem analysis described in Ben-Israel, *Philosophie du renseignement*.
43 To reveal the identity of "Gerald," the mole working for the USSR, Smiley has a message sent to the head of the "Circus" that forces "Gerald" to seek an emergency meeting with his Soviet handler at an already identified safe house. This is the test that allows Smiley to identify "Gerald" while breaking into the safe house. In John Le Carré, *Tinker Taylor Soldier Spy* (London: Hodder & Stoughton, 1974).
44 See discussion in Goldstein, "Cyber Weapons and International Stability," for a comparison between the "digital" domain that establishes cyberspace and the "confidential information" domain that establishes the realm of traditional intelligence.
45 See Ben-Israel, *Philosophie du renseignement*.
46 See Richard Clarke and Robert K. Knake, *Cyberwar* (New York City: HarperCollins, 2010), pp. 249-54.
47 Ibid.
48 L. M. Hill, "The Two-Witness Rule in English Treason Trials: Some Comments on the Emergence of Procedural Law," *American Journal of Legal History* 12 (1968): 95-111.
49 See Glenn Shafer, "The Combination of Evidence," *International Journal of Intelligent Systems* I (1986): 155-79.
50 See the game theory analysis of the 1948 Berlin Crisis in Frank C. Zagare, *The Dynamics of Deterrence* (Chicago: University of Chicago Press, 1987), pp. 11-28.
51 See Thomas C. Schelling, *The Strategy of Conflict* (Cambridge: Harvard University Press, 1963), p. 69.
52 See Greg Rattray, Chris Evans, and Jason Healey, "American Security in the Cyber Commons," in *The Future of American Power in a Multipolar World*, eds. Abraham M. Denmark and James Mulvenon (Washington, D.C.: Center for a New American Security, 2010), pp. 151-72.
53 See Daniel L. Shapiro, "Negotiation Theory and Practice: Exploring Ideas to Aid Information Eduction," in *Educing Information*, eds. Robert A. Fein, Paul Lehner, and Bryan Vossekuil (Washington, D.C.: Intelligence Science Board, National Defense Intelligence College Press, 2006), pp. 267-80.
54 Quote from M.P. Rowe, "Negotiation Theory and Educing Information: Practical Concepts and Tools," in *Educing Information*, eds. Robert A. Fein, Paul Lehner, and Bryan Vossekuil (Washington, D.C.: Intelligence Science Board, National Defense Intelligence College Press, 2006), p. 295.
55 Stuxnet used compromised digital certificates from Taiwanese companies Realtek and JMicron. See Falliere, Murchu, and Chien, *W32. Stuxnet Dossier*.

56 David E. Sanger, "Obama Order Sped Up Wave of Cyberattacks against Iran," *New York Times*, June 1, 2012.

57 Starting for example with nations from the Technical Cooperation Program ("5 eyes nations") and/or other nations that have a history of cooperating closely in critical programs, in joint cyber operations for example or intelligence-sharing programs, as with the example of nations participating in Base Alliance against al-Qaeda. See Dana Priest, "Help from France Key in Covert Operations," *Washington Post*, July 3, 2005.

58 See Thomas C. Schelling, *Arms and Influence* (Yale University Press, 1966), p. 47.

59 See initial formulation in evolutionary biology, Leigh Van Valen, "A New Evolutionary Law," *Evolutionary Theory* 1 (1973): 1-30; see application to cyber arms race, Rattray, Evans, and Healy, "American Security in the Cyber Commons," the section "Adaptation and counter-adaptation," p. 154; see a first account by a practitioner in Kevin Mandia, "Cyber Threats and Ongoing Efforts to Protect the Nation," Permanent Select Committee on Intelligence, US House of Representatives, October 4, 2011, in particular the lack of deterrence or costs for the attacker.

60 Edward Rhodes, "Conventional Deterrence," *Comparative Strategy* 19, no. 3 (2000): 221-53, in particular 222-23.

61 Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984); see p. 13 for explanation of the "shadow of future" and p. 124 on "enlarging the shadow of the future" to promote cooperation.

62 Edward N. Luttwak, *Strategy: The Logic of War and Peace*, rev. and enlarged ed. (Cambridge: Belknap Press of Harvard University Press, 2001); see Chapter 11, "Nonstrategies," p.168-84.

63 Julian S. Corbett, *Some Principles of Maritime Strategy* (London: Longmans, Green & Co, 1911), p. 90: "Command of the Sea, therefore means nothing but the control of maritime communications, whether for commercial or military purposes."

64 See discussion about digital logos in Goldstein, "Cyber Weapons and International Stability."

65 Marc Andreessen, "Why Software is Eating the World," *Wall Street Journal*, August 20, 2011.

66 In 2010, the fastest supercomputer was the Cray Jaguar, running at 1.8 10^15FLOPS; see top500.org, November 2009-2010. Performances over one exaflop or 10^18 FLOPS could be available by 2020; see Agam Shah, "SGI, Intel Plan to Speed Supercomputers 500 Times by 2018," *Computerworld*, June 20, 2011.

67 Zettaflop capabilities (10^21) could achieve full-weather modelling – the accurate prediction of weather over a two week time span; see Erik P. DeBenedictis, "Reversible Logic for Supercomputing," in *Proceedings of the 2nd Conference on Computing Frontiers*, Sandia National Laboratories (2005), pp. 391-402.

68 Researchers have talked of an "Artificial Intelligence Winter" during at least two periods: in 1974-1980 and 1987-1993. See Jim Howe, "Artificial Intelligence at Edinburgh University: A Perspective," November 1994, School of Informatics, University of Edinburgh; Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach,* 2nd ed. (Upper Saddle River, New Jersey: Prentice Hall, 2003), p. 24.

69 By the mid-2000s, the mood had been reversed on AI and there was talk of a "spring" in AI. See for example John Markoff, "Behind Artificial Intelligence, a Squadron of Bright Real People," *New York Times*, October 14, 2005.

70 See the discussion on rules relaxation and lengthening the game in Zagare, *The Dynamics of Deterrence*, pp. 48-56.

71 See the discussion on the "shadow of the future" in Axelrod, *The Evolution of Cooperation*, p. 13.

72 See Goldstein, "Cyber Weapons and International Stability."

73 Zagare and Kilgour, *Perfect Deterrence*, pp. 293-96.