NAVAL POSTGRADUATE SCHOOL ESSAY CONTEST—1ST PRIZE WINNER

Sponsored by NPS Foundation

SHUTTERSTOCK

AUTOMATION WILL CHANGE SEA POWER

BY JOHN ARQUILLA AND PETER DENNING

Artificial intelligence is changing the balance of power, and a networked fleet is the best way for the Navy to maintain maritime superiority.

Artificial intelligence (AI) will profoundly influence sea power in the coming years. It will change the game for organization, doctrine, policy, and operations. And, most important, it will change the areas in which naval personnel are expected to become knowledgeable and proficient.

Since World War II, carriers have been the U.S. Navy's capital ships. They have been the highest expression of sea power in the industrial era, and they remain so today. The rapid rise of digitization and networking, however, signal the beginning of a new era that may take the Navy beyond the carrier's primacy. When that change comes, the next capital ship is likely to be virtual: a swarm of platforms, including carriers, plus countless digitally controlled entities—some remotely controlled, others fully autonomous.

THE RELIABILITY CONUNDRUM

In the 19th century, mathematician and inventor Charles Babbage convinced the British Navy that using a machine to calculate navigation tables could reduce shipwrecks by eliminating error-prone hand calculations. Unable to get a machine of gears and levers to work reliably, he invented another machine of simpler design: a mechanical prototype of a modern programmable computer. He died before completing that machine.

Belief in machine reliability was a key part of the justification for the first electronic automatic computers in the 1940s. It has become clear, however, that computing machines are so complex that no one can be sure they are working correctly. How to organize computers and software for reliability remains a central question of computer science, and its resolution is essential for the implementation of automated systems into military and naval operations.

Achieving high reliability will be a major challenge as the Navy is drawn deeper into automated systems—so-called intelligent systems. To maximize their chances for crafting reliable systems, designers must be conservative, keeping their designs as simple as possible.

There is another source of unreliability for AI machines. Despite their ambitious title, they are profoundly "unintelligent." Most of the intelligence attributed to them arises from the speed at which they operate. These machines get their speed because their operations depend solely on "context-free" rule sets. Humans are good at sensing context; machines are not. Humans care about outcomes; machines do not. But the rise of neural networks has led some to argue that machines are getting close to being "intelligent."

THE COMING OF NEURAL NETWORKS

Big data, increasingly important to decision support, is being paired with a new generation of learning machines called neural networks, which reflect a technological tipping point in the analysis of large volumes of data. They are powerful, yet they present some troubling behaviors.

Neural networks are machines that take inputs X and generate outputs Y. They are built of electronic neurons in a dense network of connections, loosely mirroring structural principles of the human brain. Each neuron can be depicted in a "0" or "1" state. When total input of a neuron from all incoming connections exceeds its built-in threshold, the neuron "fires" and enters the "1" state. That "1" is fed to connected neurons, causing some of them to fire in the same way. An input to the whole network triggers a cascade of neuron firings until the network settles. The settled "1" and "0" values of the neurons designated as outputs become the network's output.

These networks are organized as a series of layers that process stimuli through successive levels of abstraction until they become responses. Because they have dozens of levels, they are called "deep-learning networks."

A neural network is not "programmed." Instead it is "trained" by an algorithm. The process of training a neural network can take several days; however, once a network is trained, it is exceedingly fast. Trained networks generally can produce a response within milliseconds of receiving a stimulus.

When training is done, the result is a matrix that specifies the weights of connections from any given neuron to all the others. These matrices can contain billions of entries. The question of explaining how a network reached its conclusion thus cannot be answered by an examination of its connection matrix. It would be like trying to determine the dark motives of criminals by examining their brain scans.

Inscrutability—the inability to explain how or why a network reached a particular conclusion—is one of two big open problems with neural networks. The other problem is fragility—the network response can be unpredictable beyond its training data set. Networks also can be confused and attacked by presenting slightly noisy versions of trained inputs.[1]

THE ALPHAGO INFLECTION POINT

AlphaGo is a deep-learning system built by the DeepMind subsidiary of Google to play the ancient strategy game of Go, or Weiqi. AlphaGo changed the landscape of machines that play games against humans. In its 2016 debut, it beat Grandmaster Lee Sedol of South Korea. *Science* magazine declared AlphaGo one of its "breakthroughs of the year." The designers thereafter called it AlphaZero because they were able to use the same basic machine to also learn chess and shogi.[2]

One of the AlphaZero breakthroughs was in the method of training. The classical neural network takes a large number of X-Y pairs and teaches them to the network—usually requiring analysis of past recorded games. Instead, AlphaZero trained by playing against itself with no input other than the rules of chess, shogi, or Go. It learned to play grandmaster chess in 9 hours, shogi in 12 hours, and Go in 13 days.

DEEP LEARNING AND MILITARY STRATEGY

There is a strong analogy between Go and military war games. The goal is to place assets and make tactical moves so that the game produces a win by achieving greater control of the battlespace—in naval parlance, "command of the sea." While the current rules of wargaming are complex and the tactical moves subtle, it is not hard to imagine a version of AlphaZero that could learn to play and win a war game in a few days. The implications for future battle management could prove enormous. The Navy should invest in such a deep-learning capability, as it may become a key to future naval mastery.

Deep-learning algorithms also are being explored for their ability to enable automation of decisions about weapon use. Yet, the uncertainties around fragility and inscrutability raise the possibility that automated systems accidentally could trigger escalation of conflict before human operators could intervene—for example between automated Chinese and U.S. fleets in the East or South China seas.

THE BOEING COMPANY

U.S. NAVY

U.S. NAVY

The Navy must bet big on the network as its next "capital ship"—a swarm of platforms, including carriers and digitally controlled entities, some fully autonomous.

Critics have said AlphaZero is good only at learning to play games. However, a future version may learn how to win war games—and translate this capability into winning real military and naval campaigns. Such a system could augment human strategic planning, as it would be operating in the absence of the cognitive and motivated biases that plague human decision-making.

A SWARM OF PLATFORMS

Despite its risks, automation is key to dealing with an emerging threat to modern navies: swarms.

Swarms are large networks of small coordinated entities, each able to inflict some damage and together capable of enormous damage. The entities can be any combination of autonomous boats, ships, drones, human kamikazes, or even cyber codes. The goal of swarming is to overwhelm defenses, fatally exposing the target.

Technology that enables swarming already exists. The Israeli military, for example, has built small drones ("Harpies") that can carry deadly explosives and navigate autonomously, with the intent of diving on a target. A swarm of such drones could overwhelm normal ship defenses.

The best answer to this threat is for the Navy to develop a "reverse swarm" able to attack and neutralize an incoming swarm. This is a major design challenge. Drones in the defensive swarm will require jam-proof, interdrone communications and a low-bandwidth, jam-proof connection to the ground controller. If the bandwidth is jammed, or the battle-management supercomputer is degraded or disrupted, the swarm will fall out of the air or become chaotic.

To solve this problem, individual drones must be equipped with their own computers and control programs. The connection to the ground controller must work at normal wireless channel speeds. Designing a reverse swarm is a challenge for autonomous software. Then-professor Tim Chung met this initial challenge in a demonstration of a 50-drone swarm at a Naval Postgraduate School test site in September 2015.

The capital ship protected by a reverse swarm no longer would be a juicy target for the attacking swarm. But what if the concept of the capital ship itself evolved into an enormous swarm of land, sea, air, and cyber elements coordinated by a network of autonomous controls? Indeed, the dominant naval tactic of the future is likely to be swarming, which is sure to spread to many nations—especially because U.S. naval preeminence makes it hard for competitors to employ traditional battle doctrines with much hope of success.

To counter the coming swarms, naval automated systems of the future will have to be good at detection and rapid action. In fact, they will have to *be* swarms—automated, too, because the pace of operations will rule out the possibility that humans alone can meet this threat.

HUMAN-MACHINE TEAMS

The naval future is not one of just humans or machines; it will feature the blending of humans and machines. The only viable path is one in which humans and machines work together, combining the speed of machines with the wisdom of humans.

The skillful blending of automata and humans may hold the key to naval mastery in the 21st century. Tomorrow's navies may go into battle with humans and computers fighting side by side, the former providing insight and judgment, the latter lightning-swift calculations guiding choices in fast-moving, uncertain situations.

At the fleet level, a thoughtful network combining humans and automata may herald an era in naval affairs in which the capital ship is no longer a particular vessel type. Instead, the essence of sea power may become the network itself. Thus the "big thing" in naval affairs may actually be a swarm of little things. This echoes the thinking of the late Vice Admiral Arthur K. Cebrowski in his notion of "network-centric warfare."[3] But whereas he thought in terms of creating separate grids for sensors and shooters, automation will fuse them—making them all faster and more accurate.

The network-as-capital-ship concept, combined with a battle doctrine of swarming, is likely to serve well in a range of future naval warfare scenarios. For example, it makes good sense to send in a swarm to confront maritime challenges in the East or South China seas, where the People's Liberation Army Navy is trying to build a network with smart weapons. Similarly, Iranian coastal forces have a swarming doctrine for operations in the Arabian Gulf, which will require a response from a U.S. swarm. The U.S. Navy should embrace automation in the form of networked, swarming human-machine teams.

Naval history is replete with examples of "broad preparation"—investing in a bit of everything instead of making a few "big bets." It's time to place a big bet on the network as the capital ship. The network, encompassing a range of platforms—including carriers—will be robust on defense and capable of lightning-swift offensive action.

In his *Design for Maintaining Maritime Superiority Version 2.0*, Chief of Naval Operations Admiral John Richardson uses the term "networked fleet" to address this same concept. Indeed, the network is ideally suited to the emerging Navy notion of distributed lethality.[4] For nothing is more distributed than a network, nor more lethal than a well-armed one.

THE ROLE OF EDUCATION

Networking, swarming, and machine learning involve powerful tools and sophisticated practices, requiring skilled engineers, mathematicians, statisticians, and others to design, build, and sustain them.

The Naval Postgraduate School offers a program for computer science students to compete in annual "Capture the Flag" contests in which teams try to invade rival networks and steal critical data without being detected. Almost all attacks are conducted at the lowest levels of network protocols and operating system kernels. Without deep knowledge of these technologies, teams cannot win, because they will be blindsided by attacks they do not understand and cannot stop. Although they may have defensive tools to help them detect and block attacks, teams must be prepared to build new tools *on the spot*.

The same will be true for teams engaged with networked, machine-learning technologies. They must know their terrain intimately and be able to build new defensive and attack strategies and tools on the fly.

To better prepare for a more automated fleet, naval education should emphasize the technologies of computer science, information science, networking, machine learning, sensor-data acquisition and analysis, and advanced statistics. In other words, deep learning by humans.

ARTIFICIAL INTELLIGENCE IS HERE

Historically, militaries have focused enormous energy on training to act in highly coordinated, automaton-like ways under the strains of combat. Now, the ambition of artificial intelligence is to pour insight and human-like judgment into machines that will have the potential, when reliable enough, to team with humans in many combat situations.

In today's state of technological play, the goal of artificial intelligence—creating human-quality insight in machines—remains far off. But automation and machine learning are here, and can be employed, particularly in human-machine teams.

[1] One study showed that a network designed to read road signs was confused into saying it saw a speed limit sign when presented with an image of a stop sign that had bird droppings splattered on it. How can you trust a driverless car that cannot accurately read road signs?

[2] David Silver et al., "A General Reinforcement Learning Algorithm that Can Master Chess, Shogi, and Go through Self-Play," *Science* 362, no. 6,419 (December 2018): 1,140–44.

[3] VADM Arthur K. Cebrowski, USN, and John H. Garstka, "Network-Centric Warfare: Its Origin and Future," U.S. Naval Institute *Proceedings* 124, no. 1 (January 1998).

[4] See VADM Thomas Rowden, RADM Peter Gumataotao, and RADM Peter Fanta, USN, "Distributed Lethality," U.S. Naval Institute *Proceedings* 141, no. 1 (January 2015).

■ **DR. ARQUILLA AND DR. DENNING** are distinguished professors and department chairs at the U.S. Naval Postgraduate School. Dr. Arquilla is in defense analysis and has authored several books and articles in military and security affairs. Dr. Denning is in computer science and has authored many seminal works over the past 50 years.