Super Computers in Warfare -- A Million Mistakes a Second

by Paul Scharre via lynx - Foreign Policy *Monday, Sep 17 2018, 8:13am*
international / prose / post

> Ultrafast computing is critical to modern warfare. But it also ensures a lot could go very wrong, very quickly.

Militaries around the globe are racing to build ever more autonomous drones, missiles, and cyberweapons. Greater autonomy allows for faster reactions on the battlefield, an advantage that is as powerful today as it was 2,500 years ago when Sun Tzu wrote, “Speed is the essence of war.” Today’s intelligent machines can react at superhuman speeds. Modern Chinese military academics have speculated about a coming “battlefield singularity,” in which the pace of combat eclipses human decision-making.

The consequences of humans ceding effective control over what happens in war would be profound and the effects potentially catastrophic. While the competitive advantages to be gained from letting machines run the battlefield are clear, the risks would be grave: Accidents could cause conflicts to spiral out of control.

Consider what has already happened with stock markets, where computers use algorithms to make decisions so quickly that microseconds make a difference of millions of dollars. Such trading has made brokers huge amounts of money—but has also produced extreme flash crashes that can send markets tumbling in minutes. Regulators have managed these risks by installing circuit breakers that can take a stock offline if the price moves too quickly, but battlefields lack these fail-safes. Flash crashes are bad enough; a flash war would be downright disastrous.

Humans have already ceded control to machines in certain military domains. At least 30 countries—with Israel, Russia, and the United States leading the pack—employ human-supervised autonomous weapons to defend bases, vehicles, and ships. These weapons systems, such as the ship-based Aegis combat system, can detect incoming rockets and missiles and, if human supervisors do nothing, respond on their own by firing to eliminate the threat. Such automated responses allow the systems to defend against what are known as saturation attacks, in which salvos of missiles or rockets are launched at a target with such little notice that they could overwhelm human operators.

For the time being, autonomous weapons such as these are used purely to protect human-occupied installations or vehicles. Humans supervise the weapons’ operation in real time and can intervene if necessary. Future autonomous weapons could lack these safeguards, however. A number of advanced militaries—including those of China, France, Israel, Russia, the United Kingdom, and the United States—are currently developing stealth combat drones intended to penetrate an adversary’s airspace. Once deep behind enemy lines, these drones might find their communications jammed, so they’re being designed to ensure they can continue to operate on their own.

Most countries have not explained how their drones will operate under such circumstances and what rules of engagement they will follow.

Countries could require their drones to get human authorization before launching any attacks. Doing so would allow the drones to bomb preapproved fixed targets but would require them to report back

and get permission before attacking any newly discovered quarries. Such an approach sounds good in theory, but the problem is that these days many high-priority targets, such as air defense systems and ballistic missile launchers, are highly mobile. This mobility will increasingly tempt military planners to delegate lethal decision-making authority to machines, since doing so could give them an edge in reaction time.

No battlefield is static, and the ability to rapidly react to a dynamic environment is critical to mission success—whether in the air, on the ground, or in cyberspace. Air combat strategists call this the OODA (Observe, Orient, Decide, and Act) loop in dogfighting. In the OODA loop paradigm of combat, pilots win dogfights not simply because they enjoy the best hardware but because they assess and react to their situations faster than their opponents, although better sensors and maneuverability might help shorten reaction times. Since machines can react faster than humans, automation will offer tremendous advantages in this competition. That means that the same competitive pressures that led to the creation of systems such as the Aegis could soon be introduced on a wider scale.

A military that fully integrated its autonomous systems could always stay one step ahead of its enemy in combat and present a constantly shifting threat. As Gen. Paul Selva, the vice chairman of the U.S. Joint Chiefs of Staff, told the Senate Armed Services Committee in July 2017, "It is very compelling when one looks at the capabilities that artificial intelligence can bring to the speed and accuracy of command and control and the capabilities that advanced robotics might bring to a complex battlespace, particularly machine-to-machine interaction in space and cyberspace, where speed is of the essence."

Yet speed is not an unadulterated good. Forces that react to the enemy so quickly that their own commander does not understand what is happening could risk a breakdown in command and control, a problem that military leaders have struggled with for millennia. Today, email and chat messaging have replaced horses and flags, but the fundamental problem persists. Militaries counter this inherent friction between orders from above and the reality on the ground with a concept known as commander's intent: succinct goal-oriented statements issued to subordinates that explain the desired goal of a particular mission, and thus ensure that they stick to the general plan, but also allow them the flexibility to adapt to events on the ground. Such statements prevent forces from becoming too predictable and give subordinates the freedom to overcome obstacles in novel ways.

The problem is that automated systems—at least those using current technology—tend to be brittle.

Machines are good at handling routine tasks under predictable circumstances, such as flying a commercial airliner. But automation can sometimes fail dramatically in new situations, which is a major reason why self-driving cars, which must contend with extremely dynamic and uncontrolled environments, have proved so much harder to develop than self-flying planes.

Learning systems, which focus on a set of rules for processing and incorporating data into behavior instead of strict mandates, exhibit more flexibility but are still limited by the quality of information. Machine learning can fail if the real world proves different from training models—a major problem in the military context, where adversaries are unlikely to offer easy access to their tactics and hardware. Human intelligence is robust and adaptable in ways that machine intelligence is not—yet. The most effective militaries will thus be those that find ways to successfully marry human and machine intelligence into joint cognitive systems—an approach that defense analysts call centaur warfighting.

Despite humans' advantages in decision-making, an arms race in speed may slowly push humans out of the OODA loop. Militaries are unlikely to knowingly field weapons they cannot control, but war is

a hazardous environment and requires balancing competing risks. Faced with the choice of falling behind an adversary or deploying a new and not yet fully tested weapon, militaries are likely to do what they must to keep pace with their enemies.

As mentioned above, automated stock trading provides a useful window into the perils of this dynamic. In 2010, the Dow Jones Industrial Average lost nearly 10 percent of its value in just minutes. The cause? A sudden shift in market prices driven in part by automated trading, or what's come to be known as a flash crash.

In the last decade, financial markets have started to suffer such crashes, or at least miniature versions of them, on a regular basis. The circuit breakers installed by regulators to pull a stock offline can't prevent incidents from occurring, but they can stop flash crashes from spiraling out of control. Circuit breakers are still regularly tripped, though, and on Aug. 24, 2015, more than 1,200 of them went off across multiple exchanges after China suddenly devalued the yuan.

In competitive environments such as stock markets and battlefields, unexpected interactions between algorithms are natural. The causes of the 2010 flash crash are still disputed. In all likelihood, there were a range of causes, including an automated sell algorithm interacting with extreme market volatility, exacerbated by high-frequency trading and deliberate spoofing of trading algorithms. To prevent the military equivalent of such crises, in which autonomous weapons become trapped in a cascade of escalating engagements, countries will have to balance advantages in speed with the risk of accidents. Yet growing competition will make that balancing act ever more difficult. In 2016, Robert Work, then-U.S. deputy defense secretary, colorfully summed up the problem this way: "If our competitors go to Terminators, and it turns out the Terminators are able to make decisions faster, even if they're bad, how would we respond?"

Again, stock markets show how important it is that countries answer this question in the right way. In 2012, an algorithm-based trading accident nearly bankrupted the high-frequency trading firm Knight Capital Group. A glitch in a routine software update caused the firm's computers to start executing a lightning-fast series of erroneous trades, worth $2.6 million a second. By the time the company reined in its runaway algorithm, its machines had executed 4 million trades with a net loss of $460 million—more than the company's entire assets. To give a sense of scale: In 1994, it took more than two years of deception for the rogue trader Nick Leeson to bankrupt Barings Bank. In what came to be known as the Knightmare on Wall Street, a machine managed to inflict the same damage in 45 minutes. In that case, of course, although a company was destroyed, no lives were lost. A runaway autonomous weapon would be far more dangerous.

Real-world accidents with existing highly automated weapons point to these dangers. During the initial invasion of Iraq in 2003, the U.S. Army's Patriot air defense system accidentally shot down two friendly aircraft, killing three allied service members.

The first fratricide was due to a confluence of factors: a known flaw that caused the radar to mischaracterize a descending plane as a missile, outdated equipment, and human error.

The second blue-on-blue incident was due to a situation that had never arisen before. In the hectic march to Baghdad, Patriot operators deployed their radars in a nonstandard configuration likely resulting in electromagnetic interference between the radars that caused a "ghost track"—a signal on the radars of a missile that wasn't there. The missile battery was in automatic mode and fired on the ghost track, and no one overruled it. A U.S. Navy F-18 fighter jet just happened to be in the wrong place at the wrong time. Both incidents were flukes caused by unique circumstances—but also statistically inevitable ones. Coalition aircraft flew 41,000 sorties in the initial phases of the Iraq

War, and with more than 60 allied Patriot batteries in the area, there were millions of possible interactions, seriously raising the risk for even low-probability accidents.

Richard Danzig, a former U.S. secretary of the Navy, has argued that bureaucracies actually systematically underestimate the risk of accidents posed by their own weapons. It's also a problem that it's nearly impossible to fully test a system's actual performance outside of war. In the Iraq invasion, these accidents had tragic consequences but did not alter the course of the war. Accidents with fully autonomous weapons where humans cannot intervene could have much worse results, causing large-scale fratricide, civilian casualties, or even unintended attacks on adversaries.

Attempts at arms control go back to antiquity, from the Bible's prohibition on wanton environmental destruction in Deuteronomy to the Indian Laws of Manu that forbade barbed, poisoned, or concealed weapons. In the intervening centuries, some efforts to ban or regulate certain weapons have succeeded, such as chemical or biological weapons, blinding lasers, land mines, cluster munitions, using the environment as a weapon, placing weapons in space, or certain delivery mechanisms or deployment postures of nuclear weapons. Many other attempts at arms control have failed, from the papal decrees denouncing the use of the crossbow in the Middle Ages to 20th-century attempts to ban aerial attacks on cities, regulate submarine warfare, or eliminate nuclear weapons. The United Nations began a series of meetings in 2014 to discuss the perils of autonomous weapons. But so far the progress has been far slower than the pace of technological advances.

Despite that lack of success, a growing number of voices have begun calling for a ban on autonomous weapons. Since 2013, 76 nongovernmental organizations across 32 countries have joined a global Campaign to Stop Killer Robots. To date, nearly 4,000 artificial intelligence and robotics researchers have signed an open letter calling for a ban. More than 25 national governments have said they endorse a ban, although none of them are major military powers or robotics developers. But such measures only tend to succeed when the weapons in question are of marginal value, are widely seen as especially horrific or destabilizing, are possessed by only a few actors, are clearly distinguished from other weapons, and can be easily inspected to verify disarmament. None of these conditions applies to autonomous weapons.

Even if all countries agreed on the need to restrain this class of arms, the fear of what others might be doing and the inability to verify disarmament could still spark an arms race. Less ambitious regulations could fare better, such as a narrow ban on anti-personnel autonomous weapons, a set of rules for interactions between autonomous weapons, or a broad principle of human involvement in lethal force. While such modest efforts might mitigate some risks, however, they would leave countries free to develop many types of autonomous weapons that could still lead to widespread harm.

Humanity stands at the threshold of a new era in war, in which machines will make life-or-death decisions at speeds too fast for human comprehension. The risks of such a world are real and profound. Autonomous weapons could lead to accidental death and destruction at catastrophic scales in an instant. The unrestrained pursuit of fully autonomous weapons could lead to a future where humans cede control over what happens on the battlefield, but the critical decisions about how this technology is used still rest in human hands.

https://foreignpolicy.com/2018/09/12/a-million-mistakes-a-second-future-of-war/

Inverse Times Open Publishing. http://inversetimes.lingama.net/news/story-292.html