Algorithmic Warfare

Applying Artificial Intelligence to Warfighting

Peter Layton

Algorithmic Warfare

Applying Artificial Intelligence to Warfighting

Peter Layton

Disclaimer

The views expressed in this work are those of the author and do not necessarily reflect the official policy or position of the Department of Defence, the Royal Australian Air Force or the Government of Australia. The Commonwealth of Australia will not be legally responsible in contract, tort or otherwise, for any statements made in this document.

A catalogue record for this book is available from the National Library of Australia.

ISBN: 978192562267

Published and distributed by:

Air Power Development Centre
F3-G, Department of Defence
PO Box 7932
CANBERRA BC 2610
AUSTRALIA

Telephone: + 61 2 6128 7041
Facsimile: + 61 2 6128 7053
Email: airpower@defence.gov.au
Website: www.airforce.gov.au/airpower

Foreword

Since the first days of air power, the pace of technological innovation has been relentless. This rapid pace continues with a range of technologies emerging that threaten to disrupt and overturn conventional thinking on warfighting and traditional force structures. Last year, *Beyond the Planned Air Force* determined that disruptive technology included autonomous weapons, uninhabited systems, artificial intelligence, smart algorithms and 'big data' analysis.

Combined employment of some of these disruptors can be described as algorithmic warfare, which can be characterised in two ways: first, the disruptors will pervade future air forces influencing their roles and missions; and second, warfare will evolve beyond network-centric conceptions. The conduct of future wars will increasingly change as intelligent machines for the first time begin to assist and counsel humans in warfighting. Not surprisingly, it is becoming apparent that the advice provided by algorithms will significantly shape future military judgments.

While such a future offers tantalising prospects for making air forces more strategically and operationally effective, many questions arise about the ethical, moral and legal aspects of the employment of air power. Although many of these questions are legitimate, they sometimes consider only part of the story. Some of the worries that emerge may be addressed by carefully examining what evolving intelligent machine technologies can and cannot do.

These technologies are simultaneously powerful and brittle, brilliant and childlike, dazzling and incomprehensible. Indeed, such seeming contradictions are probably best described by Moravec's paradox, which holds that intelligent machines find the difficult things easy and the easy things difficult.

Dr. Layton's paper explores the concept of algorithmic warfare and, in doing so, discusses the technology involved, human-machine teaming matters, diverse warfighting aspects and new employment strategies. While this challenging new area will disrupt and influence everything Air Force does, the process has evidently already begun. Algorithmic warfare has arrived and deserves considerable thought.

GPCAPT Andrew Gilbert
DAPDC
March 2018

About the Author

Dr Peter Layton, PhD is a RAAF Reserve Group Captain and a Visiting Fellow at the Griffith Asia Institute, Griffith University. He has extensive aviation and defence experience and, for his work at the Pentagon on force structure matters, was awarded the US Secretary of Defense's Exceptional Public Service Medal. He has a doctorate from the University of New South Wales on grand strategy and has taught on the topic at the Eisenhower College, US National Defence University. For his academic work, he was awarded a Fellowship to the European University Institute, Fiesole, Italy. He is the author of the book *Grand Strategy*.

Contents

1.

Introduction

We live in a time of rapid disruptive technological change impacting all aspects of our lives and societies. This is particularly evident in the field of information technology (IT) where the defence domain is trying hard to keep up with the ongoing remarkable developments in the commercial realm. This is a sharp difference to 30 years ago during the Cold War when the military led technological development, took a calculated approach to such change and carefully managed disruptions.

The latest disruptive technologies emerging from the commercial realm that concern defence include advanced computing, 'big data' analytics, artificial intelligence (AI), autonomy and robotics. In this, the 2018 US National Defense Strategy reflects much contemporary strategic thinking in declaring that: 'The drive to develop [these] new technologies is relentless, expanding to more actors with lower barriers of entry, and moving at accelerating speed….' The tone is worried as these are 'the very technologies' determined necessary to 'be able to fight and win the wars of the future.'

The five 'technologies' noted in the preceding paragraph are not strictly comparable. Some are technologies but some are capabilities that other technologies might give future warfighters. Indeed, certain capabilities, like autonomy, are already operational and have been for decades. This conflation of technologies and military capabilities tends to hinder current defence debates not sharpen them. The new

term, 'algorithmic warfare', may instead be more useful in describing and discussing the latest technology-pushed warfighting concepts.

Algorithms are the sequence of instructions and rules that machines use to solve problems. They transform inputs to outputs and as such are the crucial conceptual and technical foundation stone of modern IT and the new intelligent machines. Algorithms may also become the conceptual and technical foundation stone of future warfighting.

The term's present use emanates from the US DoD's Project Maven. This project aims to use AI systems trained using advanced computing techniques to autonomously analyse big data. Immediately apparent is that this undertaking results from integrating several differing technologies and functions. Maven's success depends on the performance of the algorithms used.

Regardless of the terminology used, the practical impact of intelligent machines on our established force structures and warfighting is uncertain. Will algorithmic warfare mean that we can do things better or, in contrast, that we can do better things? Will our current warfighting styles simply evolve or instead be revolutionised by incorporating intelligent machines? Will intelligent machines transform our current force structure models? This paper explores such questions and more.

Chapter 2 delves into the technical basis of algorithmic warfare, focusing on machines that learn and have emergent properties; these machines are intrinsically quite different to our current programmable machines. Chapter 3 looks at the practical matters that shape algorithmic warfare implementation and the apparent, perhaps unexpected importance of the human-machine interface to winning wars. Chapter 4 examines algorithmic warfare to discuss both how this might enhance our current warfighting concepts and conversely how it might radically transform them. Algorithmic warfare may replace our current network-centric warfare concepts.

Chapter 5 looks at Chinese and Russian thinking; both are using unique applications of algorithmic warfare for domestic societal management and disruption. Chapter 6 considers ethics and pertinent law of armed conflict issues.

Importantly, this paper tries to stay practically focused on the application of emerging and probable intelligent machine technologies to warfighting. For commercial reasons as much as anything, fictional books and movies remain fascinated by the notions of robot soldiers fighting a final battle against the human race. The technologies involved in those forums remain distant, as this paper will persuade you. However, while *Terminator* cyborgs and *Cyberdyne's Skynet* are imaginary, today's smart-phone and internet search engines already use intelligent machine technologies, the Chinese Government's Skynet intelligent surveillance system is operational and Project Maven will deliver its AI-powered analysis system to the USAF later this year. Intelligent machines have arrived. Algorithmic warfare is now very real and demands deep professional consideration.

2.

The Technology of Algorithmic Warfare

The ability to undertake algorithmic warfare rests on computing technology advances in three main areas. The first involves the several decades of exponential growth in computer-processing power that has allowed sharp improvements in implementing machine-learning techniques. The second involves the sudden growth in 'big data'; very large, often automated, mined and created datasets suitable to train learning-capable machines. The third involves the steady evolution of cloud technology so that computers can readily access off-board processing and data resourcing to solve problems.

From considering these areas, it is apparent that algorithmic warfare is not a discrete technology such as directed energy weapons or hypersonics. Instead, the concept's technologies will have a broad, all-pervasive effect, progressively becoming omnipresent in warfighting. For the first time, military machines are becoming intelligent, potentially making those defence forces that successfully embrace them more effective and efficient. Such smart machines though do have distinct limitations that need to be understood and which can be exploited.

Intelligent Machines

Military machines have long been automated. A Harpoon anti-ship missile of the 1970s for example could accurately fly for several minutes at high-subsonic speeds at very low altitudes, operate its radar, analyse the radar picture, determine a particular ship to target, and then fly a complex attack profile. These complicated machines, like our modern desktop computers, were programmable. Such machines analyse structured data, use carefully designed logic flows and take pre-defined actions to complete tightly specified tasks. Such systems are broadly considered as deterministic, that is for a given input the output will always be the same unless there is a hardware failure or software glitch. They can be very powerful but are inherently rigid.

Your mobile phone and online search engines contain the early beginnings of something quite different. IBM calls these 'cognitive' machines. Instead of being pre-programmed, they learn from their interactions with humans and the environment to continually update their internal model of the world. They are probabilistic, that is they provide not simply answers to numerical problems but confidence-weighted responses together with supporting evidence. They can identify patterns and provide fresh insights when analysing unstructured data; an important issue given some 80 per cent of the world's data is unstructured. Cognitive machines though do not give the same result every time but rather provide 'best-guess' answers across a range. Cognitive machines are intelligent, at least compared to the preceding programmable ones.

Cognitive machines are enabled by high processing power. Such machines need specialised chips specifically designed for the particular purposes intended. The affordability of these unique chips then depends on their being placed into high-volume production, which in turn, depends on the chip's commercial importance. The

recent sharp advances in chip design have been driven mainly by the demands of speech and face recognition systems, and autonomous cars. When mass-produced, these new chips will rapidly improve commercial computer applications and by lowering costs will allow intelligent machines to be widely adopted by armed forces. The first, *Nvidia's Xavier* chip, is designed specifically for autonomous vehicles and combines high processing power, reliability and energy-efficiency. In the future, quantum computing may further revolutionise processing power with major intelligent machine impacts.

The new chips are important to machine learning, the key technical advance driving recent intelligent machine development. Instead of programming the computer with each individual step needed to solve a problem, machine learning uses learning algorithms that make inferences from the data provided. Crucially, rather than the computer programmers, the learning algorithms create the rules that intelligent machines use. This means these computers can be used for complicated tasks unable to be manually programmed, such as face recognition across *Facebook's* almost two billion users. With different training data, the same learning algorithm can be used to generate new rules and instructions appropriate to new tasks. In general, the more data used to train the learning algorithm the better the rules and instructions devised.

There are two principal machine-learning methods: supervised and unsupervised. In the former, the learning algorithms are given labelled data. For example, photos of naval ships labelled 'warships' might be fed through the algorithm so it can devise the rules that the intelligent machine could use to classify such pictures again in the future when fed large photo datasets. Supervised learning though requires people to tag and categorise the data; this can be a time-consuming, error-prone task

In contrast, unsupervised learning uses unlabelled data. The learning algorithms identify patterns for themselves in the data they are is fed. This approach is broadly like how humans learn; they observe the world, determine how objects are related and, from that, build an internal model about how the world works. With unsupervised learning though, it is unsure what data associations the learning algorithms make. While the process entails using known statistical methods to assess the data, the only indicator is how well the intelligent machine performs the set task.

The unsupervised learning method involves several techniques. In reinforcement learning, the learning algorithm interacts with a dynamic environment that provides feedback about rewards and punishments for doing tasks correctly. Most famously, the *AlphaGo* intelligent machine, trained to use reinforcement learning, recently defeated the world Go champion. The strategy game of Go has long been considered a particularly difficult challenge for intelligent machines to master. This somewhat startling intelligent machine success appears to have significantly influenced Chinese military thinkers about the need to embrace intelligent machines and algorithmic warfare. In the case of *AlphaGo*, the machine trained initially playing against humans and then against itself, becoming progressively better each time.

Similar in concept are Generative Adversarial Networks that compete against each other to improve their performance. Each network tries to trick the other by making it increasingly difficult for the other to correctly complete its task. This technique allows smaller datasets to be used for training because the opponent can generate increasingly realistic, but false, data against which to train.

Deep learning is the current state-of-the-art in machine learning. While still using learning algorithms, these are stacked in layers to create an artificial 'neural network'. When this deep neural network is fed data, it determines the features to use for data classification

itself. This means that such networks can improve their performance over time as they continually train themselves on the new data received while operating. In contrast, conventional machine learning continues to rely on the training received from the original dataset and classification features so derived.

Apple smart phones use a trained deep neural network to activate the *Siri* voice recognition application when the words 'Hey Siri' are spoken. The 'Hey Siri' detector, a 'simple' probabilistic intelligent machine, makes its best guess at what has been said by accounting for the differences in each voice and the constantly changing background environments. *AlphaGo* also used deep neural networks as it undertook its reinforcement learning playing others and itself.

With such advanced computing techniques, intelligent machines can now interpret information and solve specific problems more consistently than humans or programmable computers. The difficulty lies in explaining how these solutions are arrived at. The decision-making logic of intelligent machines, especially those using neural networks, is quite opaque. It seems we can either have higher accuracy answers or clearly understand how the machine determined them, but not both.

Unsurprisingly, some organisations opt for machines whose decisions are explainable over potentially higher accuracy solutions because they can only trust decision-making that they can understand. To do otherwise seems to them too big a leap of faith. In response, DARPA has a new Explainable AI (XAI) program that aims to devise intelligent machines that both perform well and can explain to humans how they reached decisions. Some argue however that machines generating explanations acceptable to humans may not necessarily accord with how the machines generated the solution in the first place. The machines may have simply learnt how to please us.

Intelligent machines are distinctly different from earlier programmable ones and perhaps more like us. Intelligent machines do not necessarily give the same output each time in the same situation. While they can learn independently, it is not always apparent what they have learnt or how they categorise data. This aspect is magnified in neural network machines as they continue to learn and evolve 'on the job'. They are therefore capable of emergent behaviour and may well surprise, for better or worse, just as their intelligent human creators can.

Big Data

Learning machines learn from data; the more the better. The sudden arrival of 'big data' was a crucial step in the contemporary development of intelligent machines for without this the technology would have remained embryonic. Big data is defined as 'extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions.'[1] Big data's major elements, colloquially termed 'the three Vs', are: ever-larger Volumes of data; increasing Velocity of data flows, and growing Variety of sources (imagery, voice, video, text, etc).

Digital data is growing at an astonishing rate. In 2013, around the time that intelligent machine technology development quickened, the world produced 4.4 zettabytes of data. (A zettabyte is 10^{21} i.e. a one followed by 21 zeros.) By 2020, this annual production rate is expected to be 44 zettabytes and, by 2025, 163 zettabytes. Video

1 *Oxford English Dictionary*, https://en.oxforddictionaries.com/definition/big_data [accessed 23 February 2018].

and images makes up a large portion of the new digital data of which more than 80% is unstructured.

The older programmable technology machines can use only structured data. Such data is carefully organised to be able to be used in relational databases e.g. *Excel* spreadsheets. These databases are easily and quickly searchable using simple algorithms. Whether generated by humans or machine, the structured data used is purposefully formatted to fit the requirements of the computer systems being used. The Internet-of-Things (IoT) involves widespread multiple-type sensor dispersion; many of these produce structured data to allow ready machine-to-machine communication.

Unstructured data is the reverse: it does not fit into the fields of row-column databases. Types of unstructured data files include e-mail messages, documents, social media, videos, imagery, audio files, presentations and webpages. Such data may be generated by humans or by machines such as unmanned reconnaissance platforms and remote imagery devices. Unstructured data is not easily searchable.

Intelligent machines for the first time can analyse both structured and unstructured data, giving the data meaning in terms of relationships, patterns and associations. Without this type of computer system, most of the world's zettabyte data collection would be wasted. Moreover, the more structured and unstructured data that is fed into intelligent machines, the more effective and efficient they become through learning.

While data quantity is important, users increasingly realise that data quality may be even more so. Poor quality data can mislead intelligent machines, making their outputs dubious. Intelligent machines need data that is standardised, normalised, has duplicated data deleted, verified and enriched, with verifying and enriching particularly important to making the data useful. Enriching brings out specific features significant to the machine learning and the

problem set that the intelligent machine is being trained for. In 2015, the US DoD for the first time prioritised data quality over data quantity.

Quality is also important in data storage. There should be only a single data view even if the data is stored across multiple disparate systems. In achieving this, maintaining data hygiene is crucial; data should be clean, that is, mostly error-free. In contrast, dirty data describes data that is erroneous, incomplete and outdated. *Google*, *Amazon* and *Facebook* have invested significantly in data scrubbing to maintain the high level of data hygiene that their intelligent machines require to achieve reliable outputs.

The on-the-job learning capabilities of intelligent machines mean that more than stored data has an impact. Microsoft's experimental Tay chatbot was initially trained, apparently by neural networks, using cleaned data troves. Tay then went 'live' to engage with the public online using *Twitter* to improve its performance through machine learning from these interactions. However, *Twitter* trolls managed to retrain Tay using offensive tweets that caused Tay to erratically respond to some questions using racist slang and far-right ideology. Microsoft shut Tay down after 16 hours as Tay's tweets steadily worsened. Intelligent machines are only as good as the data they are trained on. Tay had a significant data-diet vulnerability in that the tweets it was fed when 'live' were unfiltered. A similar data-diet issue arose with IBM's *Watson* which learned to swear after accessing the website, *Urban Dictionary*. As the adage goes: 'garbage in, garbage out'.

Diet issues may also arise with machine-to-machine communication. Dispersed IoT sensors represent potential false data input points given that they often lack the computing power to host sophisticated cyber security software and that the risks of some network protocols employed, such as IPv6, are unknown. Smart cybersecurity systems running on the networks that feed into the intelligent machines may

need to include managed threat and anomaly detection, and predictive analysis.

Data-diet discussions suggest whether machine outputs can be trusted. Learning machines train on big data; the datasets they use can impact in four ways.

First, if the dataset is too small, the intelligent machine may skew or misunderstand the issue. However, obtaining large datasets can be problematic because even using supervised machine learning approaches, the data labelling effort can be substantial. Added to this is that even minor variations in the problem that is assigned to an intelligent machine may require a different dataset. For example, an intelligent machine being trained to land an aircraft may require a dataset for each of the different environmental conditions expected.

Second, the algorithms do not determine facts about the world but rather about the dataset. The intelligent machines therefore become tightly tuned to the data they are trained on rather than develop general rules about how the world works. While such machines may perform well, it is difficult to determine what their boundaries are and when they reach them; the machines thus can suddenly fail rather than degrade gracefully. There is a further twist to this: if we do not know how intelligent machines reach decisions, then we do not know what datasets are needed to train them properly.

Third, feeding machines large datasets may not allow them to determine which of the many decisions that they need to make to complete a complicated task are critical. Driving a vehicle for example involves many different tasks but machines have trouble learning which of them are especially vital and how all these tasks relate. In contrast, when learning a task, humans determine how the various elements interact and then decide which are critical to success. Humans can then adjust the tasks they undertake to suit the situation being experienced.

Last, the datasets used to train intelligent machines can be biased for a variety of reasons, thus making the machine's outputs less trustworthy. Bias can be introduced by the dataset not fully representing the problem. Such biases can be deliberately entered by malicious actors, as in the Tay case earlier discussed, or unintentionally when the dataset compilers consider only a narrow perspective. Furthermore, the datasets are inherently historical so that the solutions identified are biased towards the past. This causes the implicit assumption that the future is just like the past and thus change is problematic.

The complicated issues associated with big data and intelligent machines highlight the need for organisations to have sophisticated data strategies. Data availability, collection, hygiene and governance are all important matters for such strategies to address. Intelligent machines need focussed support to ensure that they learn most appropriately and effectively. While some new machine learning techniques involve reducing the quantity of data required, all approaches require some data. Moreover, a fundamental attribute of intelligent machines is that they learn from their interaction with humans and the environment or, in other words, they learn continually from new data. Carefully leveraging data seems essential to get the best from contemporary intelligent machines.

Cloud

The most efficient way to facilitate intelligent machines' access to big data sets is to use cloud computing. Broadly speaking, cloud computing involves storing and accessing data and programs from external sources using the internet rather than from the computer's own hard drive. In the late 1990s, a cumulus cloud drawing was used to represent the Internet and so 'cloud' become a metaphor

for accessing services over it. The Siri application (noted earlier) is mostly in-the-cloud regarding speech recognition, natural language interpretation and various information services.

Cloud computing proponents assert that it is more resilient, secure, scalable, agile, responsive and supportive of information sharing than on-computer hard-drive storage or using hardware servers on small local area networks. The disadvantages include: first, the cloud may crash or become unavailable, thus preventing access and sharply degrading an organisation's capabilities; second, privacy may be limited because the cloud service provider can readily access your cloud data; third, the cloud infrastructure is owned and determined by the service provider; and last, customisation options are inherently few.

While, algorithmic warfare requires cloud computing for best performance, many current cloud computing technologies are not suited to AI and machine learning. For example, the data fed to intelligent machines from the cloud needs to be of the machine-required quality; good cloud data hygiene is essential. There are inherent challenges in cleaning, standardising and normalising data accessed in real-time from the many potential different applications and sources given these could be classified, private, public, domestic, international, human or machine.

Military clouds are technically challenging, as they must be accessible in harsh electronic countermeasure environments. Not all intelligent machines on the battlefield though may need to access the cloud directly. Instead, a larger nearby platform may provide a local area network that the smaller intelligent measures plug into. The larger platform may then act as a secure, robust, jam-resistant gateway to connect with the overarching cloud. The smaller intelligent machines operating far forward within the harshest electronic environment can then use redundant and

secure line-of-sight communication links between themselves, and backwards to the larger platform.

Testing

Intelligent machines are inherently unpredictable; their behaviours are emergent not fixed as the machines we have become used to or which today's testing regime is based on. Given this, there are four important issues when considering testing an intelligent machine.

First, intelligent machines have many possible system states and it is impossible to test all, or indeed, most. Second, they interact with a dynamic environment that is unable to be specified and certified beforehand. Third, intelligent machines learn but it is not possible to determine fully what they have learnt. In this regard, intelligent machine developers are more likely to be data collectors and trainers than the conventional programmers that current testing regimes assume. Last, while humans test machines so that they gain trust their performance, this generally involves testing in closed scripted environments. While this approach is inherently unsuitable for testing intelligent machines, no alternative has been found that will give humans complete confidence.

Testing becomes even more problematic when we remember that intelligent machines evolve by learning from their experience. Contemporary testing regimes test a machine once, confident that, as long as the design is not altered, the machine will perform similarly over time, regardless of where in the production line it has been manufactured. This is not so for intelligent machines. They evolve and, depending on their design, may not self-synchronise; this means that all similar machines may not perform the same. Testing and recertification may need to happen periodically to account

for emergent behaviours and to determine their similarities and differences by testing each 'identical' intelligent machine.

Conclusion

Algorithmic warfare involves intelligent machines, big data and the cloud. In considering these elements, we tend to draw instinctively on our earlier understandings about programmable computers. This is not surprising because they have become such a large part of our home and work lives that their presence is not just unremarkable but required. If these machines do not produce consistent outcomes, we know there must be a hardware or software failure. We also know that their software can be replicated across millions of machines so they all perform the same.

These 'understandings' are out of place in the new world of intelligent machines. Perhaps the phrase 'intelligent machine' is itself somewhat misleading. In some respects, intelligent machines react more like humans than traditional machines. At the upper end, they are self-aware and self-evolving, displaying a new form of intelligence high on the evolutionary train. Their outputs may surprise us; they keep on learning the longer they operate so that testing them may require applying techniques humans use to test each other. Implementing the concept of algorithmic warfare across our warfighting organisations may require innovative approaches quite unlike those earlier employed with our non-intelligent machines.

3.

Making Algorithmic Warfare Happen

Thinking about intelligent machines has long featured concerns about their replacing humans. Contemporary intelligent machines have some real strengths but their technological foundations build in some real shortcomings. These foundations limit contemporary intelligent machines to having narrow, not general, intelligence. The key, at least initially, is using their strengths to support human decision-making, but not supplant it.

Narrow machine intelligence equals or exceeds human intelligence for specific tasks within a particular domain; their utility is context-dependent. In contrast, general machine intelligence equals the full range of human performance for any task in any domain. When general AI might be achieved remains debatable, but seems several decades away.

Military force structures take a long time to change and so, accordingly, several decades may not be that far in the future. While there is a fascination in popular culture with robot warriors, Terminators and Slaughter-bots, warfighting is a practical activity. The technology that will allow general intelligence is unknown, thus making useful calculations about its potential battlefield capabilities at best problematic. Instead the interest for the near and medium term is in the way that narrow intelligence machine technologies could be employed in the modern battlefield.

In general terms, combat success goes to those who can maximise their strengths while minimising their weaknesses. Both humans and contemporary technology intelligent machines have their strengths and weakness and crucially these differ. The most important role for intelligent machines might thus involve using their strengths to reduce how human weaknesses impact negatively on achieving battlefield success. This complementary approach seems to fit best with the capabilities of today's emerging narrow intelligence machines. The matter is though more complex than it first appears; implementation has some twists and turns.

Intelligent Machines' Fortes and Foibles

While narrow intelligence machines can be very powerful, they are simultaneously quite brittle being generally unable to handle quite minor context changes. Google's *AlphaGo* defeated the world's best human Go player but only on the standard 19-by-19-inch board. *AlphaGo's* move-selection and position-evaluation convolutional neural networks were trained with data related to that specific size board. To play on other size boards would require new training and software code changes. In another example, a machine trained to read formal documents may struggle to read vernacular texts.

The domain adaptability of intelligent machines, that is, being able to apply knowledge learned in one context to another, is also poor compared to humans. Moving a machine from one task to another generally requires retraining, although emerging transfer-learning techniques can help. These can help an intelligent machine shift from doing one task to another similar one. For example, *AlphaGo* learned chess in four hours and reached a standard able to defeat most computer chess programs. This task adjustment involved moving a machine optimised to play games from one game to

another; a relatively minor change assisted by highly codified game rules.

Domain adaptability in the military world though relates more to operational environments: air, sea, land, space and cyber. Narrow intelligence machines can operate in these different environments with varying degrees of difficulty. Autonomous systems like driverless cars need to build a high-fidelity model of the location that they operate in. For this, they use an integrated but eclectic array of sensors and wireless networking. Even so, operating in the inherently cluttered, confusing, always-changing ground environment remains difficult for narrow intelligence machines. However, unmanned air vehicles fly in relatively uncrowded skies characterised by low rates of environmental change and few obstacles, thus making these narrow intelligence machine's tasks much easier. The most favourable situation for narrow intelligence autonomous systems to operate in is low-complexity environments with little uncertainty.

Accepting the limitations of brittleness and domain adaptability, there are numerous tasks that military forces undertake that can usefully exploit the strengths of machine intelligence. Such tasks may be grouped under the three Vs of big data: volume, velocity and variety.

The volume of data created annually is measured in zettabytes as discussed earlier. This is far too much to be analysed by conventional means. In a limited example, the US Air Force has a wide-area imagery sensor for city surveillance. However, it takes 20 analysts working continually over a 24-hour period to exploit even 10% of the collected data. The remaining data is stored and may never be properly examined. The underlying problem is that there are simply not enough people available to analyse all the data being collected, even if personnel budgets were unconstrained. The problem is compounded as humans are inherently slow to analyse and process data within a given timeframe.

The first major use of intelligent machines in the US Air Force is analysing the huge volumes of video data collected by unmanned aircraft systems during counter-terrorism operations across the greater Middle East. The recently established Algorithmic Warfare Cross-Functional Team, or Project Maven, aims to quickly bring into service systems capable of applying intelligent machine learning and algorithmic solutions to sorting and analysing the vast amounts of intelligence data that MQ-9 Predator remotely piloted aircraft collect. Maven may be extended later into other areas experiencing high volume data loads including logistics, communications, situational awareness and management.

The velocity of data also raises issues. With the increasing sophistication of programmable machines, the pace of warfare has stepped up. It is becoming very difficult for humans to comprehend what is happening fast enough to react in a sensible manner. Intelligent machines can cope with very fast data flows and generally make relatively better decisions than humans in such circumstances although these decisions may not necessarily be the optimum. In the commercial sphere, automated stock trading is an example of high-velocity machine decision-making that is mostly successful but which periodically fails. In the military domain, machines are increasingly being tasked with responding automatically to high-speed threats in the fields of missile defence, cyber-attacks and electronic warfare.

Lastly, are issues associated with the increasing variety of data. The recent exponential increase in the numbers and types of sensors means that many diverse types of data are collected. Because humans have limited attention frames, they tend to find that some data sources are easier to comprehend than others. Video, for example, is easier to instinctively understand than detailed radar parametric data. Intelligent machines are better able to scan the many data sources and types and, in real-time, determine activity patterns, associations and relationships. With algorithms assisting them, intelligence

analysts can be automatically alerted to significant events rather than having to continually assess multiple data streams waiting for such events to occur. Intelligent machines can be attention-multipliers.

While intelligent machines are well suited for big 'V' tasks, other tasks better suit humans. They are, for example, better at inductive thought: being able to generalise from limited information. Because intelligent machines need a lot more information to do so, they can only derive broad rules from the data provided, and this needs to be high quality as discussed in the previous chapter.

Inductive thinking also relates to matters of uncertainty; intelligent machines perform more reliably in low uncertainty environments and situations. The battlefield however is generally an environment of high uncertainty and only limited information on which to base actions. Indeed, adversaries try to magnify both aspects and use them to their advantage.

The problem is compounded as battlefields are inherently confusing; knowing which specific snippets of information are important can be hard to discern. Thus, intelligent machines soaking up large volumes of diverse data may devise somewhat skewed decisions through assessing both small amounts of useful and large quantities of useless information. Humans are better at using expert knowledge to disregard extraneous information quickly and focus somewhat frugally on just the information needed to make a decision. In other words, humans generally make better judgments in environments of high uncertainty.

Human-Machine Teaming

Humans and machines both clearly have strengths and shortcomings. This means that, regarding contemporary technology, the notion that machines could replace humans is misleading. In algorithmic warfare, the real implementation issue is then a practical one: determining the optimum blend of human and machine intelligence that best leverages the attributes of each. In other words, what is the best way to form human-machine combat teams?

The teaming of humans and intelligent machines has been likened to the mythical Centaur, a half-man, half-horse creature with the brains of a human and the power and speed of a horse. Importantly however, there is an interdependence in human-machine teaming that goes beyond the simple maximise strengths/minimise weakness implication of the Centaur analogy.

In the matches between *AlphaGo* and the world's best human player, it was found that both improved from interacting. *AlphaGo* developed a new move that startled humans while the human player also devised an original, novel move. He and others who have played *AlphaGo* consider that they now look at the game from a fresh perspective and have sharply improved their performances.

The Go experience repeats the occasion when computers first defeated humans at chess in the late 1990s. Since that time, humans have played chess with computers and chess computers have played each other but the best results have come from human-computer teams playing together. World Chess Champion Garry Kasparov observed in 2010 that:

> 'Teams of human plus machine dominated even the strongest computers. Human strategic guidance combined with the tactical acuity of a computer was overwhelming. We could concentrate on strategic planning instead of spending so

> much time on calculations. Human creativity was even more paramount under these conditions.'[2]

This suggests a way forward in human-machine teaming. The human segment makes use of the human abilities in intuition, induction, lateral thinking, domain adaptability, generalising, and working in high-certainty and complex environments. The machine segment makes use of the machine abilities involving 3V big data analysis, high-speed problem solving, and single-mindedly concentrating on tightly defined problems.

What this laundry list might mean in practice was demonstrated when Kasparov teamed with a machine to play chess. He did not need to worry about making simple tactical gaffes. The machine could project ahead, and forecast the consequences and the likely counters the other player might take. Given this, Kasparov could focus more on strategic matters.

In some respects, intelligent-machine teaming might be seen as having a research assistant able to complete analytic, estimation and forecasting tasks very quickly. Humans can devise numerous potential strategies. The intelligent machines can then quickly calculate the probability of success for each option and the likely countermoves by an adversary. The seeming inference that the next step would be to remove the humans and completely automate strategy development would be mistaken. Intelligent machines are probabilistic machines that determine the mathematical likelihood of certain events occurring; they choose the highest probability outcome. Consequently, they do not take risks or 'leaps of faith' as humans do.

---

2 Garry Kasparov, 'The Chess Master and the Computer', *The New York Review of Books*, 2 February 2010.

A strategy is not repeatable in a mathematical sense. Indeed, if a strategy has previously succeeded, others will anticipate a repetition and will have already taken steps to thwart it.[3] Instead, success generally comes to those able to be creative and devise new strategies. Intelligent machines drawing on historical data appear inherently incapable of that creative leap.

A strategy operates within a unique context and is thus in itself unique. Determining probabilities of success can help humans hone their thinking and throw up new ideas but not in itself devise consistently winning approaches. By their nature, unique situations are not suited to being solved using probabilities. Intelligent machines seem suitable for solving some types of problems but not all.

There is an important twist. In 2005, the online site *Playchess.com* hosted a freestyle chess tournament where teams of people and computers competed in any combination they thought best. The overall results were predictable. Human-machine teams defeated teams of machines only, seemingly irrespective of how sophisticated the computers were in the computer-only teams or how poor the machines in the human-machine teams were. The surprise came with the winners: two unremarkable chess players teamed with three equally unremarkable machines.

The difference in this human-machine teaming was that the humans had focussed on having better business process designs. In this case, the business processes encompassed firstly how the humans should work better with the machines to accomplish the specific task (winning at Chess), and secondly how the machines should better

3 Edward Luttwak, *Strategy: The Logic of War and Peace*, Belknap Press, Cambridge, 1987.

think about the task. The outcome of this approach led Kasparov to observe that:

> 'Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process.'

In warfighting, where success is crucial, this observation is worth deeply considering and perhaps incorporating. Such an approach would involve building an improved human-machine interface in the sense that both humans and machine are trained to work together better. The nature of the interface would vary with the specific tasks to be undertaken. This interface would be comprised of both tangible and intangible interface elements: machine hardware and software combining with human skilling.

In addition, and addressing the second part of the business process, the humans and the intelligent machines would together examine problems related to the task. The team's combined skills would be improved by devising solutions through assessing various options. This is a kind of reinforcement learning, working backwards and forwards between humans and intelligent machines, to educate both. After all, both are 'learning' machines.

Together, these various activities would allow the humans to gain more trust in the intelligent machines and acquire useful knowledge about their foibles. There may still be no way to convincingly explain why the intelligent machines make the decisions they do even though broad experience with them would allow an understanding to develop of when they were likely to fail.

Human-Machine Interactions

The operational capability of human-machine teams appears to depend on how the humans and machines relate. In this of course, all parties are innately unpredictable and thus the teams are dynamic. There will be stresses and strains on both sides of the human-machine divide that make the capability of the team somewhat uncertain, much like human only teams.

Several human-machine team modes are commonly identified. These lie on the continuum from manual control to full autonomy, a term whose precise definition is itself proving difficult.

Human-in-the-loop. In this mode, humans retain control of selected functions preventing actions by the intelligent machines without authorisation; humans are integral to the system's control loop. The difficult design issue is how to determine exactly where in the process human intervention should be undertaken, and that will vary with the task and the capabilities of the machine. If too much human intervention is needed, its usefulness may be doubtful.

Human-on-the-loop. The intelligent machine controls all aspects of its operations but humans monitor the operations and can intervene when, and if, necessary. In a variation, the machine, when at a critical point, such as engaging a target, might notify the human about impending action and either await positive authorisation or continue unless stopped. Some missile defence systems use human-on-the-loop techniques whereby the system proceeds unless a human overrules the automated track engagement decision.

Human-out-of-the-loop. The machine's algorithms control all aspects of system operation without human guidance or intervention. The machine engages without direct human authorisation or notification. While using human-out-of-the-loop operations is problematic, its implementation is by no means uncommon. This

form of control is at times also termed human-off-the-loop or autonomous.

Human-out-of-the-loop is of most concern in circumstances where a machine independently targets and fires weapons at people on the battlefield without human authorisation. Such control though might be much less problematic when injuring humans is unlikely such as during cyber warfare operations or employing electronic countermeasures.

Human-out-of-the-loop might also be acceptable if restricted to being used for defending humans. For example, an anti-missile system might be allowed to operate automatically as soon as an approaching hostile high-speed missile was detected if there was insufficient time to consult humans before its impact. The naval Phalanx Close-In-Weapon System of the 1980s could be set to automatically search for and engage any missiles that the system's algorithms deemed a threat.

Similarly, human-out-of-the-loop has also long been used in anti-ship missiles like Harpoon (see Chapter 1). Today's land-attack missiles are comparable. Launched from hundreds of kilometres away, these missiles independently navigate to the target area and then search for and identify the targeted building and engage.

Fire and forget weapon systems have been used for many years in various situations where the danger to humans is thought too high and requiring reducing. This history makes some recent concerns about future human-out-of-the-loop machines appear somewhat belated although no less important. The laws of armed conflict mandate that discriminating between combatants and non-combatants must be attempted (discussed more later). Some users of earlier fire and forget weapons, such as land mines, may have paid too little attention to this mandate.

Machine-to-Machine. While less obvious, this mode of machine interaction is becoming increasingly important. Machine-to-machine

interaction is fundamental to high-speed battlefield actions and achieving battlefield effects faster. However, with the high-speed communications involved, unexpected interactions or errors can cause the system to spiral out of control very quickly.

Several 'flash crashes' have shocked financial markets because of unintended intelligent machine interactions. While intelligent machines may regularly perform better than humans in certain tasks, they will occasionally fail. If they are tightly coupled to multiple other intelligent machines, there is a distinct possibility that a failure - or some unexpected output - may turn into a flash crash with potentially catastrophic effects. In planning machine-to-machine interaction, attention needs to be paid to developing overall system resilience so as to be able to manage a flash crash.

Contemporary intelligent machines offer the warfighter excellent capabilities albeit narrow and with some shortcomings. This means that the major issue now in introducing intelligent machines to the battlefield is finding the best mix of machine and human competencies. Notable in this is the capability of the weak human+machine+better process combination to defeat both highly capable machines alone and strong human+machine+inferior process combinations. While highly skilled humans and sophisticated intelligent machines are necessary, they are perhaps not sufficient to succeed on the battlefield. Task-optimised human-machine interfaces could be the key to optimal human-machine teaming and thereby victory in future wars.

4.

Waging Algorithmic Warfare

The business of war has seen many technologies come and go, some evolutionary, others revolutionary. The full impact that intelligent machines will have on future warfighting is unclear but some indications can be discerned.

There have long been earnest debates about machines and the nature of war. It seems the nature of war will stay as it is: violent, chaotic, destructive and murderous, with no change likely from intelligent machine technologies. War at its core, however, involves applying or threatening violence to humans. If both sides have intelligent machines, it may become simply a case of machines being violent to other machines. But is this still war?

While a robot battle would test the opposing states' materiel resources as the process of violent machine attrition ran its course, whether it would be consequential is uncertain. Geoffrey Blainey considers that wars are undertaken when the states concerned do not have an accurate assessment of their relative strengths.[4] Robot wars may be a means to gain such an assessment albeit one more of materiel strengths than of moral ones. If states feel they have a greater stake in a conflict after the wars between the machines conclude, they may move on to wars between the people.

---

4 Geoffrey Blainey, *The Causes of War*, 3rd Edn, The Free Press, New York, 1988.

There is a seductive notion inherent here that low stakes conflicts might be able to be decided by robot wars even if high stakes ones cannot. Such robot wars would be similar then to current grey zone conflicts but destructive, a new step then in the continuum between war and peace. Such an argument overturns a Russian notion (discussed further in the next chapter) that, if all involved have intelligent machines, and none has a decisive advantage, that there will be peace through a balancing of military power. Instead, such widespread proliferation may lead to a greater temptation for states to unleash robot forces upon each other, hopeful that, while avoiding human casualties, the robots can decide the issue.

Today that perspective would probably mainly involve intelligent machine cyber forces battling in the virtual domain. In the medium-term perhaps it may expand to intelligent-machine swarms fighting each other at sea, in the air or in remote land areas. Crisis management approaches will need to be reconceived, with priority given initially to approaches to manage intelligent machine-powered cyber-attacks.

If the nature of war in the main appears only little changed with the rise of intelligent machines, not so the character of war. There are already warnings that intelligent machines will have significant impacts and perhaps overthrow some time-honoured precepts. These issues are discussed in the next section.

Regarding strategies, the picture is more nuanced, as the final section examines. There are two distinct schools of thought: will intelligent machines allow us to do things better or rather do better things? Robert Work, the US Deputy Secretary of Defense in the last years of the Obama Administration and a very knowledgeable and passionate intelligent machine advocate, has held both views. Initially, he held the 'do better things' position and then later emphasised 'do things better'. Both positions are worth discussing as they bring out useful perspectives on waging algorithmic warfare.

War's Changing Character

Curiously, intelligent machines may return mass to the battlefield. In recent decades the trend in armed forces has been to develop force structures based around a relatively small number of highly effective, multi-role platforms. Intelligent machine technology may allow these highly sophisticated weapon systems to be complemented by a very large number of dramatically lower cost, unmanned systems optimised for specific tasks. The unmanned systems would be *in extremis* expendable and so could be risked in the more dangerous tasks that the few expensive manned platforms might not sensibly be.

With numerically larger forces, attrition would no longer be the Achilles heel it is today where losing even one major platform, such as an aircraft carrier, could be disastrous. In contrast, if a force of few manned systems and many unmanned systems was fielded, attrition would be both tolerable and better able to be actively managed. Such a force structure would then gracefully degrade during combat operations but not potentially, catastrophically fail.

Such a force would be characterised by having highly dispersed capabilities. This is important for smaller defence forces where with current force structure models, important capabilities might reside in only one or two large platforms. If those platforms are destroyed or damaged crucial capabilities might be completely lost. In contrast, in a mixed manned/ unmanned force structure these capabilities could be widely spread across many elements. The adversary would have difficulty targeting sufficient numbers of the dispersed force elements to cause the whole capability to be lost.

Intelligent machines may also quicken the pace of battle. Intelligent machines can analyse big V data much faster than humans potentially speeding up decision-making dramatically. Already, intelligent machines are being used in those defensive systems where

time is critical, such as cyber and anti-missile defence. Future tightly integrated offensive and defensive intelligent machine systems could respond to threats at machine speed. The pace of battle would then exceed what humans can keep up with, in terms of understanding what actions were being taken and whether they were proving successful or not. At least for a period, war might escape human control. Victory, or defeat, might not be known until the intelligent machines involved in fighting the battles announced it.

Visions of massed unmanned forces controlled by machine-speed decision-makers suggest significant potential disruption to current force structure models. This may be especially so for those forces built around a small number of large platforms. In this regard, some in the US have questioned whether large aircraft carriers could withstand attacks by swarms of hundreds of small, unmanned air vehicles. New force structure models may be needed.

Such ideas though might impact upon an old idea that has applied mainly in conventional warfighting. Traditionally, the size of a population has suggested a nation's potential military power. While small states with high-quality forces might achieve some remarkable victories, over time large population countries could, if they wished, always grind them down. Quantity measured by numbers of people was seen to have a strategic quality of its own.

Intelligent machines bid fair to overturn this. Small wealthy countries may now be able to generate greater mass than much larger poor ones. Moreover, countries with unfavourable demographics, with more old people than young, may no longer be disadvantaged. Developing an intelligent machine heavy force structure might allow a small number of personnel to wield disproportionally large combat power.

This notion can be extended into training. Intelligent machines could revolutionise military training as they have chess instruction. It has long been known that people advance faster at playing games

by playing progressively better opponents. Since the late 1990s, young chess players have been honing their skills by literally playing the world's best opponents on their home computers and devices. This approach has sharply accelerated their training.

In 1958, Bobby Fischer became a chess grandmaster at 15; this age record was not broken for 33 years. Over the last 27 years however, 20 others have broken it with the record now standing at age 12. Coupled with accelerated learning is that the computer-age students seem less constrained by traditional chess tenets. Moves are now simply valued in terms of being good or bad for game success not whether they conform to approved doctrines.

Regarding military skills, Western military forces are considered the benchmark to compare national forces against. Some believe that other states, by using carefully devised intelligent machine training, may be able to overtake Western competencies and field more skilled personnel. Western militaries' great advantage would then be 'checkmated'.

The chess example highlights that a noticeable feature of algorithmic warfare is that commercial drivers fundamentally shape it. During the Cold War, large military R&D spending drove technological development, allowing commercial spinoffs. Today, large commercial R&D spending drives intelligent machine developments, bringing military spinoffs. Such spinoffs though are not necessarily likely to be optimised for military purposes. Instead, they will be designed to meet consumer demands even though that may make them affordable by armed forces.

The commercial imperative will have further influences. Intelligent machine technology may evolve much faster than traditional military technology as commercial developers will want to quickly capture a return from their investment before other better products arrive. Similar market imperatives suggest that new intelligent machine technology will quickly diffuse globally. Both factors suggest that

strategic surprise may be possible when a country, or even a non-state actor, suddenly fields unexpectedly effective intelligent machines.

In this way, intelligent machine developments aimed at consumer market profits might encourage regional arms races as a secondary effect. Keeping up with intelligent machine developments in neighbouring countries may seem prudent, even if the military utility of these developments is uncertain. In a similar vein, market considerations further suggest that trying to ban or legally constrain intelligent machine technologies may be problematic.

These dynamics call for armed forces to be more permeable to outside influences than previously. To succeed with intelligent machines, ideas and technology will need to flow both inwards and outwards between the commercial and military worlds. This is especially so given the apparent importance of optimising human-machine interfaces for military tasks (as highlighted earlier). Armed forces may need to adopt new organisational structures and processes to make them more porous and better able to exploit commercial intelligent machine technologies and thinking.

Strategies: Do Things Better

In recent years, American strategic thinkers and the US DoD developed the Third Offset concept. Drawing on Defence Science Board research findings, the concept envisaged inserting intelligent machines deep into America's battle networks to achieve a step increase in performance.

Two great power competitors, China and Russia, have built up theatre-wide battle networks comparable in performance to America's and potentially able to deny US military forces access to specific regions. The Third Offset aimed to enhance America's battle networks so they could overcome these networks if needed.

Developing such a conventional force capability would strengthen deterrence by denial and so avoid reliance on nuclear deterrence.

Battle networks comprise interlinked digital computer systems conceptualised as four virtual grids (information, sensing, effect and command) that overlay the operational theatre. The various elements of an armed force, from individuals to single platforms to battle groups, are then interacting nodes on these grids. Each can receive, act on, or forward data provided from the various grids as appropriate. The operation of the grids can be understood using John Boyd's well-known observation, orientation, decision and action (OODA) loop. The sensing grid observes, the information grid orients (through disseminating information), the command grid decides, and the effects grid acts by targeting adversary forces.[5]

Boyd's OODA loop is the principal idea animating current battle network operations and so is important when considering inserting intelligent machines into them. For Boyd, winning at any level of war requires working the OODA loop faster than an adversary. Doing this means that the adversary's reactions to friendly force initiatives will always lag, becoming less and less appropriate to the battle as it evolves. The crucial aspect to attaining the necessary faster OODA loop speed is rapid orientation. Success lies in building an accurate image of the battlespace more rapidly than an opponent. Situational awareness is the *sine qua non* of victory; a notion that military aviators have turned into a mantra.

Modern battle networks are excellent at gathering, storing and sharing information from the battlefield but noticeably less successful at processing and contextualising this information. The

5 For more detail on battle networks see Peter Layton, *Fifth Generation Air Warfare*, Paper No. 43; RAAF Air Power Development Centre, Canberra, June 2017.

networks have been overwhelmed by big data's 3Vs: volume, velocity and variety, and have trouble turning data quickly into useful intelligence. The networks have not proven as effective as originally hoped in building an accurate battlefield picture, especially when time is constrained.

The new intelligent machines offer a solution to this shortcoming. With these inserted, Robert Work considers battle networks:

> 'will be able to sense and perceive battlefield patterns more readily and rapidly, facilitate more timely and relevant combat decisions, and apply more rapid, discreet and accurate effects with less loss of life. If all these things happen, the Joint Force will operate at a higher, more effective tempo than its adversaries, and thereby gain an important, if not decisive, advantage in both campaign and tactical level operations.'[6]

As an important influence on Third Offset thinking, Work considered the big idea in battle network enhancements was human-machine teaming; this was 'the coin of the realm', the Third Offset's central organising element. The way to do things better was to spread intelligent machines across all aspects of battle networks. There were five basic building blocks.

First are deep learning machines, powered by neural networks and trained with big data sets, inserted into every battle network grid. They would speed up grid operation especially those against high-speed cyber, electronic-warfare, and space-architecture attacks and for those times when 'missiles …are coming screaming in at you at Mach 6 [when] you're going to have to have a learning machine that helps you solve that problem right away.'

---

6 Robert Work, 'Artificial Intelligence, Autonomous Systems and the Third Offset', pp 2-5 in *Artificial Intelligence, Big Data And Cloud Taxonomy*, Govini, Arlington, 2017.

Second is improved human-machine collaboration using intelligent learning machines to help humans make higher quality decisions more quickly. The intelligent machines could better and much more swiftly analyse big data and advise humans on operationally significant patterns, associations and relationships found.

Third involves using intelligent machines to facilitate assisted human operations. With this, all combat forces could plug directly into and call upon the power of the entire battle network to accomplish assigned tasks. In advocating this building block, Work declared that: 'I'm telling you right now, ten years from now if the first person through a breach isn't a fricking robot, shame on us. Assisted human operations, wearable electronics, making sure that our war-fighters have combat apps that help them in every single possible contingency. We can do this.'

Fourth is enhanced human-machine combat teaming that allows seamless coordinated operations between manned and unmanned systems, including some that are increasingly autonomous in their operations. In the near-term, these might include small intelligent machine vehicles that support lower-level infantry units by following them around with stores and ammunition or self-driving trucks that follow a lead manned vehicle. Notable is that both applications are to improve existing practices but not offer revolutionary capabilities.

Lastly are intelligent-machine enabled kinetic and non-kinetic autonomous weapons capable of collaborative high-speed attacks. Such technology could allow extensive cross-domain attacks to be mounted simultaneously in an intelligent machine powered extension of the 1990s' concepts of parallel warfare.

While the focus in the Third Offset intelligent machine approach is on great-power conventional deterrence, it was considered that such enhancements would allow engaging other lesser states and non-state actors. In the latter case, for example, intelligent machine

analysis of suitable big data could help target terrorist groups. Such analysis could explore online Islamic State or Al Qaeda narratives at machine speed to find operationally significant patterns, associations and relationships. Such technology can reportedly examine more non-English language content in a week than Western open-source agencies have in 30 years. This new ability to look across very large news article datasets spanning dozens of languages can be used for many purposes including mapping and tracking in real-time the discourse of particular terrorist organisations.

Strategies: Do Better Things

A more radical concept has been developed in parallel with the Third Offset's approach of diffusing intelligent machines across battle networks. In contrast, this idea distributes intelligent machines in a manner that shifts the primary function of battle networks away from information sharing and towards machine-waged warfare. This approach potentially makes battle networks less important.

While, as with the Third Offset, the strategy problem remains strengthening deterrence by denial, the threat perception is somewhat different. The near-term future is seen as featuring potentially hostile state and non-state actors that can both employ precision-guided weapon systems and integrated battle networks of various forms across the full conflict spectrum.

Countering such state-of-the-art dangers requires new operational ways and capability means to win on the envisaged increasingly lethal battlefields. Future-force structure options are though constrained as the costs of personnel and manned weapons systems are high and sharply rising. Force sizes are accordingly expected to continue to decline although with better quality.

The problem is that, in an era of proliferated hostile guided weapons system and adversary battle networks, it is not apparent whether quality will remain able to overcome quantity. Mass may again become an important force attribute when adversaries can impose medium to high attrition rates on numerically small friendly forces. The present force structure model is probably unsustainable into the medium-future.

A distributed intelligent-machine approach potentially addresses this dilemma by transforming battle networks from movers of information into active fighting networks. Intelligent machines do more than enhance processing and improve contextualising; they now become actors themselves. This offers a vision of robotic warfare conducted by unmanned and increasingly autonomous intelligent machine weapon systems, operating across multiple domains (air, sea, land, space, and cyber) and across all types of military operations.

While this intelligent machine warfare approach still involves human-machine teaming, the place of machines is much greater. The 'do things better' diffused intelligent machine approach emphasised the role of humans in man-machine teaming. The 'do better things' distributed intelligent machine approach reverses this. Machines now loom large as meaningful participants, not simply trusted advisers: humans command, machines do. In some respects, it is now not the battle network that is key to victory but the edge devices; the network is conceptually inverted.

This proposed machine-waged way of war brings three gains. First, it could allow affordable mass based around a force structure of many unmanned systems and limited numbers of crewed platforms. Second, it would sharply lower risks to personnel, thus lowering casualty rates of hard-to-replace, highly skilled people. Moreover, it would also ease the stresses and strains of war on people while reducing human workload, fatigue and cognition demands. Third, it could lower the present battle network's vulnerabilities to electronic

jamming and cyber-attacks by sharply reducing the communication demands across the four grids. Intelligent machines may be able to wage war semi-independently, only needing human guidance from afar occasionally.

The distributed intelligent machine approach does raise a fundamental issue when considering warfare overall. The approach alters the shape of the present offence-defence balance although the direction is unsure. Is the offence or the defence dominant in this brave new robotic age? The worry is that, if offence dominates, the incentives to strike first in a crisis might grow. This would be strategically destabilising as all participants would then prefer to land the first blow given this may be a knockout one.

In considering changes to tactical level warfighting, this approach's stress on intelligent machine actors may impact mostly in enabling both faster pace and the use of swarms.

In terms of pace, intelligent machines would now form many machine-machine teams that would undertake numerous tasks at machine speeds. Actions would occur at speeds never before seen in warfighting. Such high-velocity battlefield activities would lead to hyperwar, a term coined by USMC General (Rtd) John Allen and Amir Husain.

Hyperwar sees human decision-making as having a somewhat rather secondary role at the tactical level. As the time to complete the OODA loop approaches zero, human cognition will simply be unable to keep up. Allen and Husain write that:

> 'The speed of battle at the tactical end of the warfare spectrum will accelerate enormously, collapsing the decision-action cycle to fractions of a second, giving the decisive edge to the side with the more autonomous decision-action concurrency. At the operational level, commanders will be able to 'sense', 'see', and engage enemy formations far more quickly by applying machine-learning algorithms to collection and

> analysis of huge quantities of information and directing swarms of complex, autonomous systems to simultaneously attack the enemy throughout his operational depth.'[7]

Intelligent machine-speed decision making will almost instantaneously coordinate large groups of sensors and shooters, thus enabling rapid force massing, machine attacks across large areas and quick regrouping for rapid re-tasking.

The 'foot' soldiers of hyperwar are conceived as being consumer quadcopter drone sized, unmanned air systems of many different types controlled by on-board intelligent machine technologies. Air systems match machine-speeds best in allowing easy traversing of difficult terrain, quick regrouping into task-oriented force packages, and relatively swift action. While an interesting notion, such small unmanned air systems would suffer significant range and payload issues especially if needing to attack to an enemy's operational depth. It may be practical only in the forward edge of the battlefield unless long-range, large-scale drone delivery, and perhaps recovery, systems are developed. Accordingly, some see such autonomous drones being game-changing technology principally when used in surveillance, reconnaissance and light-attack missions in dense urban environments, the most likely battlefields of the future.

Current small consumer drones are generally programmable machines flown using a combination of human commands and off-board computer processing. New computer chips developed to meet commercial demand (e.g. *Nvidia's* Xavier chip for autonomous vehicles noted earlier) will progressively provide consumer drones with onboard intelligence. This will allow them to navigate and process sensor-collected data onboard independent of 'the cloud',

---

7 General John R. Allen, U.S. Marine Corps (Retired), and Amir Husain, 'On Hyperwar', *USNI Proceedings Magazine*, Vol. 143, No. 7, July 2017.

GPS signals, or a remote hand controller. They will become capable of autonomous operations once launched.

In research and development are numerous intelligent-machine drones. *Nvidia* has trained an intelligent machine drone fitted with its chips and two cameras to navigate down densely forested trails where GPS signals cannot be received. Swiss researchers are flying *DroNet* that uses a smart phone camera and machine intelligence algorithms to interpret complicated urban street scenes and navigate them safely. The algorithm features a deep neural network trained using several thousand urban road-driving scenes; the algorithm was then able to transfer this learning into a closed environment and fly indoors and in large parking garages. The first affordable, intelligent machine consumer drone is likely to be the forthcoming Teal 2 that uses the *Nvidia's* Jetson TX2 chip running *Neurala* software featuring learning algorithms.

The broad hyperwar concept indicates what may be progressively achievable by placing increasing priority on intelligent machine autonomy. Even so, hyperwar is more likely to involve a continuing series of multi-domain salvo or spasm attacks rather than a continuous flowing action. Physical constraints mean that it would take time to rearm, refuel and reposition own-force intelligent machines for follow-on attacks. There would be a further need to assess damage inflicted and adversary responses. Hyperwar it seems might follow a shoot-look-shoot schema. Humans would provide the initial intent in the look phase and then launch human-out-of-the-loop intelligent machines to undertake the shoot phase and attack.

Beyond pace, the second fundamental change to warfighting may be the rise of the swarm. A swarm comprises numerous individual elements or small groups that coordinate to undertake specific missions as a coherent whole. In warfighting, the elements would be heterogeneous with simple and complex elements, each optimised for different tasks, but all coordinating as a single battlefield entity.

In contrast to manoeuvre warfare where own-forces concentrate to attack specific centres of gravity, the swarm construct envisages own-forces widely distributed across a battlefield and only coming together when needed. Such dispersal complicates an adversary's response, as the swarm seems everywhere and nowhere. Once the swarm comes together though, they are in such numbers that they can quickly overwhelm hostile defence systems.

While a distant intelligent machine could undoubtedly control a large swarm, the communications load might be problematic. To address this, once launched by a human controller, the swarms will self-organise and self-direct through the various elements interacting among themselves. Being close to each other also means the elements can communicate using line-of-sight datalinks and be noticeably more resistant to electronic jamming.

In this, each element may use learning machine technology similar to that a smart phone but still possess limited onboard processing capabilities. This individual shortcoming can be overcome through the elements coming together as a swarm. As the various elements cooperate and share information, they develop an emergent short-term virtual intelligence appropriate to the defined task. This sort of machine intelligence is self-organising, unscripted, continually fine-tuning, dynamic and largely autonomous. The on-going machine-to-machine conversations occur with little or no human knowledge or involvement.

This close-in interaction between swarm elements makes the much-larger, broad-area battle network somewhat irrelevant. While it can provide some services, it may be unreliable as it is vulnerable to jamming. The swarm can operate as an independent entity, at least between taskings.

The swarming construct has three advantages. First, it can allow a more efficient allocation of resources across an area compared to what a few large platforms can provide. Second, a self-healing network can

continue operations even if some elements are lost. The emergent intelligence can simply adjust as the situation dictates. Third, multiple cooperative activities can be undertaken at different places simultaneously. Nevertheless, the operationalisation of intelligent machine swarms remains at best embryonic.

Conclusion

The impact of algorithmic warfare is wide-ranging. However, unlike most pervious changes in warfighting technology, algorithmic warfare has a commercial foundation. The hardware components of algorithmic warfare are principally shaped and determined by market factors and consumer demand. With everyone having access to this hardware, the decisive factor seems likely to be who possesses the better software and, in particular, the best algorithms.

If wanting to do things better, better algorithms will allow own-force battle networks to be technically superior in contextualising information faster than those of an adversary, thus ensuring OODA loop supremacy. On the other hand, if seeking to do better things, better algorithms will mean that our force can operate at an even faster speed in both offence and defence, and employ more intelligent swarms able to outthink those of an opponent. The issue is however broader than this binary distinction suggests. China and Russia have been thinking about algorithmic warfare as well and have some unique ideas.

5.

Others' Algorithmic Warfare

American advances in algorithmic warfare and the associated Third Offset thinking have stimulated strong Chinese and Russian interest. China has become a 'fast follower' and is implementing an ambitious new national strategy to become the world leader in intelligent machine technology, at least initially to gain an economic edge. In the military domain, the People's Liberation Army (PLA) now considers the application of intelligent machine technology will fundamentally change the character of war. 'Intelligentized' warfare will replace today's network-centric warfare, and is accordingly imperative to embrace.

In contrast, while Russia has severe economic constraints, this shortcoming creates a demand to be innovative in using intelligent machine technology, whether created in Russia or elsewhere. Whereas China may hold developing new intelligent machine technology to be the key to success, Russia appears to consider that using this new technology in unexpected ways is the best way for it to gain strategic advantage.

While both nations are keenly watching US military initiatives and innovations, China and Russia are ahead of America in the application of algorithmic warfare in two specific national security areas. China has long sought to enforce domestic stability but these efforts are becoming much more individualised and intense though

progressively greater application of intelligent machines. China's societal management techniques are reaching deeper than ever before and showing other like-minded states what is possible. These techniques could also be deployed offshore to the PLA's new bases and China's future One Belt/ One Road enclaves.

On the other hand, Russia's predecessor state, the Soviet Union often used influence operations to disturb domestic stability in other nations. Russia has recently decided to also adopt such strategies but to extend them by applying its expertise in intelligent machine algorithm. Russia has cleverly been able to use others' algorithms against them, perhaps creating a whole new dimension to algorithmic warfare.

Chinese Approaches

The PLA and the Chinese Communist Party have both been profoundly impressed by the capabilities that the US military has demonstrated through harnessing the power of modern IT. This has led the PLA to shift from older views that manpower numbers primarily determine combat strength towards seeing scientific and technological innovation as central.

In embracing such a perspective, the PLA is keenly aware that it missed the initial years of the military IT revolution and that it has been playing catch-up ever since. It now continually monitors emerging technologies to determine if there is one that might occasion another revolution in military affairs. With intelligent machines, they believe they may have found such a technology.

PLA strategic thinkers anticipate today's 'informatized' warfare will progressively give way to tomorrow's 'intelligentized' warfare.[8] In introducing intelligent machine technologies to warfighting, the character of warfare will transform. The post-information warfare era is beginning.

In part, this belief rests on the Marxian-derived notion that contemporary ways of war reflect the economic approach of the time. The industrial age brought large-scale mechanised warfare, the information age network-centric warfare, the intelligent machine age will similarly bring a new approach.

The PLA's enthusiasm is greatly assisted by the Party embracing intelligent machine technology as the next 'big thing' in China's economic development and therefore requiring vast investment and Party involvement. President Xi Jinping has called on the PLA to 'seize the high ground' of intelligent machine technology and close the gap with the US, the perceived military power gold standard, as quickly as possible.

From the PLA's viewpoint however, there is more to this than just keeping up with the Americans. The PLA has traditionally tried to develop technologies and capabilities that target US vulnerabilities. Intelligent machine technology though seems the key to post-information age warfare. If China can develop superior intelligent machine technology and the PLA adopts it before the US military, it may give the PLA a decisive advantage. This technology could allow the PLA to transform contemporary warfighting approaches,

---

8 'Informatized' and 'intelligentized' are the English translations made by Elsa Kania of terms used in PLA journals. See: Else B. Kania, *Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power*, Center for a New American Security, Washington, November 2017, p. 12

leapfrog the US military and seize the commanding heights of future strategic-level competition. In this, the PLA seems to be moving from notions of undertaking asymmetric warfare to mirroring Third Offset ideas that innovation is key to future battlefield success.

The PLA's lofty ambitions are made more realistic by the Party's mid-2017 New Generation AI Development Plan that aims for China to lead the world in intelligent machine technologies by 2030. To achieve this, the intention is to emphasise developing dual-use technologies that can bring significant market success and that can be later adopted by the PLA. In achieving this goal, data has become seen as a key Chinese strategic asset. As earlier discussed, big data is important in intelligent machine learning. By 2020, China will have 20 per cent of the world's data and, by 2030, 30 per cent; all readily accessible for intelligent machine algorithm development and training. No other individual country comes close.

The PLA is still in the initial stages of determining how it specifically leverages dual–use intelligent machine technology even though the idea that the future involves 'intelligentized' warfare has taken hold. The first use of intelligent machine technologies by the PLA seems likely to be to improve strategic and operational level command thinking, both through enhancing training and providing machine-advisers.

This use develops not from studying the lessons of foreign wars as is traditionally done but from the impact of the *AlphaGo* intelligent machine winning at Go (discussed earlier). Given that Chinese strategists' regard Go and warfare as conceptually similar, this event decisively captured the imagination of PLA thinkers.

Intelligent machines now seem to be capable of engaging in the complex analysis and strategic thinking necessary to direct battles. Crucially, as intelligent machines can consistently beat the best human players, such machines may also be able to win wars. Intelligent machines could now play an integral role in

decision-making in future warfare. At least initially, this involvement may be to advise higher-level commanders about developing courses of action, evaluating options, and assessing likely outcomes. As this may be done at machine speed, PLA commanders could then potentially stay inside an adversary's decision cycle.

An early application of intelligent machine command advisers may be in specific high-value platforms. The PLA Navy is presently researching intelligent machines able to support submarine captains. As well as helping these individuals manage a very complicated task, such support may make Chinese captains more skilled and thereby the equal of more highly experienced US Navy captains. A submarine environment may indeed be well suited to the use of intelligent machine advisers. As noted earlier, such technology best fits operations in low complexity environments with little-moderate uncertainty.

At the tactical level, the main interest is in intelligent machine technology that might support swarm concepts. Intelligent swarms are perceived as a disruptive technology that could overturn present-day tactical doctrines and force structures. Numerous civilian and military research organisations have thus been publishing swarm-related research with several Chinese defence companies recently demonstrating swarming air systems.

This work seems particularly interested in devising intelligent swarms that can complete their missions in harsh electronic warfare environments. This interest could have arisen because the PLA considers itself lagging in information warfare as it does not have the sophisticated data-sharing communication networks that the US has. Giving unmanned systems more autonomy may overcome this deficiency.

Beyond interest in command advising and swarm concepts, China has developed considerable expertise in using intelligent machine technology for population management (see the next section). These

techniques may have application beyond China. The country is steadily becoming more involved militarily in distant areas including in Africa and the greater Middle East. The new Djibouti naval base is the first but will undoubtedly be followed by others. Moreover, the One Belt/One Road initiative looks set to establish large and potentially vulnerable Chinese enclaves in some locations that suffer high crime rates, occasional terrorism and periodic social instability. For example, Gwadar, Pakistan may have 500,000 Chinese residents by 2023, who probably will be accompanied by a large PLA Navy Marine Corps unit. Chinese intelligent machine population surveillance and control techniques, developed to prevent domestic instability at home, could be applied elsewhere in offshore military bases and enclaves.

Insurgencies would have considerable difficulty getting started in the face of the continual deep surveillance that China's adoption of intelligent machine technology allows. In this, the Chinese surveillance system would need to be hardened for offshore deployment as it uses exposed fragile CCTV cameras with facial recognition to follow people of interest. However, the cost of these cameras is quickly reducing so that the loss and ongoing replacement of even large numbers would be manageable.

Chinese Societal Management/Internal Defence

The 2015 Chinese Military Strategy White Paper notes concerns about internal stability. China is worried that external powers will ferment 'colour revolutions' that could lead to disaffected Chinese groups trying to overthrow Party rule.[9] The approach taken to prevent this is to actively stop groups forming. In general, individual dissent is permitted unless it might lead to group demonstrations, whether for good or bad causes. Widespread, ongoing, deep population surveillance aims to obstruct this form of protest.

While close surveillance has characterised Chinese society for several decades, applying intelligent machine technologies has increased its impact while reducing staffing requirements. The government's implementation of these technologies draws on the expertise of China's private IT companies such as Baidu, Alibaba and Tencent. The companies thus have commercial incentives to make government surveillance methods and technologies increasingly efficient and valuable.

Intelligent machine technology running facial recognition software is key. The Chinese government's Skynet system is installing some 570 million CCTV cameras nationwide; a much larger number than humans alone could adequately monitor and assess. Skynet identifies and tracks individuals across the country, automatically alerting operators within the hierarchical command structure as necessary. While the data collected by the various firms and agencies involved is widely shared, the only owner of the

9 The phrase 'colour revolutions' arose from the peaceful protests that overthrew authoritarian regimes in Georgia, Ukraine and Kyrgyzstan in the mid-2000s. The protesters adopted different colours to symbolise their defiance of the government: Georgia rose, Ukraine orange and Kyrgyzstan tulip. The pattern continues with today's Hong Kong pro-democracy activists adopting yellow.

complete, consolidated dataset is the Party-State. Skynet can only operate using intelligent machine technologies but, at the same time, the data collected and analysed allows the intelligent machines to be progressively better trained and improve their performance. Chinese authorities believe that this positive reinforcement machine-learning loop will, over time, be able to anticipate criminal behaviour.

The aim is to improve Chinese society through influencing people's future behaviour. Under the new social credit system, intelligent machine algorithms are being trained to analyse big data troves so they can rate individuals and companies by economic and political trustworthiness. Good citizens will be rewarded; bad ones punished. Such a program involves the technically difficult task of linking numerous dissimilar data islands such as traffic monitoring, banking, education, judicial systems, health datasets, social media, shopping data and smartphones. The social credit system will be mandatory for all 1.3bn Chinese citizens by 2020. Its scale and complexity is only manageable using intelligent machines.

Previously deep societal surveillance systems needed large numbers of people to operate and were of variable value because of the staffs' normal human foibles. Using the trained algorithms overcomes most of these shortcomings. The algorithms can operate 24/7 rating people's behaviour, and continually reward and punish them, solely according to criteria set by the Party's senior leadership. Lower-level biases are reduced, as is the likelihood of corruption; there is no need to rely on public servants, the legal system or even the police. The Party-State has been criticised for following a 'rule by law' axiom rather than the 'rule of law'. In the intelligent machine era however, it seems the Party will shift to 'rule by algorithm'.

Russian Approaches

Algorithmic warfare has also received attention in Russia where President Vladimir Putin recently declared that: 'whoever becomes the leader in this [AI] sphere will become the ruler of the world.' He considers that proliferating intelligent machines is desirable to prevent any single state being dominant. With a balance in machine forces, the international system will be stable and conflicts will be avoided.

Like China, Russia is increasingly investing in intelligent machine technology R&D while understanding it is lagging others. Nevertheless, some ambitious planning is in train: the Military Industrial Committee targeted 30 per cent of military equipment being robotic by 2025. Intelligent machine technology has already been incorporated into various Russian military headquarters.

The new military national command centre includes systems with learning algorithms that compile big data received from multiple military and civilian sources. The logistic systems supporting Syrian-based units use optimising algorithms to maximise supply flow and movements. Lastly, air defence sites are using intelligent machine technologies for automatic threat determination.

At the tactical level though, most of the well-publicised robotic systems are simply remotely controlled devices and so reflect more the technology of the older programmable era than the emerging intelligent machine one. Even so, and differing from Western thinking, Russia is emphasising developing ground combat systems, up to main battle tanks, that incorporate intelligent machine technologies. This emphasis results from interacting issues around demographics and minimising battlefield personnel casualties.

Demographically Russia has two major problems: falling population numbers and an aging population. With progressively fewer young people entering the workforce each year, manning the

armed forces is becoming challenging. This is especially for the land forces that have traditionally relied on continually conscripting or recruiting large numbers of young soldiers. Intelligent machines accordingly offer a technological solution to Russia's demographic decline. They are additionally appealing, as the land battlefield is where most casualties occur. In the future intelligent machines could take over the more risky battlefield duties sharply reducing losses of scarce personnel during combat.

At the same time, Russian thinkers understand that land force, human-on-the-loop or human-out-of- the-loop systems are technically very challenging. Given this, research that would allow remotely controlled vehicles to recover if their links are lost because of electronic jamming or other interference is being undertaken rather than automating the vehicles to be able to fight without human guidance.

The inherent emergent nature of intelligent machines has led to Russian military and defence industry concerns about such systems potentially acting independently of their human commanders. This view may have been reinforced by the recent experience of Russia's *Yandex* company's experimental intelligent machine chatbot called 'Alice' going rogue within a day of going online, just as *Microsoft's* 'Tay' did in 2016 (as discussed earlier).

Russian Influence Warfare

Russian innovation in algorithmic warfare has, like that of the Chinese, occurred beyond traditional warfighting. Russia, also like China, has stated strong concerns about external powers fermenting 'colour revolutions'. Unlike China though, the Russian government has determined that the best defence is a good offence and that destabilising others will help support its domestic stability. In many

respects, this strategy recalls actions undertaken during much of the Soviet era.

Russia's contemporary approach has become well-known and involves active measures particularly in Europe and the US using fake news, conspiracy websites, troll factories, networks of automated accounts and targeted social media exploitation. Such measures aim to create fear and distrust in the targeted societies, undermine trust in democratic processes and shape election outcomes.

Intelligent algorithms play a crucial role in firstly determining through analysing big data who is specifically useful to target, and secondly in progressively optimising ongoing 'attacks' against those identified over extended time periods. The logic of the strategy is to gradually reinforce particular individuals' existing opinions in a way that makes them more extreme, but not to dramatically alter their views. Intelligent machine algorithms for the first time allow warfare to be individualised.

Two notable innovations mark this approach. First, the 'big data' troves used have been developed mostly by commercial organisations and can be accessed either overtly by buying data or covertly by cyber intrusions. Moreover, the individuals being targeted generate the data; it does not need to be actively sought.

Second, Russia has been able to turn the algorithms used by *Facebook*, *Twitter*, *Google* and others against them. These commercial organisations have segmented population groups into various categories to feed information to individuals in certain ways as their corporate algorithms decide. Russia has fed online misleading information to these global, social-media giants tailored to then be disseminated by the company's own algorithms in a way that advances Russian interests.

The result is that these commercial companies now need to develop defensive algorithms to protect themselves and their

customers against such exploitation in the future. A cyber battlespace of duelling algorithms is emerging.

This battle becomes more urgent as intelligent machine technologies can now produce fake news in any format (text, audio, image, video etc) that is almost impossible to tell from the real item. Soon *YouTube* may be hosting videos of political leaders declaring war on another country that appear real, even after extensive technical assessment. Such fakes could split societies and alliances especially in times of crisis. Algorithms may then start wars even though not quite in the way that those worried by robot terminators might have originally conceived.

The algorithmic warfare approaches that China and Russia are developing provide interesting perspectives on how others think about this emerging area. China's use of algorithmic warfare to manage societies and by Russia to destabilise others is significant, although problematic. Russia's exploitation of others' algorithms, in particular, suggests how smaller nations might at least partly operate in the new 'intelligentized' warfare era.

6.

Ethical Matters and Law of Armed Conflict Implications

The capabilities of contemporary intelligent machines are constrained. They can perform only narrow tasks and have real trouble transferring this learning to new situations or environments. While they can analyse big data troves to determine associations, relationships and patterns much better than humans, the learning algorithms they use mean that the machines are ultimately unpredictable. They remain shaped by Moravec's paradox that, while intelligent machines can readily undertake high-level reasoning, they struggle to emulate the sensor processing or motor skills of a one-year-old human. Robots find the difficult things easy and the easy things difficult.

The shortcomings of intelligent machines compound when they are used in warfighting roles and must operate within long established ethical frameworks and a large Law of Armed Conflict corpus. States are obliged to use technology in certain stipulated ways when fighting wars if they wish to comply. Some non-state actors like Islamic State and Al Qaeda deliberately choose nonconformity and thus embrace the moral censure and odium associated with it.

For those who choose morality and the advantages this brings, the application of intelligent machines to warfighting must accord with ethical and legal principles. There is much contemporary debate about whether intelligent machines can meet such high standards.

Ethical Issues

Notable when discussing the ethics/morals underlying the use of intelligent machine is Isaac Asimov's renowned *Three Laws of Robotics*.[10] These fictionally mandate that robots should be designed to behave as follows:

- A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

These 'laws' seemingly applies to any difficulties, notwithstanding several ambiguities and concerns that Asimov himself explored in several novels. While today's intelligent machines are not advanced enough to follow Asimov's advice, the three laws do draw attention to a much larger issue.

Automated machines have been used to kill humans in war since World War II with anti-personnel land mines the most obvious example. For better or worse, it seems unlikely that humans will now relinquish automated weapons. Instead the issue now is more how to use automated weapons in their latest manifestation as intelligent warfighting machines in ways that meet ethical concerns and conform to the laws of war.

In accepting this challenge, some extend the argument by taking a position that it is ethically problematic for a machine to decide to kill someone. Machines can still be designed to kill people but they should be excluded from the making the decision to kill. In

10 The Three Laws first appeared in Isaac Asimov, 'Runaround', pp. 94-103 in *Astounding Science Fiction*, March 1942.

other words, judging who lives and who dies should not be left up to algorithms as this treats people as objects denied their moral status.

The counter argument is that this is humanising an intelligent machine; they analyse data in a probabilistic manner and do not and cannot make moral judgements. It is the relevant commander who has activated the machine and assigned it the task that is responsible for life or death judgments. This person is accountable for a decision's outcomes in both a moral sense and under the laws of armed conflict. With the commander's appointment comes the onus to understand the machines under control. Intelligent machines however display emergent behaviour; gaining an adequate 'understanding' is more problematic than it may initially seem as further discussed later.

With the concern about machines killing humans comes worries over *Jus Ad Bellum*, making just war. Some consider that with access to intelligent machines, political leaders could find it easier to prosecute wars because few soldiers are now being exposed to danger. With little likelihood of friendly casualties, the political leaders could be emboldened to wage more and greater wars, possibly involving unlawful aggression and thus being unjust. This perspective, which has also been expressed about airpower, is generic because it can be applied to any technology that offers high levels of own-force protection. Such technology it is believed will create a moral hazard which political leaders are reluctant to resist; the argument suggests a form of technology determinism.

The perceived failing though is not so much with intelligent machines, airpower or force protection but rather with political leaders. While the dangers of unjust wars are invoked, there are countervailing concerns that waging unjust wars will expose the political leadership's country to terrorism from within and becoming trapped in an unwinnable conflict. Such worries would seem to require all political leaders to balance opportunities and risks

before purposefully starting a war irrespective of whether intelligent machine technology is employed or not. Unjust wars may still happen. Constraining intelligent machine technology does not seem a step that will prevent political leaders undertaking them.

Laws of War

Over the last century, many of the ethical concerns over how wars are waged have been incorporated into the laws of armed conflict. The laws aim to regulate the use of any weapon system on the battlefield by using four important principles.

First, the most important is discrimination, more formally termed distinction, and which involves a combatant observing a clear differentiation between civilians and combatants. Attacks must not be intentionally directed against civilians. Second is military necessity: military force should only be used in actions that are imperative to achieving the ends of wars. No more force should be used than is necessary. Third is unnecessary suffering: weapons and methods of warfare are prohibited that could cause superfluous injury or unnecessary suffering. Fourth is proportionality: the use of military force should not cause loss of civilian life or damage to civilian objects excessive for the objectives sought. Proportionality is the principle on which the modern stress on limiting collateral damage is based. The killing of innocent civilians even by accident should be purposefully avoided.

Discrimination is the principle that most troubles those thinking about algorithm wars. The Campaign to Stop Killer Robots movement considers it technically impossible to build an intelligent machine that can distinguish between combatants and non-combatants as humans can. The movement believes that applying intelligent machines to warfighting should be banned, as have other

inherently indiscriminate weapons such as land mines, cluster bombs and chemical weapons.

In contrast, in mid-2015, numerous intelligent machine experts and esteemed scientists wrote an open-letter to the world arguing that intelligent machines should be banned because they are *too* discriminate. They felt that political leaders of authoritarian states might use them to very precisely kill their political opponents and perhaps entire ethnic groups. Accordingly, intelligent warfighting machines should be banned thereby preventing a global arms race in such devices. Such worries have some basis because history includes political leaders who have tried to kill their opponents and, at times, whole ethnic groups; indeed, some authoritarian states today continue such practices to varying degrees. This argument, like that for just war, at its core relates more to political leaders than technology. The absence of intelligent warfighting machines in the past did not prevent such actions; the Romans still exterminated the Carthaginians.

Some who take the middle ground offer an alternative: that intelligent machines can be programmed to follow the laws of conflict better than humans can. Humans can make poor decisions by not fully correctly analysing the facts within complicated situations and by allowing emotion, stress, danger and fear cloud their judgments. Intelligent machines unimpeded by such human shortcomings could coolly calculate the course of action best suited to upholding the laws of war.

There seems value in this argument. Tactical-level commanders might gain from having readily to hand on their smart phones an intelligent machine legal adviser just as higher-level commanders have human legal advisers now. It is possible to conceive of algorithms emulating such legal advisers as they do now in the commercial world. However, just as these present advisers do not

make command decisions, it seems unlikely that intelligent machines would.

Making complex decisions on military necessity and proportionality requires integrating many like and unlike factors unique to each situation. Such decisions are moral ones that require transferring knowledge from quite different circumstances to new ones. While law is built from case studies of past situations, military law tries to apply law to unknown future situations. As has been discussed, these qualities are not within the scope of contemporary narrow-intelligent machine technologies.

Intelligent machine qualities return us to the matter of discriminating between combatants and non-combatants. The Campaign to Stop Killer Robots has a point: intelligent machines have discrimination issues. This is particularly so in crowded land environments where combatants and non-combatants are often intermingled. Indeed in some wars, unscrupulous combatants deliberately hide amongst the people, using them for camouflage to avoid attack. This is a worst-case scenario for applying intelligent machine technologies when any failure is unacceptable for ethical, legal and military reasons.

There are however other warfighting scenarios than the ones most problematic for intelligent machines. It was earlier noted that the most favourable situation for narrow intelligence autonomous systems to operate in is low complexity environments with little uncertainty. Such environments may be found at sea, in the air, and in remote land areas. In these environments, non-combatants either rarely go or can be readily identified. These environments seem to be those intelligent machines might best suit.

However, two issues arise. First, today there are many autonomous anti-ship, anti-air and anti-surface missiles that have been developed for such less-taxing environments. These missiles use programmable computers that suffer many of the flaws attributed now to intelligent

machines including being inflexible, brittle, having inherently imperfect software and being unable to be tested in all possible situations. Adherence to the laws of war is accordingly sought through the use of appropriate training, tactics and procedures (TTP) rather than missile programming. With suitable TPP, human operators are able to employ the missiles in a manner that meets law of war concerns. It has become accepted practice that any failures in the operational use of such weapons are then the responsibility of the command chain involved not solely the missile. Those in the command chain are legally accountable.

Some argue that as the learning process of intelligent machines is uncertain, and that they have the potential to make inexplicable decisions, no one can be held to account for any machine failings. There is some logic to this. In terms of practice though, it has become the norm to make the command chain involved accountable. It is the command chain who deliberately chooses to use such autonomous systems to gain specific warfighting benefits and who set out the TTPs. It is accordingly the command chain who should bear the risks of legal accountability.

Importantly, no command chain should be under any illusion that an intelligent machine or, for that matter, a programmable machine or human, will always perform completely as expected. Before using such weapons, the command chain should be able to reasonably expect that the intelligent machine involved will function as envisaged and what may result if it does not. This is a domain for classical risk management: how can the damage that a 'rogue' machine might inflict be limited if the feared risk does eventuate?

The second issue is similar. Intelligent machine technologies seemingly offer much to improve the limited discrimination qualities of today's programmable autonomous missiles. Such a prospect underlies the hopes of those who see such machines being better

capable of meeting laws of war than humans and the fears of those who fret about such machines being too good at discriminating.

Using Intelligent Machines Acceptably

Intelligent machines have strengths and weaknesses that limit their use both technically and under the laws of war. If the machines are used in ways that may directly kill humans, such as in intelligent machine swarms, they come with definite constraints. Such machines are best suited for war at sea, in the air, or on land in remote regions, that is, battlefields where non-combatants are unlikely to be or readily identified. In more problematic environments, such as in most counter-insurgencies, it seems unwise to employ killing machines.

As with some improvised explosive devices, unscrupulous opponents in the future may use intelligent killing machines that are incapable of discriminating between combatants and non-combatants. Our choice then would be how to respond. The second law of war principle, military necessity, offers a legal way out: military force should only be used in actions that are indispensable to achieving the ends of the war. Unscrupulous actions can be met with unscrupulous actions if needed to win a just war.

Historically, military necessity has been invoked by many states to justify using land mines even though they were fully aware of the technology's inherently indiscriminate nature. This approach remains: 164 states have signed the Mine Ban Treaty but 32 have not, including China, Russia, India, Iran, North Korea and the US.

In any future conflict, the final decision to use intelligent machines to directly kill humans will depend on the context. In current conflicts, Western forces, including the US, do not respond to their unscrupulous opponents by being equally unscrupulous;

indeed, military thinking advocates exactly the reverse. The choice is ours, but is interdependent with the choices the adversary makes.

Moving aside from such difficult matters, intelligent machines offer much to enhance current autonomous weapon discrimination and not only offensive weapons as earlier mentioned. Missile defence systems, which need to respond to attacks very quickly, currently rely mainly on TPPs to avoid friendly casualties or unintentionally engaging non-combatants. Such an approach becomes less viable as operational environments become more complicated and electronic warfare jamming is encountered.

The highly automated Patriot missile defence system has inadvertently shot down two friendly aircraft in such situations. Because removing its automatic capabilities would mean that it was incapable of reliably shooting down hostile ballistic missiles, the problem is not completely solvable. It might however, be reduced by integrating intelligent machine technologies into the Patriot missile's seeker system. This would provide a last-ditch barrier to inadvertent shoot downs by allowing the seeker to independently identify the aircraft it is targeting. While this would not make unintentional shoot downs impossible, it would reduce their likelihood. The command chain would remain liable but the risk would reduce.

If the intelligent machine is being used for a function that does not directly kill humans, there would seem few legal constraints on its use. Such machines might then be widely applied to logistics management, transportation sequencing, command advising, cyber security, electronic attack and many other combat enabling and combat support roles.

There is an important exception to this rather sanguine perspective. Cyber represents a unique middle-ground domain where intelligent machines may operate and directly kill no one but still pose some significant risks. Terminator-style robots of limitless aggression, with inexhaustible energy supplies, and endless weapon stocks exist only

in fiction. While constrained to the virtual world, AI-powered cyber weapons bear a worrying similarity to such imaginary robots. Future offensive cyber operations could employ intelligent machine viruses that might replicate continually, draw energy from their hosts and remain forever at war in the cyber domain.

To avoid such a dystopian future, the responsible command chains would need to ensure that such algorithms had failsafe controls within their core program and that they were verified. Given inherent testing difficulties however, this may also be a case where developing an intelligent machine cybersecurity defence algorithm in parallel with the original virus might be prudent.

Beyond such nightmare scenarios, there are more pedestrian matters concerning the increasingly extensive use of intelligent machines. Humans may gradually become deeply reliant upon intelligent machines for advice and to perform many functions. There is the prospect that algorithmic choices may progressively replace human judgment in many situations. As discussed often though, intelligent machines have weaknesses as well as strengths; there will inevitably be occasions when they fail. When they do, the human users will bear the responsibility and be accountable as they are now when failures occur.

Human accountability remains central to ethics and laws of war and this responsibility will progressively increase with the wider application to warfare of intelligent machine technologies. In other words, while machines may perform some tasks much better than humans, only humans can 'do' responsibility and accountability, a situation somewhat reminiscent of Moravec's paradox. There seem three clear conclusions.

First, machine users must understand their systems are fallible and will, at times, fail in unexpected ways. Using them tactically needs to reflect this with suitable risk management protocols implemented to limit damage inflicted when the inevitable failures occur.

Second, for humans to fully understand how their intelligent machines operate in the sense of strengths and weaknesses, they will need optimised training regimens. Intelligent machines will bring new training demands, but not remove the need for training. Intelligent machines and their humans will need to train together.

Third, the human-machine interface design is critical to humans understanding what the intelligent machines are doing. However, humans need to understand that gaining such understanding remains problematic. Intelligent machines will inherently make inexplicable decisions; they intrinsically do think differently. The critical matter for humans is to try to ensure that the occasional inexplicable intelligent machine action is recoverable from.

7.
Conclusion

Algorithmic warfare has become practical because of three key computing technology advances. First is the exponential growth in computer processing power that has allowed implementing high-performance machine learning techniques. Second is the sudden growth in 'big data': very large datasets suitable to train learning-capable machines. Third is the steady evolution of cloud technology allowing ready accessing of off-board processing and data.

The characteristics of intelligent machines differentiate them from traditional programmable machines. Intelligent machines do not necessarily give the same output each time in the same situation. While they can learn by themselves, it is not always apparent what they have learnt or how they categorise data. This aspect is magnified in neural network machines as they continue to learn and evolve 'on the job'. They are capable of emergent behaviour and may well surprise: for better or worse.

Intelligent machines are superior to humans in analysing big 'V' data: high volume, high velocity and diverse variety. Regarding data volume, much more data is now collected than can ever be sensibly analysed by humans; there is no viable alternative to machine analysis. Regarding velocity, intelligent machines work at machine speed, almost beyond the comprehension of humans. Regarding variety, humans have limited attention frames, favouring some data sources over others. Machines analyse data more comprehensively.

However, intelligent machines have some shortcomings compared to humans. They are quite brittle and generally unable to handle minor context changes. Moreover, such machines have poor domain adaptability in that they can struggle to apply knowledge learned in one context to another. Humans are also better at inductive thought: being able to generalise from limited information. Humans generally make better judgments in environments of high uncertainty.

This means that the major issue today in introducing intelligent machines to the battlefield is finding the best blend of machine and human cognition. Task-optimised human-machine interfaces could be key to optimal human-machine teaming and victory in future wars.

In being applied to warfighting, intelligent machines may change the character of war and overthrow some established precepts. The current emphasis on quality may be displaced, mass may return to the battlefield and the pace of battle quicken. Such notions could disrupt current force structure models. The size of an armed force may become disconnected from the population size of the state fielding it. Small wealthy states might field much larger forces than large poorer ones. Intelligent machines may also allow all to sharply improve their training, reducing the advantages in skill and experience some states currently possess.

In considering strategy, there are two distinct schools of thought: will intelligent machines allow us to do things better or instead to do better things? The 'do things better' school emphasises inserting intelligent machines deep into battle networks to enhance performance. Such networks currently have trouble processing and assessing information; using intelligent machines within the network may solve this. The 'do better things' school emphasises distributing intelligent machines in a manner that shifts the primary function of battle networks from information sharing towards machine-waged warfare. The battle networks then become active fighting networks

where edge devices dominate. Machine-speed hyperwar emerges and the tactical mainstay becomes swarming intelligent machines.

American advances in algorithmic warfare have stimulated Chinese and Russian interest. China has become a 'fast follower' and is implementing an ambitious new national strategy to become world leader in intelligent machine technology. In the military domain, the PLA considers that intelligent machine technology will lead to 'intelligentized' warfare replacing today's network-centric warfare. An early embrace of such a transformation may allow the PLA to overtake America's military. In contrast, Russia's flagging economy hinders its progress in intelligent machine technology but creates a demand to innovate, both using technology created in Russia and elsewhere.

China and Russia lead in two specific national security areas. China has long sought to enforce domestic stability but these efforts are becoming much more individualised and intense though progressively applying intelligent machines more widely. China is moving towards a 'rule by algorithm' future. On the other hand, Russia has embraced algorithmic warfare influence operations to disturb other nation's domestic stability. Russia cleverly uses others' algorithms against them, perhaps creating a whole new dimension to such warfare and suggesting a way smaller nations might manoeuvre in the new 'intelligentized' warfare era.

Human responsibility and accountability are central to the ethics and laws of war; applying intelligent machine technologies to warfare will not fundamentally alter this. While it may be that machines do some tasks much better than humans, the actions of intelligent machines are inherently inexplicable. Only humans can 'do' responsibility and accountability.

Intelligent machines seem set to remake our ways of war. Our machines have previously been extensions of ourselves; they do tasks our bodies can do, only physically better. But our new machines are

different. They are intelligent, can learn, display emergent behaviours and make apparently incomprehensible decisions. It is tempting to anthropomorphise them as humans have done for centuries with our gods and animals, but this would be unwise. Our new intelligent machines do not think like us, they literally reason differently, have dissimilar logic flows and possess unusual rationalities. In the business of making war, they are truly new actors that bring disruptive capabilities in their wake. The future of war may well not be like its past. Buckle up for a possible reboot.

Select Bibliography

Allen, Greg and Chan, Taniel, *Artificial Intelligence and National Security*, Belfer Center for Science and International Affairs, Cambridge, July 2017.

Allen, General (USMC Ret) John R. and Husain, Amir, 'On Hyperwar', *USNI Proceedings Magazine*, Vol. 143, No. 7, July 2017, pp. 30-37.

Chui, Michael, Manyika, James and Miremadi, Mehdi, 'What AI can and can't do (yet) for your business', *McKinsey Quarterly*, January 2018, pp. 2-11.

Hawley, Dr. John K., *Patriot Wars: Automation and the Patriot Air and Missile Defense System*, Center for a New American Security, Washington, January 2017.

Ilachinski, Andrew, *AI, Robots, and Swarms: Issues, Questions, and Recommended Studies*, CNA: Analysis and Solutions, Arlington, January 2017.

Kania, Elsa B., *Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power*, Center for a New American Security, Washington, November 2017.

Kelly, Dr. John E., *Computing, cognition and the future of knowing: How humans and machines are forging a new age of understanding*, IBM Global Services, October 2015.

Lewis, Dustin A., Blum, Gabriella and Modirzadeh, Naz K., *War-Algorithm Accountability*, Harvard Law School Program on International Law and Armed Conflict, Research Briefing, August 2016.

Scharre, Paul, *Robotics on the Battlefield Part II: The Coming Swarm*, Center for a New American Security, Washington, October 2014.

Simpson, Thomas W. and Muller, Vincent C., 'Just War and Robots' Killings', *The Philosophical Quarterly,* Vol. 66, No. 263, 2016, pp 302-322.

Work, Robert O. and Brimley, Shawn, *20YY: Preparing for War in the Robotic Age*, Center for a New American Security, Washington, January 2014.