AUTONOMOUS WEAPON SYSTEMS FOR THE LAND DOMAIN

SUBMISSION TO THE CONCEPT FOR ROBOTICS AND AUTONOMOUS SYSTEMS 2040
BY PETER A. MILANI

> "... he who uses force unsparingly, without reference to the bloodshed involved, must obtain a superiority if his adversary uses less vigour in its application. The former then dictates the law to the latter and both proceed to the extremities to which the only limitations are those imposed by the amount of counteracting force on each side."
>
> von Clausewitz[1]

INTRODUCTION

What government would send Australians to fight a mostly automated enemy? What commander would risk casualties from close combat against robots? The sacrifice of a single soldier fighting robots, churned out by their hundreds and replaced nearly immediately, will surely prompt the question: Is the fight worth the cost? If the purpose of war is to "impose one's will on an adversary" [2], and victory is about "defeating an enemy's will to fight" [3], the deployment of Autonomous Weapon Systems (AWS) immediately disrupts a Center of Gravity (COG) based on the value of people. Australia is vulnerable to such disruption, and as illustrated by the quote above, liable to be subordinate to our adversaries if we cannot fight such threats with a greater intensity and freedom of action. This means that fielding AWS needs to be central to our strategy.

The premise of this paper is that future war will devolve to AWS. The timeline for the empty or fully autonomous battlespace is likely measured in decades. However, early AWS demonstrators, with significant limitations in targeting, state estimation and navigation, can be developed now. The long term goal will be to ultimately supplant human involvement in close combat and other dangerous tasks.

The key questions to be asked are whether remotely operated systems hold an advantage over manned systems and whether fully automated systems hold a similar advantage over remotes? If the answers to those questions are affirmative, it can be shown that the best response in an adversarial scenario is for both sides to develop AWS. The side that does not develop AWS, is at a significant disadvantage.

SCOPE

---

[1] von Clausewitz, *On War*, 101.
[2] VCDF, "Future Operating Environment 2035," 4.
[3] ADF, *LWD 1 Fundamentals of Land Warfare*, 41.

This paper will examine AWS as they may be applicable to the Land Domain. It will examine ethical requirements and then define some of the technical background to AWS, characterising Artificial Intelligence (AI) and Robotics and Autonomous Systems (RAS). The ways AWS may be exploited will be examined, followed by consideration of lines of effort that enable Australia to support the development and sustainment of AWS. Finally we look at countermeasures that are available to land force elements to disrupt and dislocate AWS and protect the human capital that forms our COG.

DEFINITIONS

Where possible, the definitions in the Robotic and Autonomous Systems Strategy (RASS)[4] are used throughout this paper. The term Autonomous Weapon System[5] is added. An AWS is any UGV, UAV or UxV which is armed with a kinetic weapon, and has the ability to use it without a human directly approving each engagement.

The levels of autonomy given on page 27 of the RASS, namely the Automatic, Autonomic and Autonomous classifications are difficult to specify clearly. In that example there was perhaps too much specification of how machine autonomy was achieved, and not specific enough about the tasks being automated. For example a Roomba is a highly autonomous floor cleaner, it is very reactive, applying standard heuristics rules, but is robust in a dynamic environment and requires little human interaction. It doesn't fit neatly into the autonomy specification given in the RASS[6]. As a result we contribute a classification scheme for AWS in annex B based on SAE International's levels of driving automation for on-road vehicles[7]. This attempts to be a little more specific about the handover of tasks between human and machine, but it does not specify the mechanism of implementation or the performance standards required to be met.

ETHICAL ISSUES

AWS immediately provokes ethical questions that need to be addressed. The International Committee of the Red Cross (ICRC) has published findings on the limits of autonomy in weapon systems[8]. They found that while AWS offered certain military advantages over unmanned systems, unpredictability in their outcomes presented challenges to safety and efficiency in military operations. They found that humans needed to maintain control over three main categories:

[4] Smith, *Robotic and Autonomous Systems Strategy*, 28.
[5] "Limits on Autonomy in Weapon Systems | ICRC," viii.
[6] Smith, *Robotic and Autonomous Systems Strategy*, 27.
[7] "Automated Driving."
[8] "Limits on Autonomy in Weapon Systems | ICRC," 36–37.

1. **Controls over the weapon systems parameters**. This includes temporal and spatial limits on AWS operation to constrain their effects, and an allowance for failsafe and deactivation mechanisms.

2. **Controls on the Environment**. Use in environments where distinction in targets can be achieved, taking into account both the environment and the performance of the AWS. This is also dependent on the capability of the AWS to perform accurate discrimination between combatants and non-combatants.

3. **Controls through human-machine interaction**. Users must be able to supervise AWS and intervene in its operation, by overriding functions, aborting a task, or deactivating AWS.

These are interesting constraints on the use of AWS that can be applied with different priorities depending on the capabilities of the system. For example, an AWS with an inferior target identification system may only be unarmed, armed with non-lethal measures or constrained to environments not containing civilians. One of the key findings is that ethical employment depends on the technical capability of the robot and also its employment[9] rather than whether AWS presents an inherent good or evil.

Ron Arkin, a veteran roboticist and robot ethicist[10], argues[11] that AWS can provide a net benefit to the ethical conduct of war. Outperforming humans in their ability to adhere to International Humanitarian Law (IHL), since robots have no inherent instinct to survive. This eliminates emotional responses such as fear, guilt and revenge that can trigger excesses in the use of force and other criminal behaviour. LWD 1 also identifies fear as one of the reasons humans make mistakes[12]. Additionally, AWS have the potential to objectively assimilate new information conditioned in prior experience that is not affected by combat stressors. Potentially producing a rational and consistent decision-making process, something humans find very difficult[13].

Professor Arkin also identifies some prevailing counter arguments to the employment of AWS. The most notable argument is a reduced threshold for entry into conflict between states. In the distant future, if no citizen need be a casualty in war, what will be the barriers to conflict? Answerable to their citizens, will democracies' wars of the future fight only to the last robot, human casualties being too expensive: politically, socially and literally? Will these AWS enable authoritarian figures to dispense with the trappings of democracy: universal suffrage, healthcare and education when citizens hold no value as combatants? Or will these

[9] "Limits on Autonomy in Weapon Systems | ICRC," 37.
[10] "Ronald C. Arkin."
[11] Arkin, "Ethical Robots in Warfare."
[12] ADF, *LWD 1 Fundamentals of Land Warfare*, 15.
[13] Kahneman, *Thinking, Fast and Slow*, 411–14.

things be even more important as a country's fortunes are increasingly tied to industrial output, as opposed to raw manpower, requiring a skilled and engaged workforce?

Multiple Tradeoffs. The path this paper charts is full of ethical dilemmas. There is no simple solution. The ethical employment of AWS centers around the risk of failure in a particular robot configuration and whether that risk can be managed to not violate IHL. It also depends on the character of the humans commanding AWS, whether they too will abide by the IHL. Considerations include: the characteristics of the AWS platform, the task and risk to non-combatants. Technical considerations and a description of the characteristics of AWS will now be the focus in the rest of this paper.

TECHNICAL BACKGROUND

It is not the goal of this paper to go deeply into the technical foundations of field robotics. Instead, it will give sufficient insight that can inform and characterise some of the limitations that are inherent in RAS and AWS, and to understand the capabilities that currently exist. Annex A of this submission provides a little broader treatment of the key technical areas, but for a deeper treatment the footnotes and bibliography contain good references.

Characteristics of Artificial Intelligence

Artificial Intelligence (AI) are the ways and means of making a machine respond to its environment, without human intervention. For this paper I am considering AI that does a single task well, such as object detection, navigation or state estimation, and provides an output in a useful timeframe. These tasks form individual building blocks that must be strung together and deployed to a machine to create autonomy. This is in contrast to an artificial general intelligence, a capability that is currently science fiction[14]. The characteristics of AI are:

- **Learned models are blackboxes.** Humans can conceptualise the relationship between two things easily: we can draw a graph. However our ability to reason about the relationship between multiple parameters rapidly deteriorates as the number of parameters increases beyond three. Most modern AI utilises tens of millions of parameters for tasks like image recognition. This huge dimensionality makes it hard to understand how an AI function reaches its outputs, how correct the output is and what its outputs will be for any given input. At the moment, the best we can reason about is the dataset the algorithm is trained on, what the distribution of data points is in that dataset and how well the algorithm performs on a similar but unseen validation dataset.

[14] "3 Reasons We Are Far From Artificial General Intelligence."

- **AI is compute intensive.** Parameters are just numbers that must be operated on, that is: added, subtracted, divided or multiplied. The large numbers of parameters discussed above must be operated on multiple times for any one inference. Depending on the algorithm, this computation can be ported to a parallel processor, or done though some efficient mechanism on a cpu, or just take a long time to compute.

- **Data is paramount.** All algorithms require data as an input. Without data, an algorithm cannot be refined or optimised. Data is used, in the case of online navigation, in accordance with some rules of a physical model. Or it can be used for learning appropriate outputs for given inputs (supervised learning), or used to find relationships between inputs (unsupervised learning). Finally it may just be a bunch of reward experiences from actions undertaken whilst following some policy (reinforcement learning). Data is paramount, without good data there is no AI.

- **Metrics should be multiple and defined upfront.** The performance of any AI must be measured. Unfortunately due to the fact they are mostly blackboxes, reasoning about their function, explaining how they come up with their outputs can be very difficult. Heuristic methods have less of a problem as they are explicitly designed and not learnt. However both need to have metrics or conditions that can be used to measure their output performance. Typically, performance against a known output or ground truth is used to evaluate algorithm performance.

- **Generalising to every situation is very hard.** There is usually a distribution of examples that is considered during generation of the algorithm, whether it be by design or learning or a combination. This takes into account the assumptions made during design, the distribution of examples in the training data etc. Algorithms can become very good at producing usable responses to similar situations, however if an algorithm meets an unanticipated or novel situation, its output can often be completely inappropriate.

Characteristics of AWS

Mathematics. All autonomy is mathematics. All inputs are reduced to numbers, operated on by algorithms that are collections of basic mathematical operators, and all outputs are numbers. These numbers represent images received by a camera, points from a lidar, or frequency of voltage changes going to motors. Numbers represent everything important to the robot, comprising the entirety of its knowledge. Therefore if there is any desired autonomy, its fundamental nature must be reduced to numbers and algorithms.

Measured uncertainty. Probabilistic methods are used almost universally in state estimation. This gives not only an expected value of a robot's state, but also the probability of the robots' other states. This is useful when fusing unexpected data. If uncertainty is measured, the

degree to which the new data should be integrated into the system is determined by probability theory.

Behaviours. What behaviours are required by AWS? Typically heuristic behaviours are hierarchical: simpler one-dimensional motor schemas are driven by more complex, task based behaviours to determine navigation. Behaviours can be combined in many different ways to produce a complex autonomy. This includes navigational behaviours to move to a goal, follow a path, avoid obstacles, de-collide from obstacles[15], scan the environment for targets[16], hide, and cooperate with other robots to achieve an objective. This includes learned behaviours[17].

State Estimation[18]. Determining the location of a robot in space is necessary for higher level autonomy and basic movement. State estimation can also produce maps as part of the process. State estimates suffer from noise, random perturbations in value, and bias. Most state estimation techniques try to estimate both these values, which they use to infer a most likely pose, as well as uncertainty of that pose. A robust state estimate should also respond well to the kidnapped robot problem[19], where the robot is required to relocalise after being transported to a new location.

Targeting. Currently the best object detection, recognition and identification algorithms are based on Deep Convolutional Neural Networks (DCNN). They are black boxes in that they are explainable only with reference to the training distribution, which state of the art methods can generalise quite well. Their performance deteriorates badly on data that is not represented well by the training dataset, and they are sensitive to high frequency noise.

Communications. Robots are capable of generating large amounts of data. The main challenge of human-on-the-loop implementations, will be to physically transmit the data required for someone to review and intervene. Balancing the data needs, the link bandwidth, whilst also being robust to topography is challenging. Mesh, relay and satellite networks provide benefits in overcoming the topological constraints of UHF communications. Though as a matter of principle, robots should not be reliant on their comms and non-RF communications should be considered as a backup.

Natural Language Processing (NLP) and speech-to-text processing of voice commands may also be an option where appropriate. Current systems require processing that is usually in excess of the capabilities of edge devices such as robots. Typical implementations require a link to cloud-based computation to be realised. State of the art methods rely heavily on

[15] Arkin, *Behaviour Based Robotics*, 66–120.
[16] Arkin, 270–302.
[17] Arkin, 306–28.
[18] Thrun, Burgard, and Fox, *Probabilistic Robotics*, 20.
[19] Thrun, Burgard, and Fox, 194.

DCNNs and suffer from similar drawbacks as articulated in the section on Targeting and in annex A.

EXPLOITATION OF AWS

The AWS preference in the short to medium term, is for small lightweight robots greater than 20kg but probably less than 100kg. These may be armed, and given their size and capabilities, probably only with small calibers. The weaponization of the system is kept in line with the shortcomings of the key functional capabilities of targeting, navigation and behaviours, such that the risk is minimised if a failure occurs. If the robots can't be trusted, then they shouldn't be capable of great and far-ranging destruction.

The principles of concentration of mass and combined arms teams will continue to be important in automated forces. Whilst human-machine teaming is one way to achieve this, a single AWS type will not be sufficient or invulnerable when operating alone. In the same way that mounted and dismounted forces complement each other, a range of AWS will need to be fielded to meet differing situations. This implies a balance between offensive and counter AWS capabilities. This paper will consider how AWS could be fielded within the decade to separate humans from the most dangerous of tasks.

Close Combat. Firstly in close combat we could expect a single soldier to control multiple robots. Controlling multiple robots takes a lot of work, even when autonomy helps basic robot functions. When supervising five robots, an operator is not capable of doing much else, but commanding and supervision.

This means that for the deployment of robots by ground forces, say an infantry section sized human grouping (8 people). There would need to be the following grouping:

- **Maneuver Team.** Possibly up to half the human grouping would be in the direct C2 of robot teams, with no hierarchical grouping of robots this would be about four to five robots per person, nominally 20 AWS deployed. They are responsible for the marshalling and supervision of the robots through the mission.

- **CSS Team.** Robot-human teams would be augmented by automated resupply. This could include automated battery changing of AWS, unloading, startup and launching sequences. However there will be other functions not easily automated such as maintenance, and weapon replenishment. There may be up to two persons in this role.

- **C2 Team.** The C2 team would backup the supervision of robots and maintain the link to higher HQ. They provide additional spotting and oversight of the deployment through UAV and direct observation. They may also manage the communications across the group by ensuring relay communications nodes are located for maximum effectiveness. It would include the section commander and one other.

It is not expected that the humans would take any part in the close combat, but instead would choose a secure assembly area to control their team. The AWS would form the bulk of the assault grouping. Other, flanking human elements would provide Support by Fire and other tasks. If they were to form an assault grouping, for safety, it would be on a separate axis.

The section will require integral transport, including medium trailers to carry all the equipment and robots. We could envisage an automated, containerised solution for loading, unloading, battery swapping, and replenishment. It may also require integral power generation for recharging batteries etc. Alternatively these could be replenishment items with the recharging/refueling occurring at First or Second Line CSS.

Amphibious Operations. AWS would be ideal to fill the first wave assaults in contested landings. We can look to the many examples throughout WWII in the Pacific Theatre as why we would only want to send humans in after a sufficient beachhead has been secured. As a Joint activity, we could consider the land component similar to the Close Combat example above, with the difference that the human supervisors would be sitting over the horizon in an Amphibious Ship such as the Landing Helicopter Dock (LHD). In this instance, satcomms would be the only viable method of feedback on the landing.

The ship to shore delivery mechanism may be automated. Automated landing craft may deliver dozens of AWS per load, traversing from the LHD to the Beach Landing Zone (BLZ). These automated landing craft may return for reload, or move off to the side to make way for follow on landing craft. In a GPS denied environment, localisation of the landing craft becomes a challenge. Many of the techniques available to ground based RAS do not work on the open ocean. Automated Celestial Navigation[20] would be an avenue to explore for localisation. Closer to the BLZ, visual based methods could take over to localise against the BLZ plan and bring the landing craft to the beach for disembarkation of multiple AWS.

Screens and Guards. Ground AWS can have a lower detection signature and may hold an advantage over UAVs in not revealing our force posture. AWS could provide mobile screens to detect and report on enemy movements from positions considered too precarious to station human forces. Alternatively, they could provide additional security to forward deployed forces, through early warning and delay.

AWS could also be employed to assist human Intelligence, Surveillance and Reconnaissance (ISR) elements to break contact. Just before being decisively engaged, a human screening force could deploy AWS in defendable locations. These AWS could defend or delay depending on the capability of the autonomy, allowing human forces to keep on withdrawing to the main body.

[20] "Automated Celestial Navigation for the Navy."

These applications can obviously be applied for the security of any isolated force element. Armoured Fighting Vehicle (AFV) hides, Gun Lines and a lot of other tasks can all benefit from greater sensor coverage, situational awareness and combat power. The ability of AWS to assist in security during rest and refit periods would be a key benefit.

Rear Area Security. Command and Control (C2) and Combat Service Support (CSS) elements are usually underprotected, typically because the manning associated with C2 and CSS are preoccupied with their primary function. The space required to fit a CSS element is many times larger in comparison to what a similarly-manned infantry unit would be expected to defend. This usually means much of the perimeter is unsecured.

AWS could be used for Screen and Guard tasks around CSS and C2 elements. They are also good locations as initial integrators of RAS. The relatively static nature of second line support areas and access to power, makes them ideal as an early adopter and demonstration location.

Conclusion. As AWS improve in capability and usefulness, they will grow to fulfill most tactical tasks. The list above targets applications that are particularly dangerous, or such an unsolved problem that AWS will immediately fill the gap. However, these exploits would be enhanced with action along lines of effort not related to the tactical employment of AWS. We next describe the lines of effort that will help form AWS into a deployable capability.

ENABLING LINES OF EFFORT

In this section, auxiliary steps to ensure the timely development and sustainment of AWS are considered. These are activities that can start immediately.

Data Collation.

As articulated, data is king when it comes to AI, particularly Machine Learning. A concerted effort to generate, manage and collate data useful for AI development should be undertaken ADF-wide. For AWS Army's data would be most valuable, for example:

- **Image data required for target recognition.** Most units contain a unit photographer, these members should be constantly increasing the size of the database with imagery, during every training activity. They should be tasked with collecting imagery off-site to help reduce bias in the dataset. ADF personnel sometimes represent a latent workforce that could assist with labelling to those datasets, which is usually a somewhat tedious task.

- **Datasets for decision making.** Reinforcement Learning is a fairly data inefficient method for learning new behaviours. Using human experience to seed the RL policies is a well documented approach. Army in particular has significant professional

education methods during career courses and in-unit training focusing on tactical decision making. Digitising these tactical exercises without troops (TEWTs) and quick decision exercises (QDEs), will provide great examples for AI and also a valuable training tool for professional development.

- **Capturing instrumented exercise data.** The Combat Training Center - Live (CTC-L) instruments its exercises so that most of the activities in the exercise can be examined later as part of a learning activity. The captured data contains information on individual and vehicle movements, weapon discharge and casualties sustained. As this data is already digitised it would be a great source to understand issues such as terrain, weapon effects, tactical behaviours and logistics.

This is not an exhaustive list, but it should be noted that training the AI algorithm is only a small part of the overall effort to get a good autonomous response. Data management, collation and critical analysis are required to ensure that we understand and can trace the dataset for the AI. This leads to an overall better chain of evidence, and to understand what an AI can do and why it fails.

Constant RAS exercising

The ADF should aim to undertake constant exercising of autonomous systems by industry and research organisations. This includes activities like Autonomous Warrior 18, but also more frequent biannual activities. This doesn't just relate to the defence prime contractors, but to anyone who can field a robot. As indicated in the DARPA SubT Challenge, a large design space exists and maximising the number of participants is good way to understand the challenges. The DARPA Offensive Swarm Enabled Tactics (OFFSET)[21] executed this in a series of six-month sprints, across virtual and systems domains to address issues around the tactics, techniques and procedures for exercising autonomous swarms.

Foster greater industrial capability

The loss of Australia's manufacturing base over the last 40 years challenges our ability to support a long term, robotic enabled force. Conceivably, stockpiles of compute, actuators and battery types could be made, however, we could end up with obsolete parts in a short time, and shelf life is important. Apart from software and power systems, it is hard to find Australian companies providing components necessary for robotic platforms.

Component Classes from outside Australia

[21] "OFFensive Swarm-Enabled Tactics (OFFSET)."

- Compute and other semiconductors are sourced from a very few global companies like Intel, Nvidia and Samsung from silicon foundries in the USA, Taiwan, South Korea, and Japan [22].
- Sensors such as cameras, inertial measurement units (IMUs), lidars (excluding Baraja[23])
- Electric actuators, such as motors, servos or linear drives.
- Hydraulic actuators except for hydraulic cylinders.
- Power sources, internal combustion engines.

Component Classes from inside Australia

- Robot chassis.
- lithium batteries[24], hydrogen fuel cell technologies[25]
- Power electronics.[26]
- Software.
- Engineers and Researchers.

In general, a greater investment in RAS technologies should foster greater demand in the above components. Most of the components will come from overseas. There needs to be a whole of government approach to ensure sufficient industrial capacity to maintain sovereign control of our AWS. Alternatively, excellent supply lines to Japan, Korea, Taiwan and the USA need to be maintained.

COUNTER RAS

At the dawn of the tank during the Somme in 1916, German infantry quickly learned the machines were impervious to their weapons. However, after the initial abortive uses of tanks gave away the strategic surprise, the Germans innovated on tactics and used whatever weapons they had on hand: arming with SmK rifle ammunition that had effects on tank armour, forward deploying field artillery batteries in the direct fire role, innovating with the anti-tank ditch and flooding no-man's land[27]. Subsequently, they developed AT rifles. There was very little in anti-tank capability initially, but it was innovated from existing equipment and performed reasonably well until tanks were used in mass and in combined arms teams.

It took 20 years and a world war for tank design to converge to the basic design balance of armour, mobility, firepower and communications in use today. 103 years on from their first employment we could ask for a different technological threat: what do we actually have now

[22] "List of Semiconductor Fabrication Plants."
[23] Baraja, "Baraja | Spectrum-Scan™ LiDAR Technology for Self-Driving Cars."
[24] "□sonnenBatterie."
[25] "Fuel Cell Industry Developments in Australia and New Zealand."
[26] "REDARC Electronics: Automotive Electronics, Accessories & Equipment | REDARC Electronics."
[27] Guderian, *Achtung-Panzer!*, 60.

that can assist in the defeat of a RAS attack. Several possibilities for tactical land combat spring to mind.

Greater All-corps Anti Armour Capability. Any ground based robot fielded will likely be resistant to current small arms fire. If massed AWS are expected, single shot throwaway weapons such as M72 will not provide the weight of fire required for multiple targets. Javelin systems may similarly be overwhelmed by the number of targets. The broad scale equipping of the force with 40mm HEDP natures would provide a relatively cheap, effective and persistent response, with standoff and penetration beyond that of 5.56mm natures. For AWS it is unlikely that they will be equipped with a survival instinct, that is they will not be able to be pinned by effective fire. Whilst tactical autonomy should be expected, only direct hits will disrupt an AWS attack. A greater use of anti-materiel rifles would also be suitable additions to the counter-RAS suite, though are limited by size and utility in other tasks.

Short-range anti air capability. Aerial AWS also poses a threat. Survivability and payload in these systems is sacrificed for maneuverability and traversability. A marsupial application, where a UAV is launched for an attack from the back of a UGV, can mean troops can face a very selective, aerial adversary capable of hunting below the treeline. Such systems can move quite nimbly, and a 40mm canister or splintex-type nature should be developed for use in bringing down these Air-AWS. Shotguns would be a decent backup.

Increased Obstacle building focus. Unless facing full-size weapon systems such as optionally crewed AFVs, physics would dictate that wire and ditches would form obstacles for lighter weight AWS. Netting would be suitable obstacles for defeating Aerial AWS. Smoke and camouflage will interfere with the operation of sensors such as cameras and lidars. Also, unless equipped with sonar or bump sensors, cameras and lidar have a difficult time detecting glass.

EW Attacks. Electronic Warfare (EW) needs to extend to wider bands beyond Radio Frequencies (RF), to optical and algorithmic countermeasures. Communication is a feature of robots, they are also laden with other electronics. Traditional RF-EW has a role in detecting and tracking robotic formations, finding C2 centers. Though RF may be low power and below the noise threshold at any decent range.

Outside RF, detecting NearIR laser light from lidar equipped systems can identify their location, as well as sound from sonar equipped systems. Electronic Attack across the spectrum could be used to destroy optical sensors by blinding them with lasers, and potentially interfering with other sensors in the vehicle. GPS spoofing or jamming will be a core function to eliminate the use of basic COTS platforms.

Finally Deep Convolutional Neural Networks (DCNNs) are sensitive to small high frequency perturbations in the input. In image recognition, these are so called adversarial attacks[28]. They cause the algorithm to incorrectly classify objects in images. Newer personnel and vehicle camouflage patterns could incorporate some of these features. Typically such attacks are highly dynamic, and the "newer camouflage" may just consist of replaceable stickers as the attack evolves.

Distance. If possible, trading space for time may be a viable alternative. One of the unstated characteristics of RAS is that they fail. So if possible, give them more room and time to get lost, stuck, run out of power or otherwise suffer mechanical or electrical failure. It is likely they will do so at a greater rate than a human or animal adversary.

CONCLUSION

AWS must be developed so that Australian forces are not faced with the dilemma of sacrificing people or the mission fighting an adversary armed with AWS. The ethical employment of AWS must allow for discrimination between combatant and non-combatant. Where this cannot be achieved by autonomy to the necessary precision, at least physical separation of combatants from non-combatants is required. Finally the behaviours that govern the AWS' actions need to be commensurate with the tasks given.

AWS main applications in the land domain will be for close combat and screening of vulnerable force elements. As their autonomy grows, so will the range and complexity of tasks that can be assigned. Crippling limitations in target discrimination will curb any widespread deployment of very heavily armed systems, so the majority will usually be fairly lightly armed. Counter-AWS measures could be achieved through 40mm natures for both ground and air types, and more widespread deployment of anti-material rifles.

AWS will perform best when fielded in combined arms teams. Currently, human machine teams will be tightly integrated. As Autonomy improves so too will the hierarchical nature of teams. There is great potential to exponentially increase the combat power available to individuals when wielding AWS. The reliability and predictability of state estimation, behaviours, target discrimination and communications just need to live up to that potential.

ANNEXES

A. Extended Technical Background
B. AWS Levels of Autonomy

[28] Haohui, "Adversarial Attacks in Machine Learning and How to Defend Against Them."

BIBLIOGRAPHY

"3 Reasons We Are Far From Artificial General Intelligence." Accessed June 19, 2020. https://www.sicara.ai/blog/artificial-general-intelligence.

Pathmind. "A Beginner's Guide to Word2Vec and Neural Word Embeddings." Accessed June 11, 2020. http://pathmind.com/wiki/word2vec.

ADF. *LWD 1 Fundamentals of Land Warfare*. Australian Army, 2014.

Arkin, Ronald. *Behaviour Based Robotics*, 2002.

Arkin, Ronald C. "Ethical Robots in Warfare," n.d., 4.

"Automated Celestial Navigation for the Navy." Accessed June 14, 2020. https://www.maritime-executive.com/blog/automated-celestial-navigation-for-the-navy.

"Automated Driving." SAE International. Accessed July 1, 2020. https://cdn.oemoffhighway.com/files/base/acbm/ooh/document/2016/03/automated_driving.pdf.

Baraja. "Baraja | Spectrum-Scan™ LiDAR Technology for Self-Driving Cars." Baraja. Accessed June 29, 2020. https://www.baraja.com/.

Clausewitz, Karl von. *On War*. London: Penguin, 1982.

Davison, Neil. "Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control." ICRC, August 2019.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *ArXiv:1810.04805 [Cs]*, May 24, 2019. http://arxiv.org/abs/1810.04805.

Fuel Cell & Hydrogen Energy Association. "Fuel Cell Industry Developments in Australia and New Zealand." Accessed June 29, 2020. http://www.fchea.org/in-transition/2019/5/20/fuel-cell-industry-developments-in-australia-and-new-zealand.

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Networks." *ArXiv:1406.2661 [Cs, Stat]*, June 10, 2014. http://arxiv.org/abs/1406.2661.

Guderian, Heinz. *Achtung-Panzer!* London: Cassell Military Paperbacks, 1937.

Haohui. "Adversarial Attacks in Machine Learning and How to Defend Against Them." Medium, December 19, 2019. https://towardsdatascience.com/adversarial-attacks-in-machine-learning-and-how-to-defend-against-them-a2beed95f49c.

James, Stephen, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. "Sim-to-Real via Sim-to-Sim: Data-Efficient Robotic Grasping via Randomized-to-Canonical Adaptation Networks." *ArXiv:1812.07252 [Cs]*, July 21, 2019. http://arxiv.org/abs/1812.07252.

Kahneman, Daniel. *Thinking, Fast and Slow*. Penguin books, 2011.

Lambert, Ben. *A Student's Guide to Bayesian Statistics*. London: Sage, 2018.

Lambert, Nathan. "Convergence of Reinforcement Learning Algorithms." Medium, April 9, 2020. https://towardsdatascience.com/convergence-of-reinforcement-learning-algorithms-3d917f66b3b7.

———. "Gists of Recent Deep RL Algorithms." Medium, January 12, 2020. https://towardsdatascience.com/getting-just-the-gist-of-deep-rl-algorithms-dbffbfdf0dec.

Lee, Michael, and Eric-Jan Wagenmakers. *Bayesian Cognative Modelling*. Cambridge University Press, 2014.

"Limits on Autonomy in Weapon Systems | ICRC." Accessed June 9, 2020. https://www.icrc.org/en/document/limits-autonomous-weapons.

"List of Semiconductor Fabrication Plants." In *Wikipedia*, May 31, 2020. https://en.wikipedia.org/w/index.php?title=List_of_semiconductor_fabrication_plants&oldid=960024452.

Matuszek, Cynthia, Evan Herbst, Luke Zettlemoyer, and Dieter Fox. "Learning to Parse

Natural Language Commands to a Robot Control System." In *Experimental Robotics*, edited by Jaydev P. Desai, Gregory Dudek, Oussama Khatib, and Vijay Kumar, 88:403–15. Springer Tracts in Advanced Robotics. Heidelberg: Springer International Publishing, 2013. https://doi.org/10.1007/978-3-319-00065-7_28.
Nielsen, Michael A. "Neural Networks and Deep Learning," 2015. http://neuralnetworksanddeeplearning.com.
"OFFensive Swarm-Enabled Tactics (OFFSET)." Accessed June 19, 2020. https://www.darpa.mil/work-with-us/offensive-swarm-enabled-tactics.
"REDARC Electronics: Automotive Electronics, Accessories & Equipment | REDARC Electronics." Accessed June 22, 2020. https://www.redarc.com.au/.
Rocca, Joseph. "Understanding Generative Adversarial Networks (GANs)." Medium, August 25, 2019. https://towardsdatascience.com/understanding-generative-adversarial-networks-gans-cd6e4651a29.
"Ronald C. Arkin." In *Wikipedia*, December 29, 2019. https://en.wikipedia.org/w/index.php?title=Ronald_C._Arkin&oldid=933095709.
Smith, Robin. *Robotic and Autonomous Systems Strategy*. Australian Army, 2018.
"sonnenBatterie." Accessed June 29, 2020. https://sonnen.com.au/sonnenbatterie/.
Sutskever, Ilya, Oriol Vinyals, and Quoc V Le. "Sequence to Sequence Learning with Neural Networks," n.d., 9.
Tangiuchi, T., D. Mochihashi, T. Nagai, S. Uchida, N. Inoue, I. Kobayashi, T. Nakamura, Y. Hagiwara, N. Iwahashi, and T. Inamura. "Survey on Frontiers of Language and Robotics." *Advanced Robotics* 33, no. 15–16 (August 18, 2019): 700–730. https://doi.org/10.1080/01691864.2019.1632223.
"Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_201806." SAE International, June 15, 2018. https://www.sae.org/standards/content/j3016_201806/.
Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics*. MIT Press, 2005.
VCDF. "Future Operating Environment 2035." Text. The Cove, March 21, 2017. https://cove.army.gov.au/article/future-operating-environment-2035.

EXTENDED TECHNICAL BACKGROUND

Artificial Intelligence

AI is a catchall phrase to describe any method by which a man-made object does something smart and acts in order to achieve some goal in the world. In broad terms, AI for robotics can be divided into subsets such as Heuristic Methods, Machine Learning, Natural Language Processing, Reinforcement Learning and Probabilistic Methods.

Heuristic Methods

Heuristic or Rule-based methods are common approaches used to achieve autonomy in robots. The algorithm and its parameters are determined by design based on principles of physics and maths. They are engineered by humans to meet the requirements and specifications that are required by the system. Whilst determinism for known input is shown, most heuristic methods can manifest emergent behaviour, due to the interaction of multiple heuristics or new environments providing unforeseen situations. They are used for control functions, behaviour generation and navigation[29] to name a few applications.

Machine Learning (ML)

Machine Learning covers a large number of algorithms whose parameters are determined by data. There are two broad families based on parameter optimisation:

1. Supervised. Desired outputs for a given input are provided to an optimiser which aims to minimise the error between the outputs of the algorithm and the example data.
2. Unsupervised. The algorithm deals with unlabeled data sets and tries to group or cluster the data so that Like Data is close and Unlike Data is distant.

Deep Convolutional Neural Nets (DCNN or Deep Learning) are mentioned specifically due to the utility in achieving state of the art in many ML tasks. It is important to note that most of the algorithms were developed in the 1980s. Their current resurgence and performance is due to availability of extremely large quantities of labeled data, availability of generally programmable parallel processors. Which makes them accurate and fast to execute compared to other methods.

They are also a universal function approximator in that they can be used to approximate any general function[30], provided enough training examples are provided that accurately describe the distribution of the function inputs to outputs. Outside that training distribution, their

[29] Arkin, *Behaviour Based Robotics*.
[30] Nielsen, "Neural Networks and Deep Learning."

performance is undefined. Typically when training, the learned model is evaluated on a set of inputs not in the training distribution, this is used to evaluate how the model can be expected to perform on new data. They are routinely applied across AI in areas such as Image Processing, Reinforcement Learning, Natural Language Processing. But as a general function approximator they can or have been applied nearly everywhere sufficient labeled data is available.

DCNN can be sensitive to high frequency pertubations in the input signal, so called Adversarial Attacks [31]. Such changes have been shown to cause the DCNN to fail. For example, in an image classification task, it could cause misclassification of the image. Early Adversarial Attacks required knowledge of network architecture as well as the parameters, however there are some black box attacks, which require no specific knowledge. It has been shown that a suitable remedy to innoculate a network against adversarial attacks includes filtering images, and training the network in an adversarial environment.

The methods above are optimisations, where the objective is to try to minimise the error against some metric. There is another method that instead uses a game between a generator and a discriminator to learn the distribution of objects from data. The Discriminator tries to learn to differentiate between true and fake images, minimising the loss from errors in that task. The Generator generates fake images from random noise such that it tries to fool the Discriminator, for an inverse objective. These are the Generative Adversarial Networks (GANs)[32]. They provide a new way of learning a distribution and the concept of a game between the Generator and Discriminator is an architecture that is yet to be fully exploited[33].

Reinforcement Learning (RL)

Reinforcement Learning aims to learn from experience in order to maximise the future reward for a selection of possible actions, given the current state[34]. They can be formulated into many different architectures which work on related concepts[35]. A key part is to try and anticipate the reward for a given situation and action (the Q function). This is usually achieved by measuring the state, applying actions according to the best option given by the existing Q function and measuring the rewards returned. There is a period between sets of these trials in which the Q function estimator is optimised to improve its prediction of the reward. The updated Q function is then applied in another epoch of experiences.

Current RL methods require many training epochs because rewards do not occur after every action, they may only occur after several hundred actions, e.g. when the agent wins, loses or scores a point. This makes them impractical to implement on real robots as the number of

[31] Haohui, “Adversarial Attacks in Machine Learning and How to Defend Against Them.”
[32] Goodfellow et al., “Generative Adversarial Networks.”
[33] Rocca, “Understanding Generative Adversarial Networks (GANs).”
[34] Lambert, “Convergence of Reinforcement Learning Algorithms.”
[35] Lambert, “Gists of Recent Deep RL Algorithms.”

examples required exceeds usual physical reliability limits. Most implementations that target robots do so by simulation, speeding up the time taken to converge, and eliminating reliability issues. As the simulated environments do not replicate the real world environment accurately, most RL policies need to handle the simulation limitations. Usually the cheapest method is by randomising most of the parameters of the simulation, so that when confronted by the real world, its differences to the simulated environment appears as additional noise to the policies it has learned.

Using simulation greatly reduces the number of real-world tuning examples required and, if the limitations of the simulation have been addressed, have produced good results in domains such as robot grasping[36] and robot control[37].

Probabilistic Methods

Probabilistic methods aim to estimate parameters from data not as point estimates with certain values, but as an underlying distribution of possible values, observed at various times as measurements by a sensor. Probabilistic methods estimate both the expected value of the parameter (mean) and the uncertainty in the parameter value (variance). As new information becomes available, it is used to update the estimate for the parameter[38].

These methods are powerful in that they can not only estimate what is known, but also give a measure for how well it is known . This quality of knowledge is given by the variance in the distribution. They do rely on the model being explicitly designed for inference which can make explaining the outputs of the model easier, and can improve the predictability of the outputs. However as these models scale up, dimensionality makes them harder to interpret and compute. However as the uncertainty is explicitly stated, often a simpler model, with uncertainty measured will provide just as usuable a model, as one that is more complicated.

A key benefit is that the models can contain parameters which are important but not directly observed, so-called hidden variables. Other parameters which are observed will have some relationship to the hidden variable as specified by the model. As observations are made, likely states for the hidden variable can be estimated that are consistent with both the observation and the model. A military example might be the estimation of an enemy's location, mission and intent, based on the observation of receiving fire by a particular weapon system in a particular direction, given an enemy situational template, and other prior knowledge.

Probabilistic methods are widely used for robotic state estimation, sensor fusion and map building[39], target tracking to name but a few applications, but is also widely used across

[36] James et al., "Sim-to-Real via Sim-to-Sim."
[37] Arkin, *Behaviour Based Robotics*, 310–20.
[38] Lambert, *A Student's Guide to Bayesian Statistics*.
[39] Thrun, Burgard, and Fox, *Probabilistic Robotics*.

science, politics, economics and social science[40]. These methods are computationally intensive but not easily parallelizable which can make inference computationally intense.

Natural Language Processing

NLP is the processing and factorisation of spoken or written language by a machine. Typically approaches aim to find relationships of words that go together [41]or to predict the next word given a string of words[42]. There are various challenges in processing key terms in natural language for relationships in statements, understanding what that language actually implies and processing in or parsing it into commands that a robot can respond to[43] [44]. Additionally NLP architectures based on deep learning are much more expensive to compute, due to the requirement to keep context and memory, making them hard to deploy to low power environments[45].

More Info

The ICRC has released a thoroughly excellent article[46] on the technical aspects of the application of AI and Robotics to AWS. Addressing whether the current state of the art in AI and Machine Learning is sufficiently robust enough to deploy to AWS. They draw on commercial experiences in the development of self-driving cars, and the ability of the current technology to be certified as being predictable and reliable.

[40] Lee and Wagenmakers, *Bayesian Cognative Modelling*.
[41] "A Beginner's Guide to Word2Vec and Neural Word Embeddings."
[42] Sutskever, Vinyals, and Le, "Sequence to Sequence Learning with Neural Networks."
[43] Tangiuchi et al., "Survey on Frontiers of Language and Robotics."
[44] Matuszek et al., "Learning to Parse Natural Language Commands to a Robot Control System."
[45] Devlin et al., "BERT."
[46] Davison, "Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control."

Annex B To
AWS for the Land Domain

AWS LEVELS OF AUTOMATION

This table is based on the levels of automation specified in SAE International's J3016 Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems[47]. It aims to separate the autonomy description from the methods used to achieve autonomy. It also specifies what that autonomy looks like against various functional areas, the breath of scenarios it is valid for and also quantifies the frequency of human supervisory involvement and its nature. It is likely that a single system may contain aspects of a variety of autonomy levels.

Lvl	Name	Narrative	Behaviours	Decisions	Targeting	Failure Recovery	Scenario	Supervisory oversight	Human interaction frequency
0	No Automation	Full time performance of human crew driving all aspects of the machine.	Human	Human	Human	Human	n/a	A machine function (crew position)	constant
1	Control Systems Only	Optionally Crewed and crew assistance automation (eg Tank Fire Control Systems)	System controls	Human	Human	Human	Some scenarios	1 machine	constant
2	Partial Automation	System takes over local navigation, senses the environment, human sequences techniques (GoTo) to achieve a task	System implements tactical techniques.	Human specifies tactical techniques	System acquires, Human approves, System engages	System basic recovery, Human fallback	Some scenarios	< 5 machines	minutes
3	Conditional Automation	System implements tactical techniques as specified by humans to meet mission.	System implements tactical tasks	Human specifies tactical tasks	System engagement conditional on predefined parameters	System more robust to failure, Human fallback	Some scenarios	machine team	tens of minutes
4	High Automation	System can sequence tactical tasks to meet the mision and respond to changes in situation	System sequences tactical tasks to achieve mission	Human specifies mission.	System engagement conditional on task	System automated recovery from failure	Some scenarios	teams of teams	hours
5	Full Automation	System can sequence tactical tasks to meet the mision and respond to changes in situation	System sequences tactical tasks to achieve mission	Human specifies mission.	System engagement conditional on mission	System automated recovery from failure	All scenarios	teams of teams	tens of hours

[47] "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_201806."