ARL Battlespace: a platform for developing novel AI for complex adversarial reasoning in MDO

Chou Hung, James Zachary Hare, B. Christopher Rinderspacher, Walter Peregrim, Sue Kase, et al.

Chou P. Hung, James Zachary Hare, B. Christopher Rinderspacher, Walter Peregrim, Sue Kase, Simon Su, Adrienne Raglin, John T. Richardson, "ARL Battlespace: a platform for developing novel AI for complex adversarial reasoning in MDO," Proc. SPIE 12113, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications IV, 121130X (6 June 2022); doi: 10.1117/12.2622227

Event: SPIE Defense + Commercial Sensing, 2022, Orlando, Florida, United States

ARL Battlespace: A Platform for Developing Novel AI for Complex Adversarial Reasoning in MDO

Chou P. Hung *[a], James Zachary Hare [b], B. Christopher Rinderspacher [a], Walter Peregrim [a], Sue Kase [a], Simon Su [a,c], Adrienne Raglin [b], John T. Richardson [a]

[a] U.S. Army Combat Capabilities Development Command - Army Research Laboratory, Aberdeen Proving Ground, MD, USA 21005

[b] U.S. Army Combat Capabilities Development Command - Army Research Laboratory, 2800 Powder Mill Rd, Adelphi, MD, USA 20783

[c] National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD, USA 20899

ABSTRACT

A barrier to developing novel AI for complex reasoning is the lack of appropriate wargaming platforms for training and evaluating AIs in a multiplayer setting combining collaborative and adversarial reasoning under uncertainty with game theory and deception. An appropriate platform has several key requirements including flexible scenario design and exploration, extensibility across all five elements of Multi-Domain Operations (MDO), and capability for human-human and human-AI collaborative reasoning and data collection, to aid development of AI reasoning and the warrior-machine-like interface. Here, we describe the ARL Battlespace testbed which fulfills the above requirements for AI development, training and evaluation. ARL Battlespace is offered as an open source software platform (https://github.com/USArmyResearchLab/ARL_Battlespace). We present several example scenarios implemented in ARL Battlespace that illustrate different kinds of complex reasoning for AI development. We focus on 'gap' scenarios that simulate bridgehead and crossing tactics, and we highlight how they address key platform requirements including coordinated MDO actions, game theory and deception. We describe the process of reward shaping for these scenarios that will incentivize an agent to perform command and control (C2) tasks informed by human commanders' courses of action, as well as the key challenges that arise. The intuition presented will enable AI researchers to develop agents that will provide optimal policies for complex scenarios.

Keywords: Artificial Intelligence; Human-AI Collaboration; Decision-Making; Adaptation; Wargaming

1. INTRODUCTION

As part of the DOD's AI strategy[1, 2], the Army Research Laboratory is developing research programs and technologies based on the Human Systems Adaptation strategy, including the goal of developing superhuman capabilities based on human-AI team decision-making and mutual adaptation. These new capabilities are necessary to address the Army's Multi-Domain Operations (MDO) strategy [3], particularly its Penetrate and Dis-integrate phases during which AI-enabled decision aids can augment the commander's ability to tackle the high velocity and volume of information and the complex dynamics of the ground, sea, air, space, and cyber domains. A key challenge is that existing AI algorithms, including leading AI algorithms that are focused on specific problems in AI learning, are wholly inadequate for *complex* decision-making and fail to generalize to MDO-relevant scenarios. Another challenge is that existing Army processes for doctrine and decision support do not integrate AI into the military decision-making process (MDMP) [4], and this gap is just beginning to be addressed by the Army's Automated Planning Framework (APF) [5]. Third, existing theories and technologies for human-AI team decision-making are inadequate, with very limited ability to provide AI transparency for complex decisions in depth, in which multiple dependencies, uncertainties, and information domains and actors intersect with complex human, materiel, and environmental dynamics. They also have limited ability to synergize with the tacit reasoning of human experts. Developing these capabilities requires an integrative and multidisciplinary research approach, including the development of AI testbeds for novel AI research and human-AI teaming.

* chou.p.hung.civ@army.mil Approved for public release; distribution is unlimited

Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications IV,
edited by Tien Pham, Latasha Solomon, Proc. of SPIE Vol. 12113, 121130X
© 2022 SPIE · 0277-786X · doi: 10.1117/12.2622227

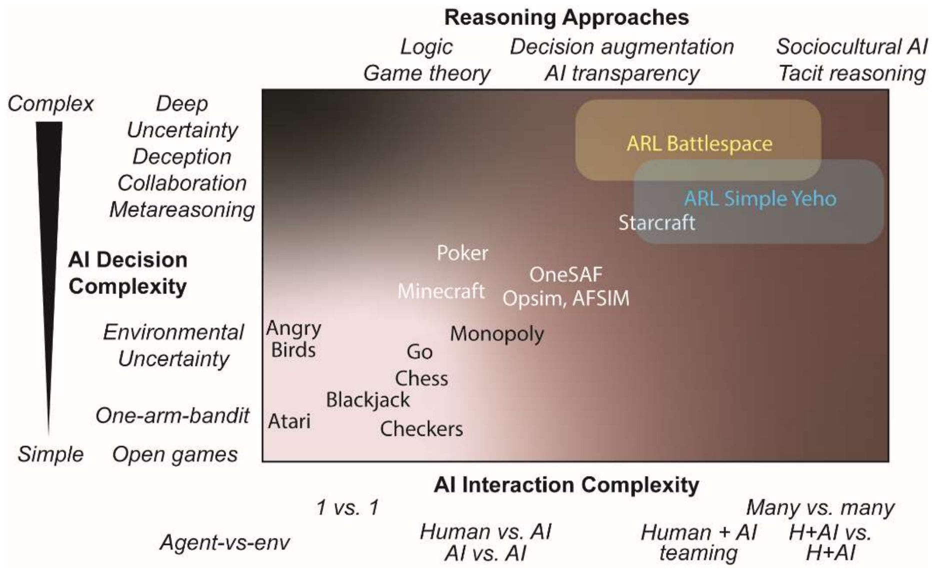

Figure 1. ARL Battlespace within the broader AI research strategy

AI research efforts can be plotted along two axes: one axis for increasing levels of AI Interaction Complexity leading to human agent teaming for decision-making, and a second axis for increasing levels of AI Decision Complexity under increasing levels of uncertainty, collaboration, deception, and metareasoning. At the bottom left corner are games like checkers, chess, and go, where the agent has access to the true state of the environment but is challenged by increasing degrees of decision depth. At the top right corner are wargaming testbeds like ARL Battlespace, in which multiple collaborative and deceptive strategies are used under high degrees of uncertainty, and new approaches are needed for effective teaming with tacit reasoning and AI transparency.

For wargaming, it is necessary to develop testbeds that can model decision-making across multiple echelons including tactical and strategic levels. Existing wargaming decision tools such as Opsim[6], AFSIM[7], and OneSAF[8] can model and simulate many factors across multiple scales to predict outcomes based on strategies, materiel capabilities, and resources, but they suffer from the limitations of aging systems that can be difficult to learn for experienced Soldiers and that are not well suited for developing AI and human+AI teaming capabilities. The recent rapid rise in AI capabilities opens up research into the development and incorporation of novel AIs as decision aids for wargaming. Recent improvements in AI reasoning, e.g. based on deep reinforcement learning, have been based on ‘open’ games in which the state of the environment is perfectly known (e.g. checkers, chess, and go) [9]. They are also based on limited cooperativity or deception. Even in cases with additional complexity such as environmental uncertainty (Angry Birds[10], Atari[11]), there is limited decision complexity, flexibility, and transferability to multiplayer wargaming (e.g. Poker, Minecraft, Starcraft, Figure 1) [12, 13]. Although these models can explore decisions in depth, they are limited to conditions in which the potential values of choice outcomes can be easily measured and quantified. Wargaming environments pose a difficult and unaddressed challenge for AI learning, because of the many sources of information uncertainty, not just from the environment but also from the human and AI agents. AIs need to adapt to changing rules and strategies, to rapidly mitigate unexpected hostile capabilities and exploit new opportunities and friendly capabilities[14]. AIs also need to mutually adapt with their human teammates, and they need to have capabilities for tacit reasoning, to synergize with human experts and for compensating for individual biases and heuristics[5] and changing cognitive states. Unlike classical approaches such as game theory, where the expected utilities of future states can be explicitly quantified for limited sets of actions depending on cooperative or non-cooperative choices, wargaming raises the possibility of interactions across environmental and social dynamics (including cooperativity and deception) and across multiple spatiotemporal scales and domains, which confound the AI’s ability to learn how decisions tie to future state values[15].

Addressing this gap requires a sustained foundational research effort with experiments focused on discovering principles and developing new algorithmic approaches for specific problems in decision-making, and the capability to tie these principles and algorithms back to MDO wargaming. For example, in complex situations with imperfect knowledge and uncertainty[16, 17], an AI that provides a landscape of near-optimal solutions may be more helpful than one that provides

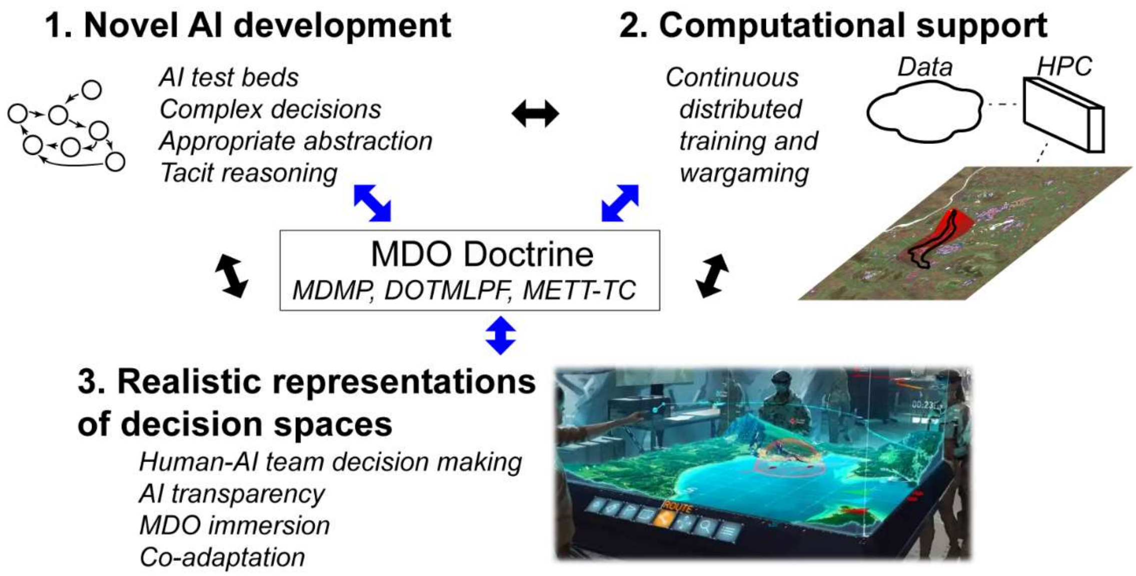

Figure 2. Novel AI development within a broader human-agent team decision-making research strategy
Novel AI development is one of three areas of development needed to develop AI-enabled decision aids for Multi-Domain Operations. It is also necessary to develop the computational support to enable continuous individualized training and wargaming, and to develop novel interfaces and theories for human+AI collaborative decision-making. The development of these capabilities will enable anytime-anywhere individualized training across multiple echelons, including individualized AI-enabled red-teaming capabilities. Continuous improvements across all three areas of development is expected to advance Army doctrine across tactical and C2 levels.

a single ‘optimal’ solution[18]. How this problem-solving ties to AI transparency also needs to be explored[19, 20]. Experimentation of conditions such as near-optimality and uncertainty, with new warfighter machine interfaces (WMIs) can lead to new algorithms, universal tools, and principles that better synergize the human+AI exploration of complex decisions[21, 22].

To create these capabilities, we have been developing two complementary AI testbeds for complex reasoning: ARL Simple Yeho, an AI testbed for tactical ground operations, and ARL Battlespace[23], an AI testbed for MDO command-and-control (C2) decision-making. ARL Simple Yeho is focused on environmental realism, with multiple map layers including roads, foliage, and elevation based on the Yehorivka Unreal environment, as well as realistic unit capabilities and mission tasks, to better align AI reasoning and transparency to the tacit reasoning of expert Soldiers. It is expected that algorithms developed in Simple Yeho will be able to support tasks such as route planning and Solder retasking, by recommending actions to a platoon commander. Because of the focus on the local terrain environment, the AI reasoning developed in this environment will be focused on fine-scale social and ecological dynamics, with limited opportunities for in-depth training on collaborative and hostile decision dynamics. Conversely, the ARL Battlespace AI testbed abstracts away the elements of local terrain, to focus the AI learning and reasoning more specifically on *complex* MDO-relevant C2 reasoning in depth (multiple decision steps including more frequent opportunities for collaboration and deception). Both testbeds develop AI capabilities for complex multi-agent (human, AI, and human+AI team) decision-making under realistic wargaming contexts. Coupled with computational support and development of mixed reality interfaces for decision-making (e.g. the Battlespace Visualization and Interface (BVI) platform, based on the Augmented REality Sandtable (ARES) platform[24]), it is expected that these AI and human-AI teaming efforts will lead to advancement and modernization of multiple Army doctrines (MDMP, DOTMLPF[25], METT-TC[26]) (Figure 2).

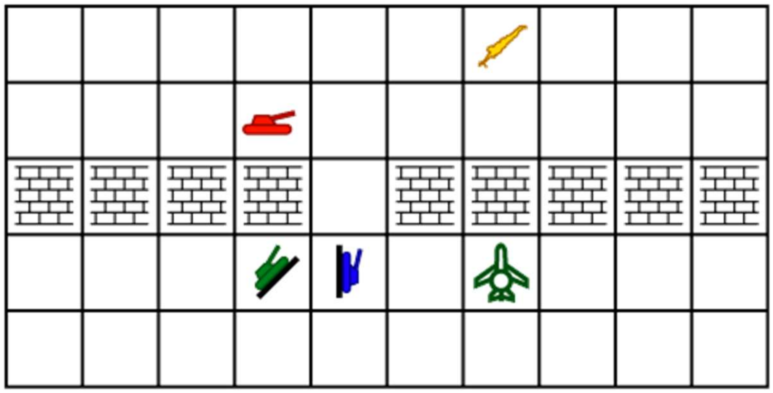

Figure 3. Breaching scenario with game theory and uncertainty

An example scenario of a forced decision requiring communication and cooperativity, with multiple opportunities for introducing decision complexity and examining how uncertainty dynamics affect the multi-agent decision landscape.

2. COMPLEX REASONING SCENARIOS IN ARL BATTLESPACE

Here, we focus on the ARL Battlespace AI testbed and outline multiple possible scenarios to challenge and develop AI reasoning for complex C2 decision-making. ARL Battlespace focuses on multi-player decision-making to develop AI reasoning strategies under cooperation and competition. The game consists of two teams, where each team can comprise multiple human or AI players, played sequentially in turn across a network (Team 1 player 1 is followed by Team 1 player 2, etc, followed by each of the Team 2 players). The game state is updated simultaneously after all players have entered their moves. Each team is analogous to an alliance, while each player represents an actor. The game is played on a common battlespace that is designed as a set of boards for each domain of the MDO. A standard board consists of 10×10 squares and multiple ground and air units, all of which may be adjusted as needed. Each unit has a limited observation distance of one adjacent square, so the agents have limited observability of hostile forces, and shared knowledge of other friendly unit positions and their observed hostile positions is dependent on maintaining friendly communication. All units are capable of firing missiles, airplanes can bomb, and ground units can ram the square ahead. Damage results from colliding with a missile, with another unit, or with a wall. We focus in particular on coordinated actions by the friendly team, the actions of the hostile team, and how these decision dynamics relate to classic and unresolved gaps in AI decision-making and decision theory.

2.1 Breaching Scenario and Classic Game Theory

We begin by focusing on the gap between game theory and wargaming in a simple breaching scenario, which is a classic problem in wargaming that is often encountered, e.g. at bridge crossings, mine fields, and mountain passes (Figure 3). In the classic game theory concept of Brinksmanship ('chicken'), the friendly blue and green tanks are incentivized to cross the gap to reach the other side. Normally these tanks would coordinate their actions, but if the communication between the blue and green tanks is disrupted, the action of one unit (e.g. the blue tank) may lead to low payoff due to collision or friendly fire with another unit (the green tank). The scenario rapidly advances beyond classic game theory if it also includes elements of Prisoner's Dilemma, as it may be necessary for both the green and blue tanks to cross together to jointly attack the stronger red tank, requiring careful coordination. The presence of additional units (e.g. the green airplane providing observation, bombing, or jamming of hostile units such as the yellow Soldier providing possible reinforcement) enables further manipulation of dynamics and environmental constraints or opportunities on the decision-making. The airplane may also discover a second gap, or the 'wall' may be permeable to create gaps (e.g. clearing the mines or establishing additional bridge crossings). Behaviors learned at a coarse scale (e.g. 10×10 board) and context can be gradually generalized to finer scales and other contexts via reward shaping. Additional map layers can also be added for domains such as rapid underground transport, to bypass walls in the ground layer. Environmental factors such as weather can also be included to alter maneuverability. Thus, even an apparently simple scenario can provide rich opportunities for manipulating factors that affect decision-dynamics and outcomes, and for exploring how interactions

Approved for public release; distribution is unlimited

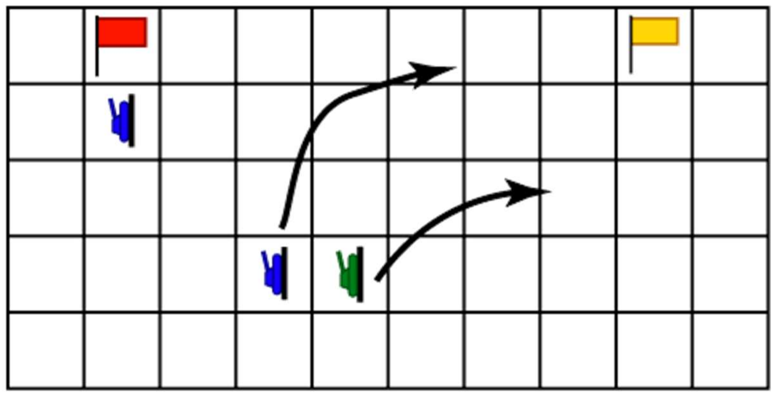

Figure 4. Metareasoning flag scenario with retasking

A scenario in which a metareasoning agent, sensing that the red flag is already nearly captured, redirects the remaining tanks to search for the second (yellow) flag.

across different types of uncertainty can alter the decision landscape to create saddle points and local minima that can confound efforts at reinforcement learning. Understanding and predicting Nash equilibria for three or more cooperative and adversarial players, under scenarios that are likely to emerge in warfare, requires a flexible wargaming platform that allows for the interdisciplinary exploration of such decision spaces. The wargaming platform would also need to enable the development, understanding, and discovery of novel interactions and synergies between the players and the AI that enable the human to use the AI to quickly find optimal and near-optimal solutions. These solutions would enable the AI to learn from human patterns of decision-making and how to optimize its search of decision space.

2.2 Metareasoning Scenario and Mission Objective

In the ARL Battlespace game, each player has a colored flag, and the game can be won by either annihilating all opposing ground units or by capturing all flags of the opposing team (a real-life equivalent is capturing all the key bridges or command centers). Depending on the state of the game, a commander may decide to alter the overall strategy (annihilation vs. capture-the-flag) to achieve the win more quickly. For example, if one tank is already nearing one flag, it may be advantageous to redirect the remaining units to search elsewhere for the remaining flag (Figure 4). For AI development, this would require that an additional, higher-level reasoning agent be constantly monitoring the state of the game, to make choices about when to switch strategies and to communicate this to the agents controlling the individual units. Incorporating metareasoning can allow constraints and various decision making approaches to be used to provide different options for courses of actions. Alternative metareasoning based choices could decide whether to prioritize exploration versus attacking known hostile units, or decide whether to prioritize attack versus defense, and which maneuvre strategy to deploy given the observable positions of hostile forces. Because of the small grid size of the ARL Battlespace environment, the games can be played quickly, resulting in frequent opportunities for metareasoning to be used and opportunities for AI to learn to combine and predict interactions across multiple types of metareasoning approaches. Because the abstract environment increases the frequency of opportunities for the AI to learn how strategies interact, this would enable AIs to learn higher order strategies such as the need to balance interactions across strategies, capabilities, and task requirements, to maintain freedom of choice and to produce strategic ambiguity to confound the opposition. Overall, the benefit of this approach is the improvments to decisions by adding the control and the monitoring mechanism that comes with including a metareasoning agent that balances the actions and the environmental constraints.

2.3 Simple Deception and AI Theory of Mind

A key aspect of adversarial decision-making, particularly in warfare, is deception. Deception can occur across multiple levels including strategy, observable information, and unit capabilities and locations. The limited observability of units in ARL Battlespace naturally creates opportunities for deception, and the capability of airplanes to observe deep in hostile

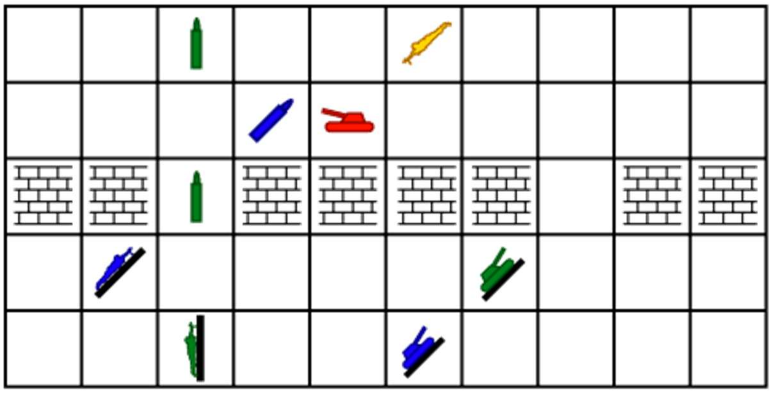

Figure 5. Simple deception scenario with AI theory of mind

A scenario in which the blue and green units deceive the red and yellow units by firing missiles through the left gap, to bias their expectations away from the right gap.

space provides opportunities to uncover deception about unit positions. Figure 5 illustrates an example of a simple deception scenario, in which the friendly blue and green units attempt to cross to the other side. The friendly Soldiers at the lower left begin by firing missiles through the left gap, because their agents reason (via AI Theory of Mind of the opposing agents[27]) that, upon seeing the missiles, the hostile agents will infer that the friendly forces are preparing to attack through that gap. This deception, by focusing the hostile agent's attention and planning to the left gap, biases them away from the right gap and creates an opportunity for the blue and green tank to enter from the right. By designing the scenario with two gaps, the scenario builds upon the two-alternative-forced-choice tasks of classic psychology, enabling the application of sensitive psychological tools for decision analysis and the development of animal models for neurophysiological and behavioral dissection of the underlying cellular and molecular mechanisms that govern context-dependent learning and decision-making. For example, one could introduce factors to bias the friendly or hostile decision-making (e.g. by manipulating the noisiness of sensors or by manipulating commands from headquarters), or apply methods such as optogenetics and chemogenetic tools to understand how the neural representation of others' perceptions, beliefs, or strategies (e.g. in the anterior cingulate and orbitofrontal cortex) contribute to decision-making computations (in the prefrontal cortex). Such investigation could also uncover factors that determine single-mindedness, heuristics, and implicit bias vs. openess to alternative hypotheses, which could help determine how best to reallocate tasks under specific conditions (e.g. when an individual is biased towards hierarchical command structure, he may be less open to pursuing sensor evidence that contradicts commands from headquarters). Such inherent biases, heuristics, and tacit reasoning are a natural component of human reasoning[28] and are anticipated in our interactions with others, and it may be beneficial for AI theory of mind to include such bias compensation and expectations to optimize human+AI teaming.

2.4 Cyber Scenarios with Man-In-The-Middle Attack and Honeypot

In human decision-making, information from different domains can combine to produce unexpected effects. The psychological McGurk effect[29] is when a strong temporal synchrony between the mouth gesture "ga" and the auditory syllable "ba" combine to produce the illusory percept "da". Although multisensory integration does not appear to have been explored in C2 decision-making, the confluence across multiple domains in MDO, particularly its high volume and velocity in the Penetrate and Dis-integrate phases, may produce unexpected non-linear cross-domain interactions (this may contribute to the "fog of war"). Figure 6 illustrates an example in which a combination of actual evidence (missiles) and tank decoys (resulting from a man-in-the-middle MITM cyber attack) could synergize to compel the hostile units toward the left gap. It is a general strategy to create converging lines of evidence for cyber deception, yet specific patterns of deception may be more effective than others. For example, the brain is thought to group similar or related evidence into chunks for efficient processing (e.g. Gestalt grouping) so that it can overcome information bottlenecks (e.g. process more than seven nominal items, thereby reducing the impact of individual items). If carrying out each

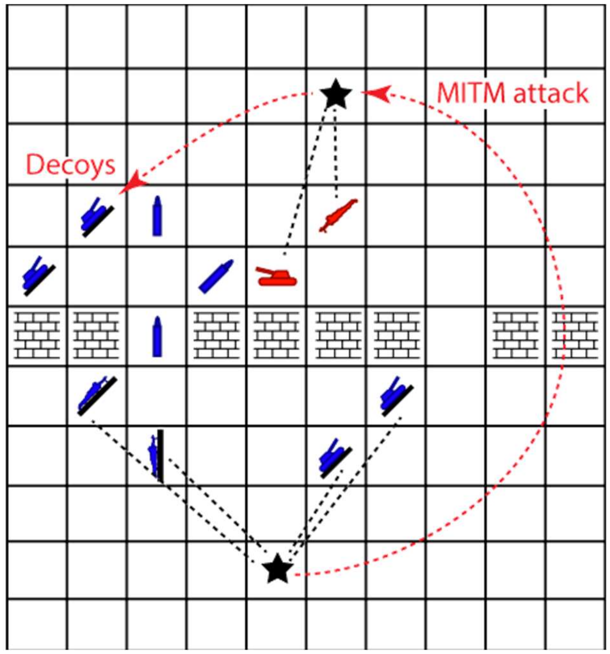

Figure 6. Cyber scenario with Man-In-The-Middle attack
To deceive the hostile team, the friendly agent applies a man-in-the-middle cyber attack to inject decoys into the communications relay used by the hostile forces. Together, the missiles and decoys act as multiple corroborating cues to further bias the hostile units toward the left gap.

instance of cyber attack incurred a certain cost or risk, it may be beneficial to understand how to distribute these costs across cue signatures to deliver the most effective impact with minimal risk (e.g. the MITM attack would probably be less effective, or even counteractive, if it produced missile decoys). It may also be informative to understand how different combinations of cues may be differentially perceived by different Soldiers. Commanders with different biases or at different roles or echelons may perceive, interpret, or act differently on the same combination of evidence (e.g. a decoy's effectiveness is likely to depend on its distance to a target commander and relevance to his decision process). More advanced strategies may include active defense, e.g. via a 'honeypot' strategy (Figure 7), to improve the effectiveness of the cyber deception. To deliver superhuman capabilities for MDO, an AI decision aid may need to aid in generating believable decoys based on the instantaneously available evidence, and to be able to rapidly adapt these presentations at the speed of cyber networks, and to maintain coherence between the virtual and real worlds, in order to maintain the effectiveness of the illusion.

3. AI TESTBED FOR CONTINUOUS TRAINING

In addition to serving as flexible platforms for developing AI reasoning, the testbeds could also be effective for parametrically training Soldiers across a wide variety of scenarios and their many combinations of decision factors. The motivation for exploring these human-AI interactions is twofold. First, current training platforms do not provide sufficient opportunities for Soldiers to practice and develop their decision-making across a wide variety of decision contexts. A typical wargaming session today consists of a group of Soldiers reasoning through one or two scenarios together, without deep dives into alternative courses of action or alternative environmental conditions or mission plans. Soldiers also do not receive interactive practice across multiple types of 'red teaming', to prepare them for a variety of hostile decision-makers. An AI testbed, running on computational resources at the DoD Supercomputing Resource Center (DSRC) via high

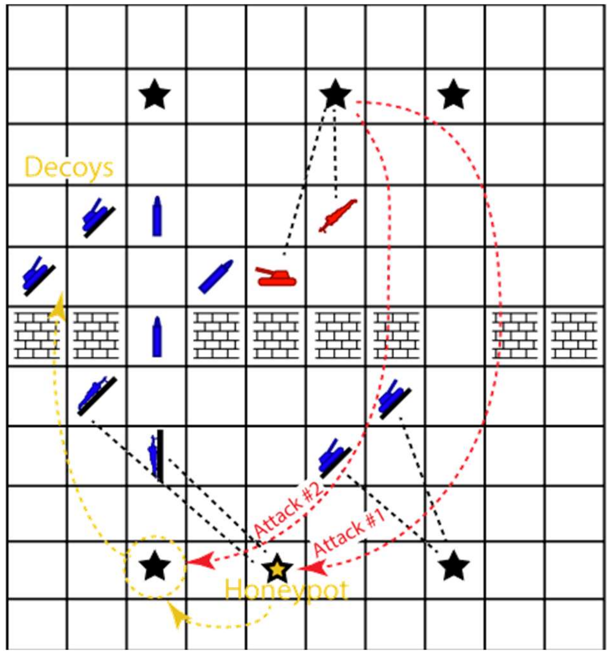

Figure 7. Cyber scenario with Honeypot

A scenario in which the friendly forces plant a secret tag or keyword ('honeypot') which is accessed during a cyberattack (attack #1). The other friendly servers monitor for hostile usage of this trigger tag or keyword (attack #2), upon which the friendly server produces false information ('decoys') for the hostile server to harvest. The use of the honeypot technique strengthens the hostile agent's belief in the decoy, improving the deception's effectiveness.

performance computing (HPC) architectures running the Persistent Services Framework (PSF) would enable anytime anywhere individualized training.

The second motivation for exploring human-AI interactions is to support human-AI teaming and co-adaptation. For the AI to be an effective decision aid, its rationale should be accessible to the Soldier for AI transparency and trust. This is especially true for complex decisions, in which human judgment is based on combinations of implicit and explicit factors. Different types of tasks emphasize different combinations of reasoning, and a flexible platform is necessary to navigate through the many permutations of the decision space and decision tree, so that the Soldier can be comfortable with how the AI arrives at decision recommendations that are highly context-dependent and lack obvious answers (e.g. Nash equilibria). Developing appropriate and effective visualizations of the AI's reasoning, and embedding game play within an advanced visualization platform such as BVI is necessary for both human-AI teaming and co-adaptation.

The converse is also true: an effective visualization platform also provides the AI access to human transparency and modeling. For complex decisions, the space of possible paths can be too broad for AI to explore in depth. Human reasoning can guide the AI to sift through these difficult choices while injecting tacit reasoning which the human may be unable to communicate easily. Over the course of many sessions and many users, the AI may be able to develop metalearning capabilities, to more effectively harness the opportunistic engagements provided by the users.

Cross-training across multiple AI testbeds may provide additional opportunities for AI generalization. For example, the ARL Simple Yeho AI testbed (Figure 8) is designed to engage platoon commanders in tactical reasoning. Some of the reasoning at this local scale (e.g. route planning and retasking) may generalize to the C2 level, and the finer granularity of the Simple Yeho environment can be an opportunity to test the complex reasoning strategies learned in ARL Battlespace, at a sparser engagement frequency and with longer time intervals between related events. Both sparseness and long term credit assignment are well known challenges to AI development, and we expect that alternating or mixing training across the two testbeds could increase resilience of the AI decision aid across scenarios and across echelons.

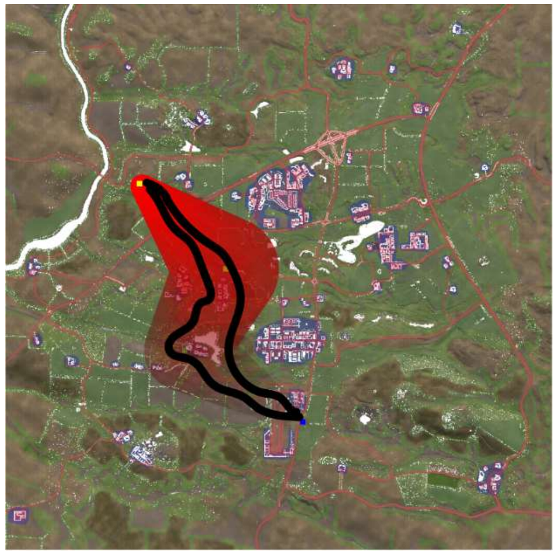

Figure 8. ARL Simple Yeho AI testbed for tactical decision aids with tacit reasoning
The ARL Simple Yeho AI testbed is based on the Yehorivka Unreal environment and consists of multiple map layers including elevation, obstacles, foliage, roads, and urban areas. Agents are provided a route (red shading) including starting point, waypoints, ending point, and left and right bounds for the mission. Result AIs are expected to be able to recommend routes and retasking opportunities to platoon commanders, consistent with Army doctrine and tacit reasoning.

4. CONCLUSIONS AND FUTURE WORK

We have discussed the motivation of developing the ARL Battlespace AI testbed and how it addresses gaps in the development of AI-enabled decision aids for MDO. We also discussed how it relates to other ongoing AI efforts, including the ARL Simple Yeho AI testbed, and how the two testbeds address different facets of Army decision-making. Whereas ARL Battlespace is more conducive for the development of complex AI reasoning, ARL Simple Yeho is better for connecting to Soldiers' tacit reasoning in a fine-grain tactical environment. We discussed many possible scenarios for developing AI and testing Soldiers across different combinations of complex reasoning. We also discussed how such exploration could benefit from cross-disciplinary investigations into biological mechanisms of decision-making, to inform the development of novel algorithms and theories. Finally, we discussed how these AI testbeds could impact training, Soldier selection, and human-AI co-adaptation. We also placed the testbed development in the context of ongoing efforts to develop HPC services and to develop AI-enabled WMIs, in support of the DoD's AI strategy.

5. ACKNOWLEDGMENTS

Research was sponsored by the Combat Capabilities Development Command Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0171. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the DEVCOM Army Research Laboratory or the U.S. Government. Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

[1] G. Allen, "Understanding AI technology," Joint Artificial Intelligence Center (JAIC) The Pentagon United States, (2020)

[2] E. S. Vorm, "Computer-Centered Humans: Why Human-AI Interaction Research Will Be Critical to Successful AI Integration in the DoD," IEEE Intelligent Systems, 35(4), 112-116 (2020).

[3] A. Feickert, "Defense Primer: Army Multi Domain Operations (MDO)," Congressional Research SVC, (2021)

[4] Army, [Military Decisionmaking Process: Lessons and Best Practices] Center for Army Lessons Learned, US Army Combined Arms Center, Fort Leavenworth, KS(2015).

[5] https://www.army.mil/article/193398/commanders_plan_autonomy_facilitates, Accessed 12-9-2021

[6] J. R. Surdu, G. D. Haines, and U. W. Pooch, "OpSim: A purpose-built distributed simulation for the mission operational environment," Simulation Series, 31, 69-74 (1999).

[7] C. T. D. West, and B. Birkmire, "Afsim," CSIAC, 7(3), 50 (2020).

[8] https://www.peostri.army.mil/OneSAF, Accessed 20-Feb-2022

[9] D. Silver, T. Hubert, J. Schrittwieser *et al.*, "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play," Science, 362(6419), 1140-1144 (2018).

[10] C. Gamage, V. Pinto, C. Xue *et al.*, "Novelty Generation Framework for AI Agents in Angry Birds Style Physics Games," 2021 IEEE Conference on Games (CoG). 1-8 (2021).

[11] V. Mnih, K. Kavukcuoglu, D. Silver *et al.*, "Playing atari with deep reinforcement learning," arXiv preprint arXiv:1312.5602, (2013).

[12] V. G. Goecks, N. Waytowich, D. E. Asher *et al.*, "On games and simulators as a platform for development of artificial intelligence for command and control," arXiv preprint arXiv:2110.11305, (2021).

[13] V. G. Goecks, N. Waytowich, D. Watkins *et al.*, "Combining Learning from Human Feedback and Knowledge Engineering to Solve Hierarchical Tasks in Minecraft," arXiv preprint arXiv:2112.03482, (2021).

[14] P. Narayanan, Z. Wu, H. Kwon *et al.*, "Overview of machine learning (ML) based perception algorithms for unstructured and degraded visual environments," SPIE Defense + Commercial Sensing: Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications. 11006, 110061T (2019).

[15] R. Köster, D. Hadfield-Menell, R. Everett *et al.*, "Spurious normativity enhances learning of compliance and enforcement behavior in artificial agents," Proceedings of the National Academy of Sciences, 119(3), (2022).

[16] J. Z. Hare, C. A. Uribe, L. M. Kaplan *et al.*, "Communication Constrained Learning with Uncertain Models," ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 8609-8613 (2020).

[17] A. Raglin, M. Dennison, S. Metu *et al.*, "Decision making with uncertainty in immersive systems," Virtual, augmented, and mixed reality (XR) technology for multi-domain operations. 11426, 114260L (2020).

[18] M. Lavine, "WHIM: function approximation where it matters," Communications in Statistics-Simulation and Computation, 1-31 (2019).

[19] J. Y. Chen, F. O. Flemisch, J. B. Lyons *et al.*, "Guest Editorial: Agent and System Transparency," IEEE Transactions on Human-Machine Systems, 3, 50, 189-193, (2020).

[20] J. Y. Chen, S. G. Lakhmani, K. Stowers *et al.*, "Situation awareness-based agent transparency and human-autonomy teaming effectiveness," Theoretical issues in ergonomics science, 19(3), 259-282 (2018).

[21] B. C. Rinderspacher, "Smooth Constrained Heuristic Optimization of a Combinatorial Chemical Space," Army Research Lab, DTIC, ARL-TR-7294 (2015)

[22] S. Su, S. Kase, C. Hung *et al.*, "Mixed Reality Visualization of Friendly vs Hostile Decision Dynamics," International Conference on Human-Computer Interaction. in Mixed Reality Technology for Immersive Analytics. Online, Springer (2021).

[23] J. Z. Hare, B. C. Rinderspacher, S. E. Kase *et al.*, "Battlespace: using AI to understand friendly vs. hostile decision dynamics in MDO," SPIE Defense + Commercial Sensing: Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III. 11746, 1174615 (2021).

Approved for public release; distribution is unlimited

[24] C. J. Garneau, M. W. Boyce, P. L. Shorter *et al.*, "The augmented reality sandtable (ARES) research strategy," US Army Research Laboratory, (2018)
[25] Army, "Charter of the Joint Requirements Oversight Council (JROC) and Implementation of the Joint Capabilities Integration and Development System (JCIDS)," Joint Staff, CJCSI 5123.01H (2018)
[26] Army, "Offense and Defense," Headquarters, Department of the Army, Army Doctrine Publication (ADP) 3-90 (2019)
[27] N. Rabinowitz, F. Perbet, F. Song *et al.*, "Machine theory of mind," International conference on machine learning. 4218-4227 (2018).
[28] J. L. Gardner, "Optimality and heuristics in perceptual neuroscience," Nature neuroscience, 22(4), 514-523 (2019).
[29] K. Tiippana, "What is the McGurk effect?," Frontiers in psychology, 5, 725 (2014).