Natural Language AI for Military Decision Support and Swarm Control for Autonomous UAS Trained in a Combat Simulation

Michael Möbius
Daniel Kallfass
AIRBUS Defence and Space GmbH
Claude-Dornier-Strasse
88090 Immenstaad
GERMANY
michael.moebius@airbus.com
daniel.kallfass@airbus.com

LTC Dr. Dietmar Kunde
German Army Headquarters
von-Hardenberg-Kaserne
Prötzeler Chaussee 25
15344 Strausberg
GERMANY
dietmarkunde@bundeswehr.org

LTC Thomas Doll
Joint Support Service Command
Abt Plg, UAbt PlgStrg/Int
Fontainengraben 150
53123 Bonn
GERMANY
thomasmanfreddoll@bundeswehr.org

ABSTRACT

The future of warfare is undergoing transformative changes through the integration of AI-assisted command systems and unmanned technologies, which will have a significant impact on combat operations and the required speed of military decision-making cycles. Future decision-support systems will assist military decision-makers in evaluating threats, developing optimal courses of action for their forces, and even executing actions through collaborative swarm behaviors of autonomous systems. To enable these systems, the combination of modeling & simulation, and advanced Deep Reinforcement Learning (RL) techniques will play a crucial role.

This paper presents the results of several studies conducted by the German Army Concepts and Capabilities Development Centre and Airbus. These studies evaluated the adaptation and utilization of simulation and AI techniques to train an AI agent capable of acting as a battalion commander in an Army combat or controlling a swarm of UAVs in an ISR mission using the RL-optimized simulation "ReLeGSim". The AI agent generates natural language commands using a language model to execute actions within ReLeGSim, enhancing communication between human advisors and AI systems while incorporating objectives and doctrines into the AI reasoning process. Through a military doctrine-aware feedback function, the AI agent assesses and improves its behavior during each training cycle.

Once trained, the AI agent can be applied to real-world scenarios, developing courses of action alternatives to a battalion commander derived from the learned AI agent policy, or directly executing them in autonomous systems to control a swarm of UAVs. This research serves as a foundation for equipping AI agents with the ability to uphold military doctrines and rules in future operations.

1.0 INTRODUCTION

In recent years, Artificial Intelligence (AI) has undergone significant advancements, with Reinforcement Learning (RL) emerging as a prominent paradigm. RL has gathered substantial attention for its capacity to achieve remarkable performance levels, even surpassing human capabilities in complex gaming scenarios such as Dota2 and StarCraft. It has become the state-of-the-art AI technique in the field of machine learning for solving complex tasks.

The primary objective of current military studies is to transpose RL techniques, originally devised for gaming applications, into the realm of military operations. The overarching ambition is to develop AI-based systems for military endeavors that can exhibit superhuman-level performance in many use cases such as [16]:

- Battlefield decision-making: RL can be used to train agents to make decisions in complex military scenarios [1] by using a simulation environment. The decisions taken by the AI can be used as recommendations to a commander, e.g., for an effective course of action.
- Autonomous systems: RL can be used to train agents to control military vehicles (e.g., drones, tanks) in a simulation [2]. The agents can learn to navigate vehicles in the environment and perform various tasks (e.g., reconnaissance, target acquisition). The trained agents can be transferred to real vehicles without the need of retraining the AI.
- Planning & Optimization: RL can for example, be used to optimize logistics planning in military simulations [3]. The agents can learn to allocate resources (e. g., troops, supplies) to different areas of the battlefield to achieve mission objectives while minimizing losses.
- Cybersecurity: RL can be used to train agents to detect and respond to cyber-attacks in military simulations [4]. The agents can learn to identify and mitigate threats to military networks and systems.
- Training and evaluation: RL can be used to train and evaluate military personnel in simulations [5]. The agents can simulate different scenarios and provide feedback on the actions taken by the trainees.

The technologies applied in RL are continuously changing and improving. New architectures like transformer models [6] and new activation functions like SiLU [7] are further improving the architecture and overall performance of AI agents trained with RL. Transformers models allow new architectures like Vision-Transformers [8] and is the foundation for all recent Large Language Models such as GPT (Generative Pre-trained Transformer) from OpenAI [9].

Motivated by these developments, this paper examines the usage of new language model architectures to tackle the problem of huge action spaces needed for military operations and to improve the AI agent's overall performance.

2.0 RELATED WORK

Complex decision-making capabilities often come with huge action spaces in RL and mitigating exploding action spaces is an active area of research. The paper "Growing Action Spaces" [10] emphasizes that random exploration is not good enough for large spaces and curriculum learning can be crucial to learn these action spaces. Recent developments use action spaces characterized by natural language and successfully exploit their flexibility of complex action generation [11].

Recent advances in natural language processing inspired developers to expand the possibilities regarding the use of natural language. Language models are typically used for Question Answering and Conversations. However, these models can also be trained to interact with an environment via RL. In their paper "Learning to Model the World with Language" [12] introduces the concept of building agents that can understand and use diverse language in numerous ways, including conveying general knowledge, describing the world's state, and providing feedback. The central idea is that language helps agents predict the future, including what will be observed, how the world will behave, and which actions will be rewarded. The authors present "Dynalang", an agent that learns a multimodal world model to predict future text and image representations and makes decisions based on simulated model rollouts. Unlike traditional agents, Dynalang uses language not only for action prediction but also for predicting future language, video, and rewards, resulting in rich language understanding. Additionally, Dynalang can be pre-trained on language and video datasets without actions or rewards, and it effectively uses language to improve task performance in various environments, from grid worlds to photorealistic home scans.

Another crucial side of RL lies in the adaptability of reward systems - the concept of providing AI agents with incentives to encourage desired behaviors. Reward shaping constitutes a technique employed to modify these reward structures systematically. In practical terms, this involves the fine-tuning of rewards to guide AI agents towards specific objectives. As an illustration, in the context of maze navigation, AI agents can be endowed with incremental rewards for exploring previously uncharted regions, thereby stimulating comprehensive exploration. An alternative strategy entails meta-learning or multi-task learning, which empowers AI systems to concurrently oversee multiple, potentially distinct, objectives. This approach resembles the simultaneous mastery of several tasks, facilitated through the sharing of acquired knowledge and skills among them. Nonetheless, the process of dynamically altering reward functions in AI is accompanied by intrinsic challenges.

If the transformation of objectives is excessively abrupt, AI systems may struggle with adaptation, requiring resource-intensive retraining efforts. Frequent alterations to objectives can potentially cause perplexity within the AI. In summary, the practice of dynamically modulating reward mechanisms within AI embodies a potent tool, although one requiring careful management. The overarching goal is to strike an equilibrium between adaptability and stability in the AI's learning process, ensuring a harmonious balance between the accommodation of evolving objectives and the preservation of effective learning dynamics.

The recent paper “Designing Rewards for Fast Learning” [13] explores the impact of reward function design on the learning speed of RL agents. It emphasizes the importance of choosing state-based rewards that maximize the action gap, making it easier for agents to distinguish optimal actions from suboptimal ones. The paper also introduces the concept of minimizing a measure called the "subjective discount" to encourage agents to make optimal decisions with less look-ahead. To address this reward design problem, the paper proposes a linear programming algorithm. Experimental results in tabular environments with Q-Learning demonstrate that the generated rewards lead to faster learning. The study identifies three key principles of reward design: 1) Penalizing each step taken facilitates faster learning compared to rewarding the goal. 2) When rewarding subgoals along the target trajectory, rewards should gradually increase as the goal approaches. 3) Dense rewards that are nonzero in every state are beneficial only when carefully designed.

3.0 RELEGS – REINFORCEMENT LEARNING FOR COMPLEX COMBAT SITUATIONS

3.1 The Simulation Environment "ReLeGSim"

ReLeGSim (*Re*inforcement *Le*arning focused *G*eneric AI Training *Sim*ulation, shown in Figure 1) is a chessboard-like simulation environment for reinforcement learning to develop self-optimizing strategies of the players in the game. Arbitrary players are supposed to reach a goal in a sequence of moves and can interact with each other. ReLeGSim can be used to model a broad-range of civil and military scenarios such as an ISR mission or a large battalion ground combat scenario. ReLeGSim allows to define actors for the chess game-like environment, to give them corresponding properties and possible actions. For this purpose, the simulation can be extended by appropriate application-specific simulation models (such as a sensor) using the Python programming language.

In the scope of the ReLeGs[1] study, ReLeGSim was configured to model a confrontation of 2 battalions, where each battalion commander has to command its assigned companies and support units. It allows players, whether human or AI agents, to generate the commands of the battalion in either an attack or defend tactical scenario. The simulation was introduced in a paper by [1] and uses the "Gymnasium" API for reinforcement learning [14].

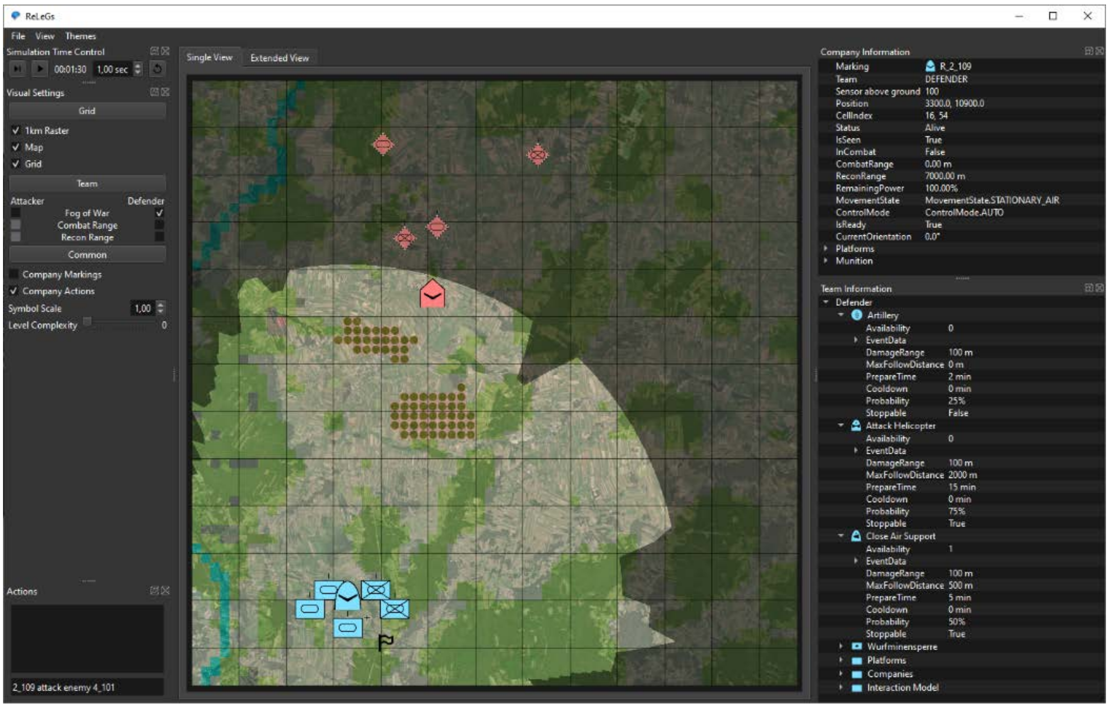

Figure 1 ReLeGSim UI

In the simulation, one player takes on the role of the attacker, aiming to capture a specific target area from the defender, who must hold it throughout the episode. Both players have access to various companies with unique capabilities, consisting of platoons and individual units. To succeed, players must understand their opponent's perspective, know their company's abilities, and effectively navigate the terrain.

[1] ReLeGs is a German abbreviation for "Reinforcement Learning for Complex Combat Situations" (German: "Reinforcement Learning für komplexe Gefechtssituationen")

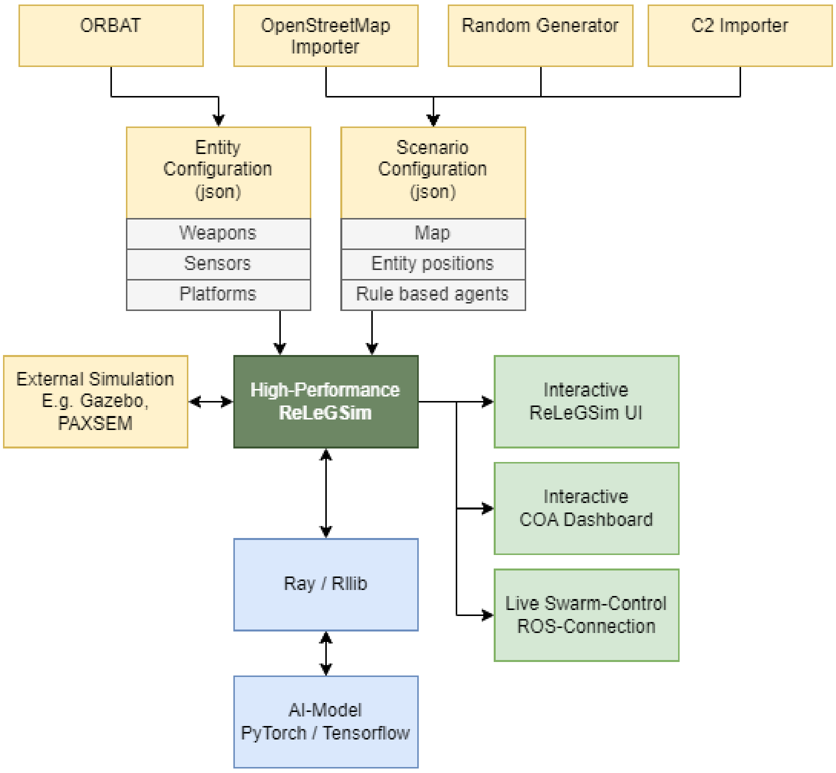

Figure 2 AI-Toolchain - ReLeGSim

The AI toolchain (Figure 2) allows the automatic creation of 3D terrain, based on real-world data, such as vector, elevation, and satellite information. The rasterized map is then used for AI training within ReLeGSim, with specific field types (like forests or roads) assigned to different areas. The simulation with the additional tools aims to provide a platform for training different AI models through reinforcement learning, but it also supports human vs. AI gameplay. Therefore, it is possible to benchmark, test, evaluate, and analyze the capabilities of the trained agents. The toolchain also includes automated testing for trained AI agents with various metrics and complex analysis based on customer needs.

3.2 ReLeGSim AI Architecture

The authors of ReLeGSim drew inspiration from DeepMind's AlphaStar [15], a leading model in complex RL problems, to develop an innovative architecture (Figure 3). Influenced by military tactics, the design utilizes scalar data and visual maps for scenario observation. Scalar data, including troop numbers and ammunition to extend the AI's view. All the input parameters are normalized for better training performance. A multi-head attention network, instead of a fully connected layer, for the scalar values improves agent quality. To understand the terrain, the AI receives a visual map with a lot of terrain information and entity encodings. A spatial encoder with convolutional layers was developed to incorporate this rich data into the AI.

The architecture is assessed and reduced to the minimum via an autoencoder setup, reducing parameters from 2 million to 47,000 and producing a pre-trained model. An optional language input can take an objective or task into consideration. In a hierarchical setup, the given task can be defined by a superior agent. Encoded values from visual, task, and scalar data are input to a core network, an LSTM component, to handle long-term planning.

The action head was initially implemented as a multi-discrete action space based on the AlphaStar implementation. Due to the exploding action space, the action head was replaced by a language model based on the latest research to predict action commands in natural language, as it will be described in the next section.

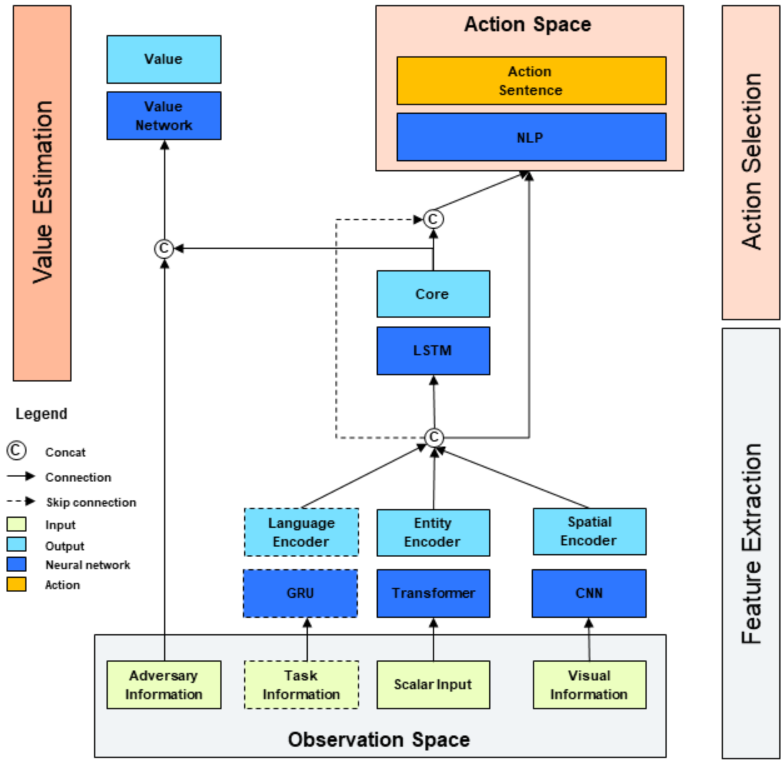

Figure 3 AI Architecture used in ReLeGSim

3.3 AI Decision Space

The issue of complex decision-making capabilities of AI comes with huge action spaces in RL, which presents a significant challenge as RL applications become more complex and realistic. Small, fixed-action spaces have limitations in terms of expressiveness, exploration, and efficiency. Researchers are continually developing new techniques and algorithms to mitigate the impact of exploding action spaces, such as function approximation, discretization, and hierarchical RL. These approaches enable RL agents to tackle increasingly complex tasks and navigate the challenges of large action spaces more effectively. As RL continues to advance, addressing the issue of exploding action spaces will remain a critical research area to enable the successful application of RL in real-world scenarios.

The approach of employing natural language to establish communication with AI, as exemplified in [2], coupled with the developments in the formulation of doctrines using natural language, highlighted in [16], set a precedent for the realization of a versatile AI capability within a multifaceted operational environment. ReLeGSim incorporates a natural language interface between the AI and the agents in the simulation with a complex parsing and execution of given commands. These commands can be on different hierarchical levels and control various agents.

Initial trials showed that a large space of unused vocabulary is disadvantageous and results in slow training. Therefore, we use a small, but effective vocabulary. The vocabulary contains only the following tokens:

(0, 1, 2, 3, 4, 5, 6, 7, 8, 9,
attack, move, observe, x, y, own_entity, enemy_entity, artillery, attack_helicopter, close_air_support, <colon>, <comma>)

The token *<colon>* splits the resulting output text sequence into multiple actions, while the <comma> token ends or pads the results. The reduction of tokens and the optimization were done manually and corresponded directly to the execution of the resulting behavior in the simulation. To tokenize the actions, we use one-hot encoding, as this allows us to use stochastic sampling over the given actions and can be easily integrated into any given RL framework through a multi-discrete representation.

4.0 EXPERIMENTS & RESULTS

4.1 Course of Action (COA) decision support dashboard in "ReLeGs"

Our AI model for ReLeGSim deeply integrates human intervention for task prioritization, enabling real-time objective changes by incorporating task information into the observation space (Figure 3). To train this behavior, we employ a curriculum learning strategy, introducing various priorities expressed in natural language, each associated with a reward guiding rule adherence. This approach encourages the AI agent to develop a broad skill set, excel in diverse scenarios, and efficiently achieve objectives.

In order to use the trained AI agent for Course of Action (COA) decision support, a COA decision support web application was developed. Based on a given combat situation and ORBAT for blue and red forces, the decision support web application generates a large number of ReLeGSim simulation runs to gain statistics how the AI agent has acted in the given situations. In addition, all possible decision factors such as number of available Joint Fire strikes is varied in order to let the user set specific filter settings, to analyze the resulting data. The resulting AI commanded simulation runs are then statistically analyzed and visualized e.g., through heatmaps in the web-based dashboard.

Figure 4 shows an example for a heatmap visualization. It shows the movement of all blue and red units for multiple simulation runs with identical start conditions. In addition, in this example 2 different attack priorities have been given to the red battalion commander: a) attack as quickly as possible vs. b) attack with minimal own losses. Figure 4 illustrates how the AI adapts its behavior based on different objectives, such as prioritizing quickness or minimizing losses.

(a) Attack as quickly as possible. (b) As few own losses as possible.

Figure 4 Comparing company movements based on the given priority (a) and (b)

This is just one way to explore different possibilities within a given scenario to help an operator generate and validate its Courses of Action. Different options from war gaming, and statistical analyses up to red forces COA predictions are available in the toolbox of ReLeGSim. This decision support tool automates scenario testing, tactic optimization, and AI model evaluation, promoting diverse exploration and adaptable decision-making.

4.2 Autonomous control of a heterogenous UAS swarms in "KITU"

In the "AI for tactical UAS (KITU)" research study of the German Army Headquarters, Airbus Defense and Space along with two German startup companies Quantum-Systems and Spleenlab is exploring the use of artificial intelligence (AI) to control tactical Unmanned Aircraft Systems (UAS) in military scenarios. This research focuses on demonstrating and analyzing AI components for autonomous UAS swarms, with an emphasis to align with the Main Ground Combat System (MGCS) and NATO eastern flank surveillance scenarios. The focus areas are coordination of heterogenous UAS swarms, target detection, and dynamic mission execution. Various automation tasks are trained using AI to understand the sensor-to-shooter chain's effort, effectiveness, and efficiency. The study also investigates data processing locations, resilience, and robustness of swarm control in case of disruptions.

Deep reinforcement learning methods are employed to develop AI capable of controlling UAS swarms under human supervision. Figure 5 shows the process from RL Training, verification to real flight tests. In order to train the controlling of the UAS swarm, the ReLeGSim simulation was adapted to be able to provide a simplified model of fixed-wing and multi rotor UAS with different flight characteristics, battery power supply and consumptions as well as the payload such as optical sensors. The action space of the so called UAS swarm controller was adapted to give the UAS search and track tasks as well as the possibility to land on ground to observe targets from ground with low battery consumption. Once successfully trained, the behavior is transferred to real UAS for testing in flight at the Airbus Drone Center. First flight experiments showed that the RL agent trained in the ReLeGSim simulation environment performs pretty well in real situations, after some Sim-to-Reality gaps have been reduced by increasing the model fidelity and calibration of the model parameters with the reality.

The insights gained aim to inform the integration of AI-learned behaviors into real UAS systems and evaluate their similarity to manual control. In summary, projects like KITU are essential for European defense programs, including drone swarms, AI, and cloud computing, with potential benefits for MGCS and Future Combat Air System (FCAS) development programs. Drone swarms offer force multiplication and increased reconnaissance capabilities, making them valuable in tactical scenarios.

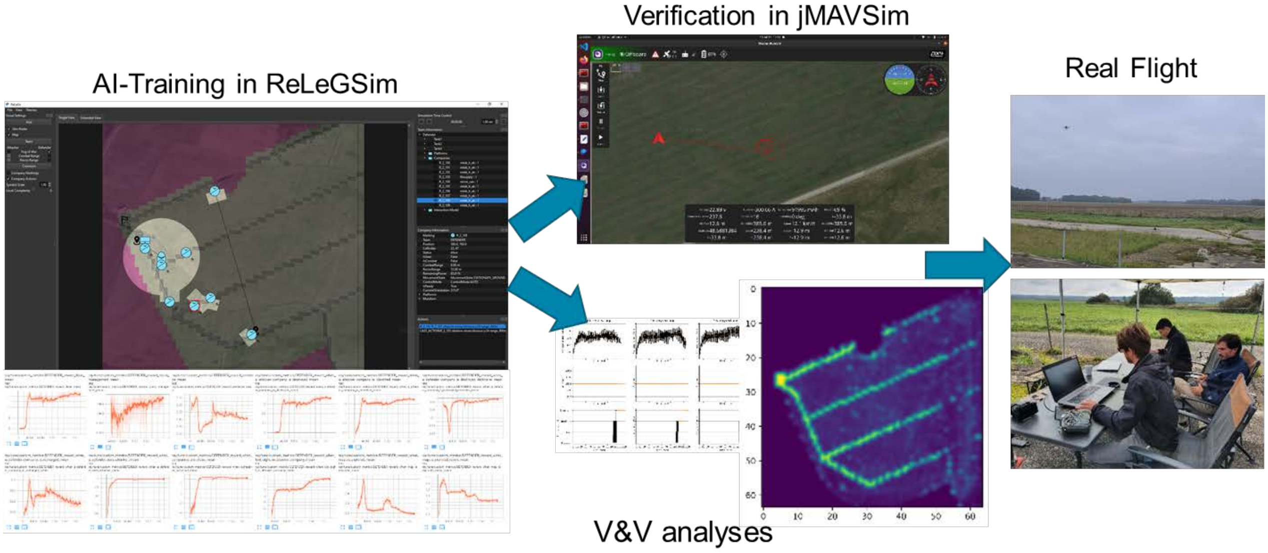

Figure 5 - Transfer of ReLeGSim AI-agent to control a real UAS swarm

5.0 CONCLUSION

In conclusion, this paper addresses the transformative changes occurring in the future of warfare, driven by the integration of AI-assisted command systems and unmanned technologies. These changes will significantly impact combat operations and necessitate faster military decision-making cycles. To enable these future decision-support systems, the integration of modeling, simulation, and advanced deep reinforcement learning techniques is crucial. These systems will assist military decision-makers in evaluating threats, developing optimal courses of action, and even executing actions through collaborative swarm behaviors of autonomous systems.

The research presented here showcases the adaptation and utilization of simulation and AI techniques to train AI agents capable of acting as battalion commanders or controlling UAV swarms using the RL-optimized simulation "ReLeGSim". These AI agents communicate through natural language commands, enhancing human-AI interaction while incorporating objectives and doctrines into the AI reasoning process. The integration of a military doctrine-aware feedback function enables AI agent self-improvement during training cycles.

While both studies “ReLeGs” and “KITU” do not aim to replace human decision-makers entirely, they provide valuable insights into the potential of AI in military operations. The development of RL agents, although challenging, has demonstrated promising behavioral patterns, including intelligent terrain utilization and strategic decision-making. As the study progresses, further insights and behavioral patterns are expected to emerge. This research lays the foundation for equipping AI agents with the ability to uphold military doctrines and rules, offering enhanced support to human decision-makers and opening avenues for AI applications in various military scenarios, training, and decision-support systems. The future of AI in warfare is marked by collaboration and augmentation, where AI serves as a valuable tool alongside human expertise, ensuring that "Man makes decisions while the machine provides support".

6.0 WAY AHEAD

Numerous national and international research efforts highlight the imperative for future operations, at every echelon, to be executed with significantly enhanced speed. Unlike in the past, where a brigade commander had the luxury of several hours for decision-making, the time available for such deliberation will need to be notably and progressively shortened to attain a position of dominance. Several factors contribute to this evolving landscape. The way ahead involves several key directions to further advance the research and practical applications:

1. Continued Training and Evaluation: The ongoing training of RL agents should be completed to further refine their behavioral patterns. This includes the development of more complex tactical behaviors, such as target prioritization, formation of reserves, and counterattack strategies. Additionally, the explanation of RL agents' behavior through Explainable AI (XAI) should be explored in greater detail to enhance human understanding.

2. Scalability and Real-World Testing: While simulation environments like ReLeGSim provide valuable training grounds, efforts should be made to scale up these AI agents for real-world testing and deployment. This includes addressing hardware and computational requirements to ensure practical applicability.

3. Human-in-the-Loop Integration: The integration of AI as decision support must continue to emphasize human control and intervention. Developing interfaces and protocols for seamless collaboration between human commanders and AI agents is essential.

4. Versatility in AI Applications: The research should expand its focus beyond decision support to explore a wide range of AI applications in military contexts. This includes the training of autonomous unmanned systems, conducting simulations for training exercises, and evaluating AI models' performance and tactics. In addition, other simulation models shall be applied with the ReLeGSim RL architecture, to train RL agents in highly detailed combat models such as PAXSEM which would be necessary to model e.g., intensive air combat or air defense scenarios.

5. Ethical and Legal Considerations: As AI's role in military operations grows, it is imperative to address ethical and legal aspects. Research should encompass discussions and solutions regarding responsible AI use, accountability, and adherence to international laws and conventions.

6. Testing and Validation: Rigorous testing and validation of AI models, particularly in complex and dynamic combat scenarios, should remain a priority. This includes evaluating the AI's performance in a variety of contexts, such as urban warfare, irregular warfare, and peacekeeping operations.

7. Adaptation to Evolving Technologies: Given the fast-paced nature of AI advancements, the research should remain adaptable and open to incorporating emerging technologies, architectures, and best practices to stay at the forefront of AI-assisted military decision support. Large Language Models (LLM) especially Multi Modal LLMs have the potential to revolutionize the understanding of the situational awareness, reasoning and to derive action plans. This technology has the great potential to significantly improve the RL agent.

In summary, the path forward involves adopting a comprehensive strategy to advance and seamlessly integrate AI into military operations while consistently adhering to ethical and legal standards. By addressing these critical dimensions, this research can contribute to the evolution of AI-powered decision support systems and their careful application within complex military contexts.

7.0 ACKNOWLEDGEMENT

We would like to thank the entire “KITU” and “ReLeGs” study groups, especially MAJ Stefan Göricke from the Army Concepts and Capabilities Development Center and MAJ Matthias Kuc from the German Army Headquarters, for their invaluable contributions to our research endeavors.

8.0 REFERENCES

[1] Doll, L. T., Behm, M., Brendecke, J. W., & Kallfass, D. (2021). From the Game Map to the Battlefield– Using DeepMind’s Advanced Alphastar Techniques to Support Military Decision-Makers. *Towards Training and Decision Support for Complex Multi-Domain Operations, NATO Modelling and Simulation Group (NMSG) Syposium held on 21–22 Ocotober 2021, in Amsterdam, the Netherlands*, 14-1.

[2] Möbius, M., Kallfass, D., Doll, T., & Kunde, D. (2022, December). AI-Based Military Decision Support Using Natural Language. In *2022 Winter Simulation Conference (WSC)* (pp. 1-12). IEEE.

[3] Yan, Y., Chow, A. H., Ho, C. P., Kuo, Y. H., Wu, Q., & Ying, C. (2022). Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities. *Transportation Research Part E: Logistics and Transportation Review*, *162*, 102712.

[4] Vyas, S., Hannay, J., Bolton, A., & Burnap, P. P. (2023). Automated Cyber Defence: A Review. *arXiv preprint arXiv:2303.04926*.

[5] Salas, E., Milham, L. M., & Bowers, C. A. (2003). Training evaluation in the military: Misconceptions, opportunities, and challenges. *Military Psychology*, *15*(1), 3-16.

[6] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. arXiv:1706.03762 [cs]. http://arxiv.org/abs/1706.03762

[7] Elfwing, S., Uchibe, E., & Doya, K. (2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural networks*, *107*, 3-11.

[8] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2010). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv 2020. *arXiv preprint arXiv:2010.11929*.

[9] Yenduri, G., Srivastava, G., Maddikunta, P. K. R., Jhaveri, R. H., Wang, W., Vasilakos, A. V., & Gadekallu, T. R. (2023). Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. *arXiv preprint arXiv:2305.10435*.

[10] Farquhar, G., Gustafson, L., Lin, Z., Whiteson, S., Usunier, N., & Synnaeve, G. (2020, November). Growing action spaces. In *International Conference on Machine Learning* (pp. 3040-3051). PMLR.

[11] Ryu, D. K., Fang, M., Pan, S., Haffari, G., & Shareghi, E. (2023). A Minimal Approach for Natural Language Action Space in Text-based Games. *arXiv preprint arXiv:2305.04082*.

[12] Lin, J., Du, Y., Watkins, O., Hafner, D., Abbeel, P., Klein, D., & Dragan, A. (2023). Learning to model the world with language. *arXiv preprint arXiv:2308.01399*.

[13] Sowerby, H., Zhou, Z., & Littman, M. L. (2022). Designing rewards for fast learning. *arXiv preprint arXiv:2205.15400*.

[14] Towers, M., Terry, J. K., Kwiatkowski, A., Balis, J. U., Cola, G. de, Deleu, T., … Younis, O. G. (2023, March). *Gymnasium*. doi:10.5281/zenodo.8127026

[15] Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., ... & Silver, D. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, *575*(7782), 350-354.

[16] Möbius, M., Kallfass, D., Doll, T., & Kunde, D. (2023, December). Incorporation of military doctrines and objectives into an AI agent via natural language and reward in reinforcement learning. In *2023 Winter Simulation Conference (WSC)* (pp. 1-11). IEEE.