23rd International Command and Control Research & Technology Symposium
“Multi-Domain C2”

Title

Artificial Intelligence for Decision Support in Command and Control Systems

Topic

Topic 9: Knowledge Systems

Authors

Johan Schubert[1†], Joel Brynielsson[1], Mattias Nilsson[1], Peter Svenmarck[2]

[1]Department of Decision Support Systems
Division of Defence and Security, Systems and Technology
Swedish Defence Research Agency
SE-164 90 Stockholm, Sweden

[2]Department of Man - Technology - Organisation
Division of C4ISR
Swedish Defence Research Agency
SE-164 90 Stockholm, Sweden

[†]Corresponding author

E-mail: johan.schubert@foi.se

Artificial Intelligence for Decision Support in Command and Control Systems

Johan Schubert[1†], Joel Brynielsson[1], Mattias Nilsson[1], Peter Svenmarck[2]

[1]Department of Decision Support Systems
Division of Defence and Security, Systems and Technology
Swedish Defence Research Agency
SE-164 90 Stockholm, Sweden

[2]Department of Man - Technology - Organisation
Division of C4ISR
Swedish Defence Research Agency
SE-164 90 Stockholm, Sweden

†Corresponding author
E-mail: johan.schubert@foi.se

Abstract

Artificial intelligence (AI) is the intelligent behavior displayed by machines. In everyday terms, the term AI is used when machines mimic the cognitive functions that people associate with learning and problem solving. The key issues within AI include reasoning, planning, and learning. In military applications, AI becomes increasingly important in systems used at different military levels, from the combat level to tactical and operational levels. This development has led to decision support systems being used at the battalion and brigade levels. Based on empirical data gathered through structured user-centered activities involving military personnel, this study investigates how AI may be used in command and control systems. We study its use in the intelligence and operations processes. We discuss how AI methods can be used for decision support for processes that provide a common operational picture, use threat analysis to predict enemy actions, and analyze own forces' alternative actions before execution. We conclude that the benefit of AI for the armed forces is that it can deliver critical system support when time is limited or when the number of choices is too large for people to be able to analyze all alternatives. We believe that the side that successfully implements AI in its command and control system can become the best and fastest at analyzing information and as a result can make quicker decisions and gain an operational advantage over its opponent.

Keywords

Artificial intelligence, command and control, OODA-loop, analysis, planning, execution.

1 Introduction

The Oxford dictionary defines artificial intelligence (AI) as follows:

"The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."

At present, it is hard to think of a more prominent buzzword than AI. Of course, with recent advances in performance, AI even surpasses humans at some tasks such as playing the game

Go [1], skin cancer detection [2], and speech recognition [3], and there are some good justifications for its use.

The common denominator of these advances is the subfield of deep learning (DL). Deep learning refers to machine learning models consisting of multiple layers of nonlinear processing units. Typically, artificial neural networks represent these models, where in this context a neuron refers to a single computation unit where the output is a weighted sum of inputs that passed a (nonlinear) activation function (e.g., a function that passes the signal only if it is positive).

Deep learning systems based on artificial neural networks are called deep neural networks (DNNs) and consist of a large number of serially connected layers of parallel-connected neurons. The combination of access to large amounts of data and powerful computers together with a series of innovations (e.g., initialization strategies and data normalizations) has led to the successful training of these large capacity networks. Representation learning is one of the main reasons for the high performance of DNNs. Using DL and DNNs, it is no longer necessary to manually craft the features required to learn a specific task. Instead, discriminating features are automatically learned during the training of a DNN.

It should be stressed that DNNs are not the silver bullet to all AI problems and that other AI concepts and machine learning models are needed depending on the specific scenario and task.

According to McCann and Pigeau [4], command and control (C2) is defined as "the establishment of common intent to achieve coordinated action". In a military context, the central problems of C2 are as follows [5]:

- How can one obtain a collective effect from a large set of resources?
- How can one handle inherent uncertainties?
- How can one produce an impact at a faster pace than the enemy?

Producing an impact at a faster pace than the enemy forces the enemy to react rather than act. A prerequisite to achieve this is to be able to process large quantities of information and to model uncertainties efficiently.

To address these problems in a structured way, C2 is always accompanied by a C2 system [5]. The C2 system consists of people, organizations, processes, methods and equipment. As mentioned by Brehmer [5], the products of a C2 system are orders, and in order to generate orders, the system needs to facilitate (*i*) data collection, (*ii*) reasoning/sensemaking (i.e., analyze information and identify what needs to be done), and (*iii*) planning (i.e., turn what needs to be done into how it could be done).

For the military sector, the benefit of incorporating AI into C2 systems is that it can potentially deliver critical system support when the time is limited or when the number of options is too large for people to be able to analyze alternative courses of action. Thus, the strategic importance of using AI at the tactical and operational levels can hardly be exaggerated. Ayoub and Payne [6] write that *"a domain specific AI could radically shift military power towards the side that develops it to maturity. Domain-specific AI will be transformative of conflict, and like previous transformations in military capability it has the potential to be profoundly disruptive of the strategic balance. (...) [T]actical and operational systems hold most promise, and that these will have a strategic impact."*

In this concept paper, we discuss the use of AI methods in *decision support systems* (DSS). Based on empirical data from a workshop held at the Swedish Armed Forces Command and

Control School, we identify areas and tasks where AI would potentially have the largest impacts within the existing C2 systems with respect to the three central problems in C2 that were listed previously. Moreover, we discuss different aspects of AI methods and their corresponding suitability for specific tasks. Specifically, being able to explain certain suggestions produced by an AI is likely to be central for an AI-based DSS.

The remainder of the paper is organized as follows. In Section 2, we introduce the dynamic observe, orient, decide, and act-loop modeling of the C2 system, and describe the user-centered methodology that is used to identify the C2 system's challenges where AI can potentially be utilized to make a difference. The findings from the user-centered activities are then summarized in Section 3. The opportunities and challenges of AI methods for some of the highlighted tasks that are identified in the workshop are presented in Sections 4 and 5, respectively. Finally, Section 6 is devoted to the conclusions.

2 Methodology

For a given mission, the C2 process is a highly dynamic process that—as inspired by Boyd [7]—can be modeled, at some level of abstraction, as a dynamic observe, orient, decide, and act-loop (a so-called DOODA-loop), as shown in Figure 1 [5]. The orders translate to military activity, which in turn causes some effects to be filtered by frictions (unknowns impairing the effect). The effects are observed by sensors (in the widest possible sense, from electronic sensors to human observations), and sensor data are collected together with data from the system's internal state, such as the mission's progress. In the reasoning/sensemaking [8] process, the events crucial to the mission are determined together with identification of tasks, resources, and constraints. Finally, the planning process from which the orders are based determines the plans, allocates assets, assesses risks, evaluates, selects, and rehearses plans, etc. The loop continues until the mission is accomplished, lost, or withdrawn. All stages in the DOODA-loop are associated with some uncertainty that has to be accounted for by the C2 system.

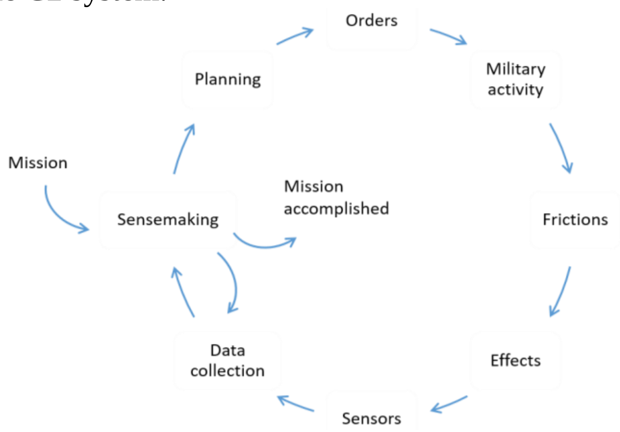

Figure 1: Illustration of the DOODA-loop [5].

The complexity and diversity of tasks to be handled in a DSS within a C2 system are large. Thus, it is unlikely that in the very near future we will have an AI that simply outputs a set of suggested orders given all available sensor data and internal system states. Instead, a gradual introduction of AI in a DSS seems more plausible. To identify where bottlenecks causing reduced pace of the DOODA-loop are in the current armed forces C2 systems and to identify where improvements can be made, a user-centered design (UCD) approach was adopted [9]. Such an approach makes it possible to give a voice to end-users and laypersons that otherwise have little opportunity to affect the future development of technology [10]. The UCD is characterized by an elaborate process making use of appropriate design methods and design activities. These methods and activities depend on the issue at hand, what expert knowledge is needed, and the accessibility of end-users, designer engineers, and policy-makers. The idea is that designers and design engineers are enablers for facilitating the design activities and being design experts, while the end-users along with other stakeholders are considered to be experts within their respective domains. This clarifies the roles and competencies so that decisions concerning future design choices and procurement can be based on relevant and accurate information.

Inspired by the UCD philosophy, a structured brainstorm was conducted with the aim of obtaining deeper knowledge concerning the users' needs and to be able to envision the development of future C2 systems. The participants consisted of officers from the Swedish Armed Forces Command and Control School, design engineers, and researchers. The structured brainstorm included individual and joint brainstorming activities, as well as the prioritization of needs.

The question to be answered during the brainstorm concerned the *purposes*, by *whom*, and *where* an AI system would be beneficial in a C2 system context. To ensure that the full possibilities of tomorrow's technology were accounted for, the participants were explicitly told to disregard any concerns regarding financial, legal, and technical challenges. The participants were asked to first write down their ideas individually. The ideas were then grouped into a number of clusters by all participants, where each cluster was assigned a representative tag. Finally, to assess the priority of the generated ideas, each participant prioritized three different ideas on a scale ranging from one to three.

3 Findings

Three different clusters were identified: analysis (and monitoring), planning, and execution. These can be tied to the sensemaking, planning, and activity processes in the DOODA-loop described earlier. Given the participants' background and experience, this is not surprising. In terms of importance, the analysis cluster was deemed to have the highest priority, followed by planning and then execution.

Within the analysis cluster, three different subclusters could be identified: finding information, compiling information, and the detection of anomalies in the information. Examples of ideas related to finding information were tailor-made searches (for individuals or for roles) and the automatic meta-tagging of information (e.g., topic or security level). An idea listed in the information compiling subcluster was automatic common operational picture updates. Lastly, identifying inconsistencies in and between reports and orders was listed as an example tied to detecting anomalies.

The ideas within the planning cluster could be split into two subclusters: planning support and tactics development. In the planning support case, the ideas that were suggested were terrain analysis (e.g., to show routes with a minimum detectability), logistics planning (e.g., to move from *A* to *B* requires *x*, *y*, and *z*), the prediction of the enemy's awareness of the situation, the prediction of the enemy's behavior (from doctrine to actual data), and the automatic proposal of action plans. For tactic development, the use of reinforcement learning[1] for air and naval combat was proposed, which would potentially lead to new military doctrines.

Ideas listed in the execution cluster focused on evaluation of action alternatives for the commander, and streamlining staff work in the military headquarters during execution of

[1] Reinforcement learning is a machine learning technique where action-state pairs are learnt with the objective to maximizing the expectation of discounted future rewards. This technique in combination with deep neural networks is the foundation of alpha-Go from Google DeepMind [1].

operations. AI for automatically generating (customized) report summaries and automatically transcribing speech to text were two concrete ideas proposed.

4 Opportunities

Today, the nonmilitary sector drives the innovations in AI. However, the developed techniques and concepts are quite generic and can be used in military systems. In this section, we discuss the opportunities for using AI in the context of some of the problems distilled during the workshop conducted at the Swedish Armed Forces Command and Control School.

The focus is on decision support for analyzing the situation and proposing and evaluating actions for our own forces. We identify several subtasks: to analyze the current situation, to provide decision support for making plans, to evaluate plans already made and during their execution, providing decision support for dynamic replanning, and to extend and refine plans as the evolution of events progresses.

The most important contribution for obtaining good decision support is to construct an adequate knowledge representation for the current issue. Knowledge representation provides the framework within which the AI methods will work. To construct the knowledge representation is an intellectual problem. Given that this has been done in a good way, the rest of decision support is a matter of mathematics within the framework of the representation and to provide a good presentation for the decision-maker. If we assume that the operation to be planned, evaluated, implemented and dynamically replanned can be fully described in the form of parameters with multiple possible values and that a valuation of such plans is done with several measures of effectiveness (MOE), then the problem is to find a plan that provides a good MOE.

4.1 AI for Analysis

In the analysis phase, one processes and combines information to build a common operational picture (COP). This includes classifying incoming information, identifying the current situation, constructing a dynamically updated COP, and checking if one's own system is being deceived. Using information fusion techniques, a tactical COP can be automatically generated based on an incoming sequence of intelligence reports [11]. The analysis phase is thus important in itself, but it also has a further purpose in that its results constitute the available inputs to decision support in the following phases of planning and execution.

The methods to understand what an identified situation means are important; *"[t]he current emphasis on understanding has resulted from UK military commanders' recognition of a military tendency to rush into precise solutions to the wrong problem, without full consideration for context. This has also been recognised by US commanders"* [12]. These commanders see a need to develop methods to define a problem's framework before performing data analysis, information fusion, etc. to construct an abstract COP and start to solve the problem at hand. This is work that has traditionally been performed in intelligence units [13], but should be integrated into all functions of the headquarters.

The workshop highlighted three different analysis subclusters that are all related to information processing: finding information, compiling information, and detecting anomalies in information.

The problem of finding information exists at many different scales. For instance, a common case would be to retrieve a set of similar documents dealing with a certain topic. If all documents are meta-tagged with their topic, then that process is fast. The meta-tagging can potentially be performed automatically using semisupervised learning. Salakhutdinov and Hinton [14] use deep learning in the form of a deep autoencoder[2] to transform very high-dimensional document input vectors (normalized word counts) into a low-dimensional latent vector space in which neighboring vectors correspond to similar documents. Learning the autoencoder itself can be done in an unsupervised fashion. An automatic meta-tagging algorithm can then be constructed by defining specific clusters in the latent space using a few topic-labeled samples.

Finding information at a different scale identifies the entities that are relevant to a certain topic within a document. For instance, extracting items and quantities related to logistics can accelerate the planning process for troop movements, etc. In natural language processing, this problem is referred to as named-entity recognition, and neural networks in combination with named entity dictionaries have shown good results [15].

The final subcluster contained ideas related to anomaly detection. State-of-the-art methods for anomaly detection are currently using deep autoencoders as a foundation [16]. The "normal" data points are assumed to lie on the nonlinear lower-dimensional embedding modeled by the autoencoder and thus have low reconstruction error when decoded by the autoencoder. In contrast, outliers tend to have larger reconstruction errors. The described method is applicable to a wide range of problems, from detecting anomalies in the incoming sensor data to flagging reports that are very different from the norm. The specific idea tied to the anomaly detection cluster was detecting conflicting information in reports and orders. Recent work has shown some promising results in classifying whether two sentences are in conflict [17].

4.2 AI for Planning

For planning operations, AI in combination with simulations is a fruitful combination. Those who have to plan military operations can perform *what if*-tests to measure the expected effects of different plans [18, 19]. The goal is to simulate as realistically as possible the different effects that military operations will have. This includes both impacts on the battlefield and effects on other factors such as morality, logistics, and refugees.

It is also important that the military knowledge obtained during exercises can be used as an aid in the decision support systems used for planning. This knowledge is needed in planning for generating the objectives to be achieved, for effective resource allocation, and during the execution of operations for monitoring operational development and to propose the replanning of activities as needed.

Plans can be analyzed with qualitative or quantitative methods during the planning process and prior to execution. In a qualitative approach for analyzing *courses of action* (COA), a framework that highlights similarities and differences between argumentation models can be used to select and refine arguments critiquing military COA [20]. Such a framework is useful in decision support systems that can argue for and against military plans. When several COA are proposed by different planning groups, the conceptual framework is used to register the domain experts' criticisms of these COA. To create structured criticism and to systematically assess certain aspects of COA, a template is provided to the experts. This method helps to provide a structured analysis of alternative COA during the planning phase.

[2] An autoencoder is a special artificial neural network architecture with an encoder part and a decoder part. In general, the encoder part compresses the high-dimensional input into a low-dimensional latent representation. The decoder reconstructs the high-dimensional input from the low-dimensional representation.

As an alternative to qualitative analysis, quantitative methods can be used. One example is the combination of AI and multiagent systems for Red Teaming [21]. Red Teaming has a long tradition in military planning and decision-making. A Blue Team represents the purpose, goals and interests of our own side, while enemies are represented by a Red Team. By allowing a Red Team to imitate the enemy's motives, intentions, behaviors and expected actions, its own side can test and evaluate its own action options, identify opportunities to exploit the weaknesses of the enemy, and learn to understand the dynamics of how blue and red interact. Red Teaming is a way to understand all devices that have the potential to affect a system and its decision-making. Essentially, an enemy is a unit that has goals that compete with ours and that take actions that prevent us from achieving our goals. Here, AI and multiagent systems can be integrated to support decision-making and planning. It allows decision-makers to explore possible event developments that can affect the goals, discover and evaluate our own vulnerabilities, learn to understand enemy behaviors and find strategies to win.

Also worth mentioning is the recent work on Developing Actionable Data Farming Decision Support for NATO (MSG-124) that uses data farming methodology (i.e., massive parallel simulations, data analysis and visualization) to analyze the outputs from simulation systems with hundreds of thousands of alternative simulations of operational plans for ground warfare [22, 23]. This is a qualitative approach that combines simulations with big data analytics.

4.3 AI for Execution

When executing operations, it is important to quickly get information from the battlefield that can be fused and analyzed by AI methods into the hands of the commander. The commander needs the information to quickly make critical decisions in stressful situations. The amount of information processed and delivered to the commander is often so high that there is a significant risk of information overload. The problem arises if the information is not presented in a logical, concise and meaningful manner that is understood by the commander.

In addition to AI and information fusion, high-level simulation is an important methodology within the framework of a decision support system where simulations can interact with AI methods. Moffat and Witty [24] have developed a model of decision-making and military command that helps to provide insight into the military decision-making process. In this model, a military operation can be seen as a sequence of subsequent confrontations. The model is based on game theory with confrontation analysis. The perceptions of the different sides of the confrontation are based on their perception of the current situation and the alternative actions they have at their disposal. The model can be used in high-level simulations to evaluate operations within the framework of a decision support system.

Since 2008, the Defense Advanced Research Projects Agency (DARPA) has developed a technology called Deep Green (DG) for military tactical command and control. DG helps commanders to discover and evaluate more action alternatives and thus proactively manage an operation. The method behind DG aims to get inside the opponent's OODA-loop. The idea is that decision-making should be so fast that the OODA-loop is broken up into an extremely fast OO-loop that provides a customized DA-loop with the current situation information being used to simulate many combinations of their own and their opponent's decisions, as well as simulate and evaluate these options. The program was transferred to the US Army in 2013. DARPA has taken further steps after DG, and has recently conducted a research project entitled Real-time Adversarial Intelligence and Decision-making (RAID) using predictive analysis, AI, and simulations to analyze opponents' actions [25].

RAID develops technology to assist a tactical commander to estimate the position, strength and purpose of hostile forces and to predict their likely tactical movements as they strive to effectively combat the opponent's actions. This includes the recognition of the opponent's intention, the prediction of the opponent's strategy, the detection of deception, the planning of their own deception, the generation of a strategy, etc. These problems occur in the military planning of operations, the execution of operations, intelligence analysis, etc. To achieve this, RAID combines AI for planning with cognitive modeling, game theory, control theory, and machine learning.

Machine learning can also be used to develop tactics for combat. However, many machine learning algorithms are not fast enough to find optimal behaviors of intelligent agents in applications such as air combat. Q-learning [26], which is a reinforcement learning algorithm, has been successfully evaluated for air combat target assignment [27]. The algorithm learns optimal state-action pairs for agent behaviors without using any large data sets or *a priori* information.

Sometimes, we need to learn behavioral rules for a sequence of similar scenarios. In this situation, we may use transfer learning to reduce the learning time. For example, when we learn combat rules for air combat in different 2-versus-2 scenarios, we may start by using agents that already have experience in fighting in 2-versus-1 scenarios. An experiment showed that experiences that were already obtained in 2-versus-1 scenarios were very advantageous in 2-versus-2 scenarios because further learning was minimal. Using transfer learning in this way can lead to the rapid development of an agent's behaviors for new scenarios [28].

As mentioned in Section 3, AI for execution is also focused on making the staff work more efficiently during the execution of operations. One of the ideas mentioned was the automatic generation of report summaries. The identified need comes from the hierarchical organization structure where each upper echelon receives reports from connected lower echelons, and thus an exponentially increasing size of information is potentially forwarded upwards if no summarization is performed. In the past, automatic text summarization has been of the extractive type, which cuts and pastes relevant full sentences from the original documents. With the recent deep learning technique of sequence-to-sequence modeling [29], abstractive methods for summarization have emerged [30, 31]. The abstractive methods are able to produce summaries where novel formulations that are not present in the original documents are produced.

Another idea that was mentioned was transcribing speech to text. Machine learning has been the foundation of speech recognition systems since the rise of computers. Today's state-of-the-art algorithms are all based on deep learning techniques. For example, the algorithm presented by Microsoft in 2017 was able to reach error-rates on a par with humans [3].

5 Challenges

In this section, we discuss some potential challenges when incorporating AI into decision support systems. Specifically, we discuss the feasibility and explainability of current AI technologies.

5.1 Feasibility

The different ideas surfaced at the workshop have different technological maturity. For instance, the Joint Assistant for Deployment and Execution (JADE), an AI-based logistic planning tool, has been in use by the US military for a long time [32] and the US Naval

Research Laboratory has developed a mission planning and training tool called Sniper-RT[3]. The latter tool is built around 3D terrain data and can answer questions of the type "what can I see" or "where can I be seen," which are crucial questions when placing sensors or protecting forces. Another technologically mature AI problem is automatic speech recognition. Microsoft, Google, Amazon and others all have products leveraging the latest deep learning technologies for speech-based dialog systems.

Among the natural processing language applications listed in Sections 4.1 and 4.3, efficient algorithms for finding similar documents are sufficiently mature to be used in real systems. Slightly less matured are the techniques for both named-entity recognition and automatic summarization. However, commercial systems exist (c.f. www.primer.ai). The most difficult problem (of the analysis ideas listed), and thus the least matured, is finding contradictions between documents. The learning algorithm to discover contradictions requires another layer of abstract reasoning compared to the more straightforward problems of classification.

As proven by the latest advances in AI, the availability of massive amounts of data is fundamental to achieving powerful AI systems. Depending on the scenario or application, this can be a challenge to obtain in some military contexts. Techniques such as transfer learning, where machine learning models trained for a similar but different application are reused and adapted to the new problem, will be important for many military applications when data is scarce. For instance, military reports and summaries are different from the civil equivalents. However, given the similarity, one would expect that having a summarization algorithm trained on nonmilitary text would be a good starting point for a machine learning model to learn the summarization for a specific military use case. Kruithof *et al.* [33] examined how much input data one needs for deep learning to achieve better classification performance compared to when transfer learning is used.

5.2 Explainable AI

A decision support system being able to explain its recommendations is crucial for decision-makers to be able to understand and rely on the system [34]. Within the explainable artificial intelligence area, the focus is on classification of heterogeneous data, planning, data generation, and creation of decision policies. The research area aims to create machine learning methodologies with explanatory models [35] in which machine learning systems are able to explain their recommendations and describe the strengths and limitations of their own reasoning.

This field of research is not new. It has been around for decades but is further accentuated by the increasing use of machine learning that operates at a subsymbolic level. There are several ways that AI systems can explain their recommendations. First, some types of models are perceived as more interpretable than others, such as linear models, rule-based systems, or decision trees. The inspection of such models gives an understanding of their composition and computation. Furthermore, interpretable models may be used to approximate the reasoning of subsymbolic AI systems. The approximate reasoning may sample either the system's whole decision region or the area around a specific decision point [36].

Additionally, hybrid systems are conceivable where a subsymbolic machine learning (e.g., deep learning) level is connected with a symbolic level where approximate reasoning is performed to combine uncertain data from different reasoning processes into a decision support basis. Such an explainable AI will connect machine learning to higher-level

---

[3] www.nrl.navy.mil/techtransfer/available-technologies/IT/sniper-rt (May 2018).

approximate reasoning and decision-making. It will provide decision-makers with explanations whenever decisions are partly based on machine learning results.

To provide insights into the working of deep neural networks, it is important to develop a probabilistic interpretation of neural networks in which weights are seen as probabilities, and the network is partitioned by a second explanatory process into subnetworks based on common information processing behaviors among neurons. This partition may indeed be performed by another machine learning module. Several different approaches may be considered for this secondary task (e.g., Kohonen networks [37]). Each subnetwork (cluster) can then be mapped to a node in a decision tree, which can be analyzed from an explainability point of view by investigating each node's influence on the overall conclusion reached by the decision tree. Such an approach will thus move from a problem-solving ability at a detailed subsymbolic level to a problem-explaining ability at an aggregated symbolic level.

Applications that may use deep learning with neuro-symbolic reasoning and explainable AI come from a pool of generic applications that either have big historical training data, data available from a simulator, or streaming data of a type that is not necessarily fully known in advance. These applications contain a problem that needs to be managed (and may develop dynamically over time) that requires high level approximate reasoning to integrate information from different sources, including machine learning processes, into a decision support that presents solutions to management problems.

Another active research area in explainable AI is feature visualization where subsymbolic reasoning is mapped back into the input space. Typically, two general approaches are used for feature visualization, namely, activation maximization and DNN explanation. Activation maximization computes which input features will maximally activate the possible recommendations [38]. DNN explanation explains the system recommendations by highlighting discriminating input features that may be calculated with a sensitivity analysis using local gradients or some other measure of variation [38].

Future explainable AI will likely approach how people in general explain other agent's behaviors in terms of their perceived beliefs, desires, and intentions. Miller [39] provides an extensive review of explanations in social sciences research and how this knowledge may be used to design explanations for AI systems. The major findings are that (*i*) explanations are contrastive in response to particular counter-factual events, (*ii*) explanations are selected and focus on one or two possible causes, and (*iii*) explanations are a social conversation and interaction for the transfer of knowledge.

Finally, for military decision support systems that already argue at a higher symbolic level, explanatory features based on sensitivity analysis are an established method that can be used to explain why a certain proposed military plan is thought to be successful [40]. Another example proposed by van Lent *et al*. [41] describes an AI architecture for explaining the tactical behavior of an AI agent in a field combat simulation system. The methodology is used by the US Army for the training of infantry officers.

6 Conclusions

The study presented herein has served to investigate how AI can be used for empowering the decision support functionality in future C2 systems. The study has pointed out different AI perspectives, identified areas where AI tools are likely to make a difference, and highlighted concrete C2 tasks that hold the potential to benefit the most from the insertion of AI functionality.

From a C2 systems modeling perspective, the study points to three primary activities in the C2 process where AI functionality ought to be considered, namely, (*i*) sensemaking, (*ii*) planning, and (*iii*) military activity, according to the well-accepted DOODA-loop as depicted in Figure 1. To facilitate the sensemaking process, tools for managing and making use of different pieces of information at various scales can be anticipated to provide easily achievable advantages. For planning, tools for working with tactical databases (terrain, logistics, doctrine, etc.) could be combined with decision support tools to make it possible for the commander to be able to evaluate different courses of action at different abstraction levels. Finally, AI support for execution may include the evaluation of action alternatives for the commander, and facilitating different kinds of staff work during the execution of operations, such as using speech-to-text tools for the quick and correct communication of different briefings.

The AI challenges to be considered from an end-user military-specific perspective mainly concern maturity and transparency. Considering the feasibility, it is not surprising that the ideas emerging from the study relate to different technical maturity levels in terms of R&D. Some tools, e.g., speech-to-text tools, terrain analysis functionality, etc., are already fairly mature and can be bought off-the-shelf, whereas other areas, e.g., game-theoretic tools for reasoning about a willful opponent [42, 43, 44], will require many more years of basic research progress before their actual functionality can be implemented. Concerning transparency, this is a crucial challenge to be considered for military decision support, where it is vital to be able to explain recommendations, and to be able to understand and rely on the system [45]. There is still much to be learnt concerning transparency with the active research field of explainable AI showing promising results.

In the future, we aim to perform a series of follow-up user-centered design activities with the aim of specifying a set of elaborated use cases, which can be used as a basis for the procurement and further testing of actual AI functionality in a military C2 setting involving military personnel.

7 References

1. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature* **550**:354–372. doi:10.1038/nature24270

2. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**:115–118. doi:10.1038/nature21056

3. Xiong, W., Wu, L., Alleva, F., Droppo, J., Huang, X., Stolcke, A. (2017). The Microsoft 2017 conversational speech recognition system. Microsoft Technical Report MSR-TR-2017-39. arXiv preprint arXiv:1708.06073

4. McCann, C., Pigeau, R. (1999). Clarifying the concepts of control and of command. In *Proceedings of the Command and Control Research and Technology Symposium*. Washington, DC: US Department of Defense CCRP, paper 19.

5. Brehmer, B. (2010). Command and control as design. In *Proceedings of the 15th International Command and Control Research and Technology Symposium*. Washington, DC: US Department of Defense CCRP, paper 182.

6. Ayoub, K., Payne, K. (2015). Strategy in the age of artificial intelligence. *Journal of Strategic Studies* **39**(5–6):793–819. doi:10.1080/01402390.2015.1088838
7. Boyd, J. R. (1987). A discourse on winning and losing, Document No. MU-43947. Maxwell AFB, AL: Air University Library.
8. Weick, K. E., Sutcliffe, K. M., Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science* **16**(4):409–421. doi:10.1287/orsc.1050.0133
9. Abras, C., Maloney-Krichmar, D., Preece, J. (2004). User-centered design. In Bainbridge, W. S. (Ed.), *Berkshire Encyclopedia of Human-Computer Interaction: When science fiction becomes science fact*, volume 2. Great Barrington, MA: Berkshire Publishing Group, pp. 763–768.
10. Simonsen, J. Robertson, T. (2013). *Routledge International Handbook of Participatory Design*. New York: Routledge.
11. Ahlberg, S., Hörling, P., Johansson, K., Jöred, K., Kjellström, H., Mårtenson, C., Neider, G., Schubert, J., Svenson, P., Svensson, P., Walter, J. (2007). An information fusion demonstrator for tactical intelligence processing in network-based defense. *Information Fusion* **8**(1):84–107. doi:10.1016/j.inffus.2005.11.002
12. Turner, P., Dodd, L. (2016). Developing the cognitive and social aspects of military ‘understanding capability’. In *Proceedings of the 21st International Command and Control Research Technology Symposium*. Washington, DC: International Command and Control Institute, paper 69.
13. Brynielsson, J., Horndahl, A., Kaati, L., Mårtenson, C., Svenson, P. (2009). Development of computerized support tools for intelligence work. In *Proceedings of the 14th International Command and Control Research and Technology Symposium*. Washington, DC: US Department of Defense CCRP, paper 48.
14. Salakhutdinov, R., Hinton G. (2009). Semantic hashing. *International Journal of Approximate Reasoning* **50**(7):969–978. doi:10.1016/j.ijar.2008.11.006
15. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research* **12**(Aug):2493–2537.
16. Zhou, C., Paffenroth, R. C. (2017). Anomaly detection with robust deep autoencoders, In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, pp. 665–674.
17. Lingam, V., Bhuria, S., Nair, M., Gurpreetsingh, D., Goyal, A., Sureka, A. (2018). Deep learning for conflicting statements detection in text. In *PeerJ PrePrints*.
18. Schubert, J., Moradi, F., Asadi, H., Luotsinen, L., Sjöberg, E., Hörling, P., Linderhed, A., Oskarsson, D. (2015). Simulation-based decision support for evaluating operational plans. *Operations Research Perspectives* **2**:36–56. doi:10.1016/j.orp.2015.02.002
19. Brynielsson, J., Lindquist, S., Luotsinen, L. (2016). Efficient implementation of simulation support for tactical-level military training. In *Proceedings of the 2016 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC 2016)*. Arlington, VA: National Training and Simulation Association (NTSA), paper 16292.
20. Bentahar, J., Moulin, B., Bélanger, M. (2010). A taxonomy of argumentation models used for knowledge representation. *Artificial Intelligence Review* **33**(3):211–259. doi:10.1007/s10462-010-9154-1

21. Abbass, H., Bender, A., Gaidow, S., Whitbread, P. (2011). Computational red teaming: past, present and future. *IEEE Computational Intelligence Magazine* **2011**(Feb):30–42. doi:10.1109/MCI.2010.939578

22. Schubert, J., Seichter, S., Zimmermann, A., Huber, D., Kallfass, D., Svendsen, G. K. (2017). Data farming decision support for operation planning. In *Proceedings of the Eleventh Operations Research and Analysis (OR&A) Conference*. Neuilly-sur-Seine: NATO Research and Technology Organisation, paper 7.2.

23. Horne, G., Seichter, S., Åkesson, B., Balestrini-Robinson, S., Britton, M., De Reus, N., Döring, S., Gruber, T., Hazard, C., Huber, D., Hörling, P., Johansson, R., Kallfass, D., Lappi, E., Morabito, G., Ng, K. Y. K., Nissinen, N., Rindstål, P., Schubert, J., Schwierz, K.-P., Svendsen, G. K., Thorén, P., Tuukkanen, T., Ürek, B., Zimmermann, A. (2017). Developing Actionable Data Farming Decision Support for NATO. Technical Report RDP STO-TR-MSG-124. Neuilly-sur-Seine: NATO Research and Technology Organisation, to appear. doi:10.14339/STO-TR-MSG-124.

24. Moffat, J., Witty, S. (2002). Bayesian decision making and military command and control. *Journal of the Operational Research Society* **53**(7):709–718. doi:10.1057/palgrave.jors.2601347

25. Kott, A., Michael, O. (2015). Toward a research agenda in adversarial reasoning: Computational approaches to anticipating the opponent's intent and actions. arXiv preprint arXiv:1512.07943

26. Watkins, C. J. C. H., Dayan, P. (1992). Q-learning. *Machine Learning* **8**(3–4):279–292. doi.10.1007/BF00992698

27. Luo, P.-c., Xie, J.-j., Che, W.-f., (2016). Q-learning based air combat target assignment algorithm. In *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics*. Piscataway, NJ: IEEE, pp. 779–783. doi:10.1109/SMC.2016.7844336

28. Toubman, A., Roessingh, J. J., Spronck, P., Plaat, A., van den Herik, J. (2015). Transfer learning of air combat behavior. In *Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications*. Piscataway, NJ: IEEE, pp. 226–231. doi:10.1109/ICMLA.2015.61

29. Sutskever, I., Vinyals, O., Le, Q. V. (2014). Sequence to sequence learning with neural networks, In *Advances in neural information processing systems*. La Jolla, CA: NIPS, pp. 3104–3112.

30. See, A., Liu, P. J., Manning, C. D. (2017). Get to the point: Summarization with pointer-generator networks. arXiv preprint arXiv:1704.04368

31. Rush, A. M., Chopra, S., Weston, J. (2015). A neural attention model for abstractive sentence summarization, In *2015 Conference on Empirical Methods in Natural Language Processing*. arXiv preprint arXiv:1509.00685

32. Mulvehill, A. M., Hyde, C., Rager, D. (2001). Joint assistant for development and execution (JADE), Final Technical Report AFRL-IF-RS-TR-2001-171. Rome, NY: Air Force Research Laboratory.

33. Kruithof, M. C., Bouma, H., Fischer, N. M., Schutte. K. (2016). Object recognition using deep convolutional neural networks with complete transfer and partial frozen layers. In *Proceedings SPIE, Vol. 9995, Optics and Photonics for Counterterrorism, Crime Fighting, and Defence XII*. Bellingham, WA: SPIE, paper 99950K. doi:10.1117/12.2241177

34. Lundberg, S. M., Lee, S.-I. (2016). An unexpected unity among methods for interpreting model predictions. In *Proceedings of the Workshop on Interpretable Machine Learning for Complex System*. La Jolla, CA: NIPS. arXiv preprint arXiv:1611.07478

35. Biran, O., Cotton, C. (2017). Explanation and justification in machine learning: A survey. In *Proceedings of the Workshop on Explainable Artificial Intelligence (XAI)*. Menlo Park, CA: IJCAI, pp. 8–13.

36. Ribeiro, M. T., Singh, S., Guestrin, C. (2016). Why should i trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: ACM, pp. 1135–1144.

37. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43**(1):59–69. doi:10.1007/bf00337288

38. Montavon, G., Samek, W., Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing* **73**:1–15. doi:10.1016/j.dsp.2017.10.011

39. Miller, T. (2017). Explanation in artificial intelligence: Insights from the social sciences. arXiv preprint arXiv:1706.07269

40. Schubert, J., Hörling, P. (2012). Explaining the impact of actions. In *Proceedings of the 15th International Conference on Information Fusion*. Piscataway, NJ: IEEE, pp. 354–360.

41. van Lent, M., Fisher, W., Mancuso, M. (2004). An explainable artificial intelligence system for small-unit tactical behavior. In *Proceedings of the 16th Conference on Innovative Applications of Artificial Intelligence*. Palo Alto, CA: AAAI Press, pp. 900–907.

42. Brynielsson, J. (2006). A Gaming Perspective on Command and Control, PhD thesis. Stockholm, Sweden: KTH Royal Institute of Technology.

43. Brynielsson, J., Arnborg, S. (2006). An information fusion game component. *Journal of Advances in Information Fusion* **1**(2):108–121.

44. Brynielsson, J. (2007). Using AI and games for decision support in command and control. *Decision Support Systems* **43**(4):1454–1463. doi:10.1016/j.dss.2006.06.012

45. Svenmarck, P., Luotsinen, L., Nilsson, M., Schubert, J. (2018). Possibilities and challenges for artificial intelligence in military applications. In *Proceedings of the NATO Big Data and Artificial Intelligence for Military Decision Making Specialists' Meeting*. Neuilly-sur-Seine: NATO Research and Technology Organisation, paper S1-5.