Interactive OODA Processes for Operational Joint Human-Machine Intelligence

Leslie M. Blaha
Pacific Northwest National Laboratory
Richland, Washington
UNITED STATES OF AMERICA

Leslie.Blaha@pnnl.gov

ABSTRACT

A key advantage to strategic thinking with the Observe-Orient-Decide-Act (OODA) framework is that it provides a systematic approach to get inside the decision-making process of another agent, either cooperative or adversarial. Indeed, current OODA concepts have supported understanding human decision processes to support agile and competitive decisions about human warfighters and human-centric operations. However, future military decision making based on human-machine teaming relies on technology and interaction concepts that support joint human-machine intelligence, not just human capabilities. This requires new OODA concepts. Herein, I define a machine OODA loop, considering the characteristics that make it similar to and different from the human OODA loop. I consider how advances in artificial intelligence and cognitive modeling can be integrated within the machine-Orient stage, providing the machine a unique advantage over humans in that the machine can integrate a level of understanding and prediction about human operators together with predictions about machine behaviors and data analytics. Additionally, I propose that effective human-machine teaming should be supported by human-machine joint decision-action processes, conceptualized as interacting OODA loops. Consideration of the interacting human-machine OODA processes offers conceptual guidance for design principles and architectures of systems supporting effective operational human-machine decision making.

1.0 INTRODUCTION

Reasoning about the decisions and actions of others is a cornerstone of military strategy and decision making. Indeed, Sun Tzu observed in *The Art of War*, around 400 BC, that probing enemy strengths and weaknesses was key to manipulating events to exploit the weaknesses and thwart the strengths. In modern times, where information warfare and cybersecurity operations are the new battlefield, reasoning about the decisions and actions of both adversaries and allies continues to be key to successful offensive and defensive missions. As intelligent machines and autonomous systems integrate into all aspects of military operations, we must additionally reason about the decisions and actions of machine agents (with machine agents broadly referring to all intelligent machines, virtual or mechanical, in operational settings). For approximately 50 years, the Observe-Orient-Decide-Act (OODA) framework has provided a means of strategizing about our own and others' strategic behaviors in military operations to meet an age-old goal: out-maneuver the adversary's decisions and actions to win.

I propose that the simplicity of the OODA conceptual model can be adapted to aid reasoning about all agents active in the modern battlespace, human or machine, adversarial and cooperative. The goal of this work is to introduce a new conceptual machine OODA process, for reasoning about intelligent machines in operational settings. I explore what makes the machine OODA different and similar to human OODA processes. I then propose ways to combine human and machine OODA processes into conceptual models for whole human-machine systems. This can aid in reasoning about human-machine teaming, human-in-the-loop systems, and autonomous systems operations. The objective is to provide a complete language with which we can consider the strategic strengths and weaknesses of joint human-machine decision-action processes for operational settings; I will not attempt to claim there is a single definitive solution nor how such solutions should actually be constructed.

This paper is structured as follows. First, I review human OODA loops and variations that have been developed within that framework. Next, I define the new machine OODA loop, and discuss how to give machines a strategic reasoning advantage over human OODA loops. Finally, I introduce ways to conceptualize human-machine teams as interacting OODA processes. Before delving into the details of the OODA loop conceptual approach, diagrammed in Figure 1, I will motivate the reasoning behind developing an OODA loop for machine intelligence.

1.1 Why OODA?

The key to strategic thinking with OODA is to use this conceptualization to get inside the decision making process of another agent, either cooperative or adversarial or even both simultaneously. On the surface, it seems an advantage gained from executing one's own OODA faster than a competing agent; that is, the objective is to get ahead of the opponent's decision process. But there are limits to the speed of human information processes, as well as severe limitations in human abilities to predict the future. So, we are limited in our abilities to only be faster than another person. Thus, the real strength of using OODA to conceptualize another agent's decision making is in getting into that other agent's process. This allows us to better predict the likelihood of various decisions and the likelihood of the subsequent actions that agent might take based on the decision options. OODA provides a means of outlining the critical information elements that are needed to accurately conceptualize the other agent's OODA process. Indeed, it has been suggested that there is a critical time window for disrupting the adversary between the Act and Observe stages (Blasch, Breton, Valin & Bosse, 2011). That is, if we can change the state of the world as an adversary is acting on it, there will be inconsistencies in the adversary's Observe stage that will disrupt his battle tempo. Thus, it is advantageous to study and develop models of the other agents in the domain of interest. These models will enhance our understanding and may lead to insights about other agents' OODA loops. By exploring conceptual models of other agents, we can also identify the potential weaknesses our own or our allies' OODA loops. In some ways, introspection about the likelihood of our teammates' decisions and actions may prove more advantageous than reasoning about adversaries. It can enable more effective teaming by predicting when teammates might make errors and when they need additional support from the rest of the team.

In the recent push for more artificial intelligence and autonomous systems operating with human warfighters (e.g., see Endsley, 2015), we are seeking solutions where intelligent machines may act as true teammates, as other warfighters might. This is a shift from considering machines as simply subordinate tools or electronic assistants. In the role of teammate, machines may reason independently, take actions independently, and even take initiative to direct human assets during missions. And surely as our own technology becomes our partner on the battlefields, so too will our adversary's technology become our adversary. Thus, we need a conceptual framework in which to reason about the decisions and actions of autonomous systems, just as we do humans. By adapting the OODA framework for machines and human-machine teaming, we can establish a common language for reasoning about all the agents operating in modern and future conflicts.

Additionally, as more of the defense domain shifts to information-centric operations in Command and Control environments, as in cyber and space missions, we will likely become more dependent on artificial intelligence to help provide the information critical for human decision making. These domains are dominated by large volumes of streaming data. We depend on machines to carry the burden of processing the volumes of big data and distilling it into information tractable for human consumption; we desire interactive streaming analytic systems that allow humans to interact with the machine intelligence to shape the analytic processes to the mission critical tasks on hand. Consequently, we need to develop human-machine systems capable of supporting interactive analytics that produce correct, timely, and useful interpretations of the world. Adopting an OODA loop perspective to reason about the capabilities of the humans and machines in this solution space will help us find robust ways to develop human-machine team solutions.

2 BOYD'S OODA LOOP FOR HUMANS

The cyclical, dynamic Observe-Orient-Decide-Act (OODA) process was originally introduced by United States Air Force Col. John Boyd as a strategic method for conceptualizing battlefield decision making. Boyd proposed that successful military decision making required fast, agile human decisions, not just larger machines or more deadly weaponry. His tactical theories derived from his studies of fighter pilot combat dogfights (both Korean and Vietnam Wars) and historical combat strategy dating back to Sun Tzu (Boyd, 1987). He argued that the goal of military tactics should be to operate in a manner to get inside of the adversary's decisions and actions, to "…enmesh the adversary in a world of uncertainty, doubt, mistrust, confusion, disorder, fear, panic, chaos, … and/or fold adversary back inside himself so that he cannot cope with events/efforts as they unfold" (Boyd, 1987, p. 7). That is, we need to understand our adversary's OODA process to think ahead of the adversary.

Dynamic decision environments, like military operations, necessitate that humans perform well under changing conditions with some ever-present and variable degree of uncertainty. Decision making under such uncertainty requires a continual process of deconstructing and reconstructing one's understanding of the situation at hand, as well as what decisions and actions are appropriate in the changing circumstances (Boyd, 1976). The OODA loop evolved as a way to describe the stages of gathering, processing, and updating one's understanding of the changing environment and the corresponding decisions and actions taken. While conceptualizing our own decision processes with OODA loops is helpful for reflecting on past decisions, the real advantage to this conceptualization is thinking through the adversaries decisions. This can support both real-time anticipation of adversarial tactics as well as after-action review and preparation for future engagements.

Less frequently discussed is the perspective that OODA loops can also enable thinking through the decision-action processes of our allies and teammates. Getting into the OODA loop of an ally means predicting what conditions and actions will result in high degrees of uncertainty that is undesirable in strategic operations. Rather than using OODA to disrupt, we want to use OODA reasoning to mitigate or intervene to keep teams working positively together and enabling teams to focus on the adversary. Boyd (1987) referred to this as identifying the interactions that foster harmony and allow operators to take initiative to exploit variety/rapidity that will disrupt the adversary. Both human-human teams as well as human-machine teams, which is the emphasis herein, benefit from these positive interactions. But before we discuss leveraging OODA models for human-machine teaming, I will review the basic components of the OODA loop.

Note importantly that the OODA loop is a conceptual or descriptive model, not a detailed process model. Indeed, it is often reported, or criticized, for being one of the more high-level descriptions of time critical decision making (Azuma, Daily, & Furmanski, 2006). The OODA loop might be thought of as a simple representation of a control process, where the internal operations of the human adjust to the external changes in the environment. However, though there is a tendency for some accounts of the OODA loop to draw a simple 4-stage loop, the OODA loop as originally conceived is not that simple. There is feedback between all stages, such that the internal state of the decision maker, the responses to the various cognitive processes, and the resulting changes on the environment become another critical source of information for the ongoing observations throughout task execution. Figure 1 captures the OODA process, derived from Boyd (1996) and consistent with the detailed descriptions by Boyd (1987) and Fadok (1995). I discuss each component stage of the process in turn below. Although I define each stage separately, a critical part of the OODA process is that it is cyclical and adaptive to changing circumstances in dynamic tasks based on the interactions and feedback between the various stages.

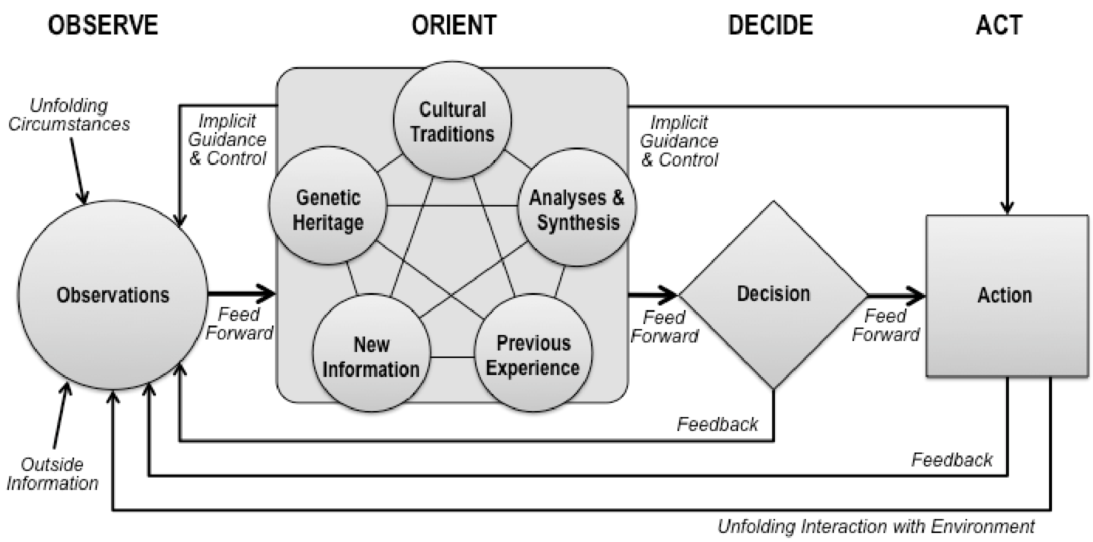

Figure 1. Boyd's Observe-Orient-Decide-Act (OODA) Loop. The image is based on the descriptions from Boyd (1987) and Fadok (1995), and reproduced from the diagram in Boyd (1996).

2.1 Observe

The Observe stage is the point at which information is gathered by the human decision maker. The Outside Information comes through direct interaction of the human with external information, which could be any information source feeding the human. Sensory input includes external sensory information (visual, tactile, auditory, taste, smell) as well as internal sensory information (proprioception, internal cognitive reflection). System information may come through computational systems, social networks, media, planning documents, or any source relevant to the domain at hand. Unfolding Circumstances surrounding the decision maker refer to the dynamic operational environment in which the decision maker is operating. This environment includes the actions of other operators, both on the decision maker's team and adversarial. There may be additional unexpected events that unfold (third party engagement) or other environmental changes (e.g., change in weather) that shape the observation process.

Importantly, the Observe stage receives feedback from every other stage in the OODA loop. This captures the concept of top-down or goal-directed shaping of perception and information gathering. Feedback from the Decide stage acts to shape the observations according to goals and information needs. Any direct consequences of the operator's Act stage on the operator are captured through the Unfolding Interaction with Environment as well as the feedback processes. With the Observe stage at the left end, the diagram emphasizes that the OODA loop is strongly, though not solely, an information driven process.

2.2 Orient

Orientation was the process Boyd discussed in the most depth. He viewed it as the *schwerpunkt*, or the crucial focal point of the OODA perspective (Boyd, 1987), because it shapes all the interactions with the environment, including the Observe, Decide, and Act processes. The Orient state includes all the internal cognitive reflection and information synthesizing needed to integrate new information with the agent's current understanding of the state of the world. According to Boyd (1976), we make sense of the world by continuously unstructuring (differentiation) and restructuring (integration) our perceptions. This is the process by which we determine if or how our understanding of the world needs to be updated. Orient, then, produces the internal representations, or mental models, that we use to examine patterns of activity in the world. Disruption of the Orient stage results in incomplete or erroneous mental models, which causes the friction and weaknesses that can be exploited.

As shown in Figure 1, the Orient stage consists of five components that collectively shape human information processing: Genetic Heritage, Cultural Traditions, Previous Experience, New Information, and Analysis & Synthesis. The first three comprise all the internal processes and biases that an operator brings to a situation, while the latter two are defined and evolve dynamically during operations. Orientation critically shapes the present observations through feedback, and influences both decisions and actions through feed-forward Implicit Guidance and Control. The loop collectively shapes the future orientations. This is a process of re-orientation, which happens continually in a dynamic decision making situation.

2.3 Decide

Following Orientation/Re-orientation processes, the human agent executes the Decide stage. At this point, at least one available course of action is selected, or the user takes the action to loop back to Observe for more information. Decide entails making choices among hypotheses about the environment, events, and adversaries, as well as possible responses to each. Through a combination of feed forward connections, decisions are driven by the operators goals and observations as interpreted by the Orient process. The Decide stage is the driver of Act, through feed-forward connections. It provides feedback to the Observe stage, to inform and guide the information needed to support the current decisions and actions.

2.4 Act

The Decide stage immediately feeds into the Act stage, wherein the selected course of action is executed. In the case that the agent determines that more information is needed before any actions can be taken (a "take no action and observe more" decision), the Act stage is truncated and the operator returns to the Observe stage. In all other cases, the types of decisions and actions are dictated by the domain and the situation. The Act stage is where the operator tests the chosen hypothesis by interacting with the environment. Receives internal guidance and control from Orient and Decide through Implicit Guidance and Control, and could be thought of as goal-directed action. Act provides direct feedback only to the Observe stage. This means that the results of actions can only influence future decisions and actions after the consequences of the action have been synthesized and reconciled with all other information sources in the Orient stage.

2.5 Variations on the OODA Loop

OODA loops are popular for conceptualizing command and control (C2) operations. In fact, some authors refer to OODA as the accepted business model for C2 (Grant, 2005; Révay & Líška, 2017). Yet, some of those same authors are quick to criticize the OODA model as too simple to fully capture the richness of the C2 strategic environment and tactical decision making. Most of the criticisms center around the descriptive nature of the OODA model, which does not make explicit any types of process model that would support implementation of the model and does not make explicit the specific cognitive functions contributing to each stage. Consequently, multiple alternative versions of an OODA process for C2 operations have been proposed that add more detail and functionality to move OODA concepts toward implementable systems.

The modified OODA (M-OODA) loop proposed by Rousseau and Breton (2004) incorporates a set of modules that can be used to explicitly describe the sub-processes within each OODA stage. The goal was to incorporate theory-driven control functions into the descriptive model to enable the construction of models of more complex decision tasks or even team activities. Each module consists of three components, Process, State, and Control, which perform a goal-directed tasks. The goal directed tasks in the M-OODA, replacing the Observe-Orient-Decide-Act nomenclature, are Data-Gathering, Situation-Understanding, Action-Selection, and Action-Implementation. By making these processes more explicit in a process-model format, the M-OODA enables agents to be programmed with internal OODA processes preforming C2 operations.

Other variations on the OODA loop seek to incorporate more dynamics explicitly into the model. In his discourse about strategy, Boyd often commented that decision-action tempo was a key component of the OODA process. The conceptual model, however, doesn't explicitly address tempo or dynamics in any way that helps understand how these processes work for time-critical situations. Consequently, the C-OODA by Breton and Rousseau (2005; see also Blasch, Breton, Valin & Bosse, 2011) and the rationally reconstructed OODA (RR-OODA; Grant, 2005) explicitly integrate processes for sensemaking and situation awareness to emphasize how the OODA loop fits into the larger process of time critical decision making in C2 operations. Brehmer (2005) developed the DOODA loop which explicitly integrated delays in to the process by combining the OODA and a dynamic decision loop. The emphasis of these efforts is that OODA for C2 must support delivery of timely, actionable intelligence.

Bryant (2006), on the other hand, argued that simple augmentations of the OODA were not enough for C2, because the OODA concept was insufficient in that information-driven environment. He developed the alternative Critique-Explore-Compare-Adapt (CECA) loop to address OODA shortcomings. CECA incorporates two parallel processing loops. The first loop entails a combination of a Conceptual Model, Situation Model, and Information Gathering. This loop provides the operational context in which the decision-action processes are to occur. Then the Critique-Explore-Compare processes act on these models to determine information needs, gather information, and compare information to the conceptual model. The gist of the CECA approach is that the cognitive functions and representational models demanded by information-centric C2 operations are made more explicit. Other modelling efforts have also sought to break down complex cognition into information processing cycles, similar to CECA, so are often compared and contrasted directly to the OODA loop as well. Examples include the Stimulus-Hypothesis-Option-Response (SHOR; Wohl, 1981), Plan-Do-Check-Act (Demming, 1951; Shewhart, 1939), Sense-Assess-Augment for cognitive state assessment (Galster & Johnson, 2013), and the Learning from Experience Loop (Kolb, 1984). As reviews like Grant and Kooter (2005) or Azuma et al. (2006) illustrate, the models have more similarities than differences, but all are concerned with detailing out relevant cognitive processes and modelling human-centric decision-action cycles.

2.6 OODA Loops in Cyber Applications

Beyond C2 applications, OODA loops have been proposed as a key element of cyber security operations. Conceptualization in this domain is grounded in the assumption that both human attackers (Zager & Zager, 2017) and malware (Bilar, 2008) already have an adversarial advantage over the cyber defenders. Giordano and Maciag (2002) suggested OODA loops in the forensics process analysing cyber attacks. Dussault and Maciag (2004) took the concept further by explicitly mapping the Observe-Orient-Decide stages from OODA loops to the Detect-Assess-Respond steps of the Protect-Detect-Assess-Respond approach to information operations. They argue that reframing information operations in terms of OODA processes could serve to improve cyber situation awareness for cyber C2 missions. Grant, Venter, and Eloff (2007) took the application of OODA loops to cyber operations further with a proof-of-concept simulation of a system administrator defending against a network intruder. The approach implemented an agent-based OODA-RR model (Grant, 2005) both the administrator and intruder to illustrate how strategic thinking about adversarial behaviors could inform or complement traditional intrusion detection or intrusion management systems. Ohlsson (2006) proposed that entire intrusion management systems could be designed by combining an OODA approach to information security with traditional risk management techniques for cyber security systems. He argued that adopting the OODA perspective could make risk management-based intrusion detection processes more agile and dynamic, improving the response time to agile malicious threats.

3.0 AN OODA LOOP FOR MACHINES

I introduce an OODA process for machines, machine-OODA, illustrated in Figure 2. In conceptualizing an OODA process for machine intelligence or machine teammates, I am not referring to an agent-based or other artificial intelligence implementation of the human OODA process. Rather, I mean developing a higher level conceptual model for how machine intelligence can operate in dynamic, operational settings. The machine-OODA operates independently of human decision-action processes and complements human-OODA. This paper emphasizes teaming, so will illustrate the machine-OODA process defined independently . Then I consider how to effectively team OODA processes. And by keeping this development at conceptual OODA level, we have a broad enough framework to consider different specific instantiations, leveraging computational architecture, process algebras, or other formalisms of choice. It also allows a common conceptual approach applicable to a broad range of machine intelligence, from autonomous systems and embodied robots to pure AI.

The machine-OODA loop is shown in Figure 2, using the same diagram structure as Figure 1 for a direct mapping between concepts. The machine-OODA process begins with a machine-Observe stage, in which the machine, largely through sensor systems, takes in information about the world as its current state. Sensors may include any number of data streams for which the machine is configured. They may include network messages, imagery, sound,

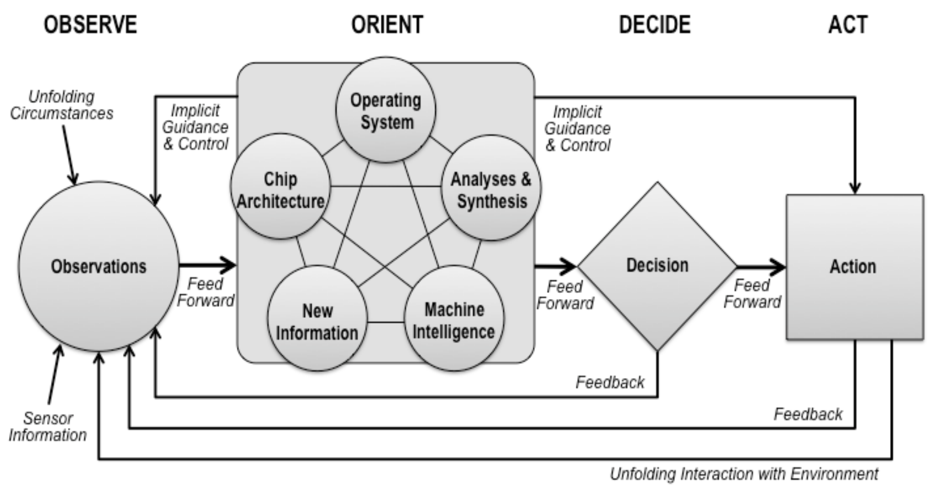

Figure 2. An Observe-Orient-Decide-Act (OODA) loop for machine decision making. Conceptually, it mirrors Boyd's human OODA loop. The key difference is that the machine's Orient stage is comprised of the machine-centric elements that influence machine intelligence but mirror the concepts known to influence human intelligence and reasoning. New Information and Analyses & Synthesis are the only two categories that are the same in both the human and machine Orient stages, though they are undoubtedly implemented in very different ways.

etc. transmitted via physical or network connections. Observation speed is limited only by transmission bandwidth. Sampling rates are limited by the sensor hardware. The machine-Observe stage also incorporates feedback from both the later machine-Decide and machine-Act stages, as well as the Implicit Guidance and Control actions from the machine-Orient stage. The nature of machine processes, when powered on, is that the system operates continuously to respond to dynamic changes within the system, over the network, and from external inputs. In this way, the machine is always in a state of observing its own activity and interactions with the environment unfolding over time.

The information obtained through the machine-Observe stage is fed forward into the machine-Orient stage. Like the human-Orient stage, this is the internal processing stage in which new observations are synthesized together with other machine "reasoning" processes to create an interpretation of the state of the world. Although conceptually similar, the elements of the machine Orient stage have distinct properties, constituting the largest differences between the human-OODA and machine-OODA processes. The set of nodes within the machine-Orient stage, reflecting properties that are unique to computing and artificial intelligence are:

- Chip Architecture: as machines are constructed, they have a Chip Architecture, rather than a Genetic Heritage, that defines the physical organization of the processes within the machine. Modern Chip Architectures include sophisticated neuromorphic computing and various field programmable gate array architectures, each bringing trade-offs in speed and decision making capabilities to the machine-Orient stage.
- Operating System: machines possess an Operating System that organizes the structure and intercommunications of the software, rather than a Cultural Heritage organizing socio-linguistic behaviors. But like Cultural Heritage, choice of Operating System (Windos, Linux, custom, etc.) dictate traditional structures, formats, and (programming) languages with which all the computing machinery is operating.
- Machine Intelligence: rather than Previous Experience, machines leverage Machine Intelligence to shape memory and knowledge. Knowledge may be found in the state space of algorithms after machine learning, or parameter values after optimization. Knowledge may also be in the form of data bases or ontological representations available to machine reasoning systems. Machine Intelligence encompasses all the available artificial intelligence and algorithms available to a system for performing its functions. We note that available memory, storage and RAM, is often much larger for machines than working memory is for humans, setting the expectation that machines will handle much of the demands of large data analytics and reasoning.
- New Information: although the implementation details are very different from humans, machines integrate New Information in the form of sensor feeds and changes to internal algorithm states. Representation is information-theoretic in nature, with bits continuing to be the fundamental computational units. There are known ways to help machines use sensor data in the form of numbers, text, images, and other hybrid forms of incoming data, and the machine-Orient should leverage anything appropriate for the mission.
- Analyses & Synthesis: again, while similar to the human notion, the implementation details will vastly differ nodes in their machine-Orient processes. This refers to how information is integrated within and across processes, like integrating information between programs or dependencies between processes. It can also refer to the approaches taken to handling parallel distribution of computations. Machine synthesis includes any meta-reasoning process needed to combine information to execute the decisions that constitute the Decide stage of processing.

It is couched amongst the interactions of these computational systems design choices that machine interpretations of the world are created or destroyed dynamically over task operations.

The outputs of the machine-Orient stage feed backward to the Observe stage and forward to the Decide Stage. The machine-Decide stage is similar to the human-Decide stage, in that the machine leverages the output of the orientation process to select among the possible response options. In many ways, the set of possible decisions is constrained by the functions for which the system is programmed, though machine learning is increasingly shifting the balance between pre-programmed and learned functions. And with processor speed and bandwidth often much faster and higher, respectively, than humans, it is possible that the machine-Decide stage entails a number of parallel decisions to be made in one pass through the loop.

Once the machine has made at least one decision, it takes the action associated with the decision in the Act stage. Predominantly, the machine actions will affect the machine's internal processes, through the feedback to the machine-Observe process. But it is also possible that the machine can act on the physical world, like human actions can. The machine may move itself, if a vehicle or robotic system, or it may adjust settings on the sensor and data streams, like adjusting the angle or focus of a camera. Rarely, and particularly not in warfighting scenarios, will we allow the machines to take actions with known negative consequences to humans. However, in some cases (e.g., pacemakers) we keep the actions of the machine in a controlled setting to help improve quality of human life. Generally, though, the machine-Act stage only influences the machine's internal processes between the incoming data stream and the display to the human user. This is particularly salient in interactive streaming analytics supporting C2 or other analytic domains.

Note also that although we have drawn the diagram for a single machine, the machine-OODA can also describe a system where multiple types of processing, machine intelligence algorithms, and even multiple types of computational hardware are integrated to meet operational demands. The machine-OODA description is agnostic to the complexity of the machine system; the general process of the decision-action cycle is consistent at this level of conception.

3.1 Augmenting Machine Orientations with Human Models

We can endow the machine-OODA process with a unique advantage over human-OODA loops: we can explicitly give the machine the ability to reason like a human. Years of work in cognitive science and psychology, as well as health and physiology, have develop formalisms appropriate for encoding human physical and cognitive mechanisms in ways that can be incorporated into machine intelligence and machine reasoning. Cognitive architectures, such as ACT-R (Anderson, 2014), essentially offer artificial intelligence languages that perform tasks the way humans do; process models can predict the perception-information processing-decision-action behaviors a human will exhibit given the context of the operational tasks and environment. But the reverse scenario is not possible; we cannot similarly simply program a human to reason like a machine. Thus, we can conceive of systems that give machines an asymmetric advantage that should help the machine to become a stronger teammate supporting the human-OODA process of allies and subverting the human-OODA process of adversaries.

To do this, a critical addition can be made to the machine-Orient stage: a Human Models node. Conceptually, this node would integrate in with the other five Orient nodes, with bidirectional connections to all other nodes in the Orient stage. The Human Models node represents the additional of computational and mathematical models of the human operator. These include, but are not limited to computational cognitive architectures, biomathematical

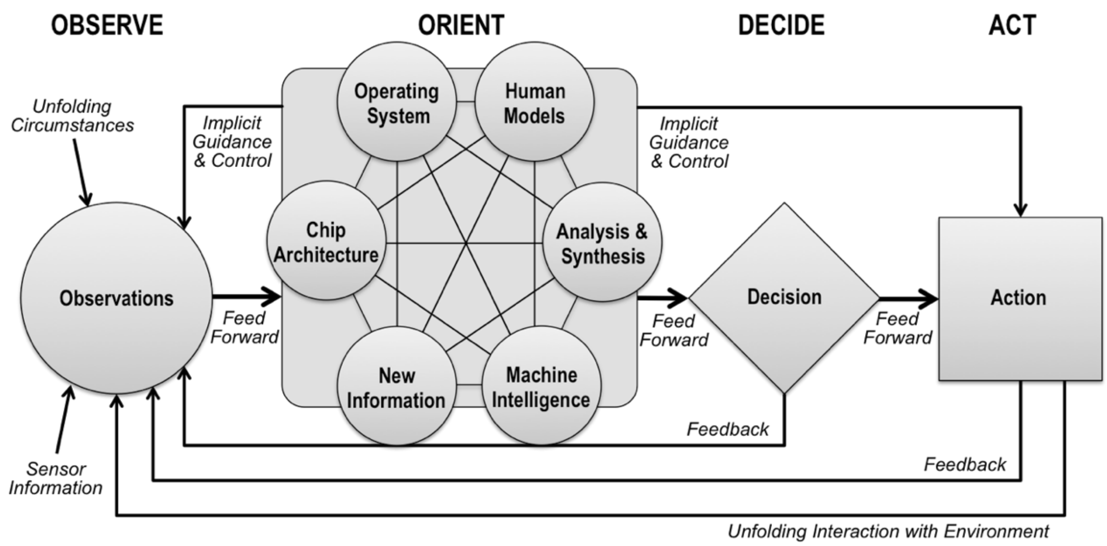

Figure 3. The augmented machine-OODA loop, where the machine's ORIENT loop contains and additional Human Models node. This represents incorporating formal behavioral and cognitive models, computational or mathematical, into the machine reasoning processes.

models of human physiology, heuristic models of decision making, and biomechanical models of human-machine physical interaction capabilities. The purpose of the Human Models node is to integrate formal representations of the human into the machine, so that the machine can interpret and make predictions about the state of the user. In short, the machine would have the capability of reasoning over both its own state and its user's state, situating both against the information coming from the Observe stage. This would enable the machine to computationally predict and adapt to the future states of the user as well as its own predicted future states and interpretation of the world.

Consider, for example, the desired future proposed in *The Quantified Warrior* (Blackhurst, Gresham & Stone, 2012): a next-generation fighter jet is equipped with analytics and machine intelligence to understand the state of the aircraft from the more than 1,000 streams of aircraft integrity data *and* the analytics and machine intelligence to understand the state of the pilot (leveraging all novel sensing and measurement technologies called for in the paper). We know how to solve the first problem, based on extensive modelling of the aircraft and flight dynamics. To equip the aircraft to solve the latter, we must integrate into the machine computers a number of models of the pilot: decision-action behaviour models of interactions with flights controls (e.g., is the pilot executing maneuver X?), physiological models of the impact of environment and flight envelope on the state of the pilot (e.g., is the pilot in danger of hypoxia-induced loss of consciousness?), stress models of voice communications (e.g., does the pilot sound calm?), and socio-behavioral models of interaction with co-pilots (e.g., does the pilot trust that UAS to be her wingman?). We have the computational bandwidth to support such integration of models into the on-board computing. But challenges remain in how to integrate machine reasoning about the human models with the machine reasoning about the machine models, which will require more advancement in artificial intelligence. And there are open challenges in how to communicate the artificial intelligence processes and outcomes to the human operators, including the pilot in the cockpit and the commanders on the ground. Reasoning about OODA concepts for human-machine teaming may help us determine viable solutions to these challenges.

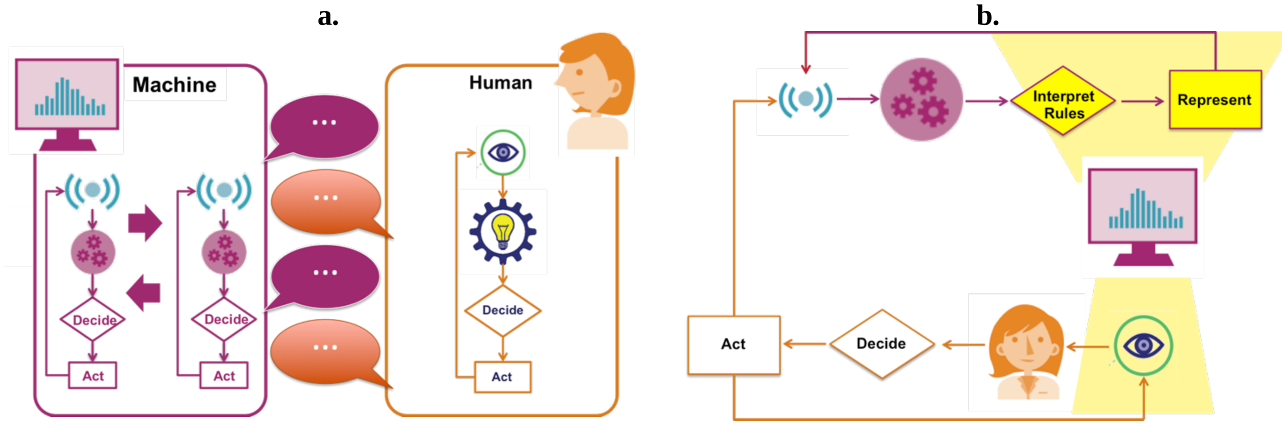

Figure 4. Concepts for interacting OODA loop-based human-machine systems. (a) Concurrent processes system concept. The machine OODA processes (at least one, but parallel computing could have more) and human OODA process are acting in parallel and need some sequential communication to inform each other's processing. Communication could be natural language, interface/display interactions, mechanical inputs/outputs, or other various sensor technology as long as the output of one agent is receivable and interpretable by the other agent. (b) Integrated OODA loops wherein the output of one system is explicitly the input to the other.

4.0 INTERACTING OODA LOOPS FOR HUMAN-MACHINE TEAMING

Human-machine teams require joint human-machine decision making. To achieve this, I propose that we must define the ways that that human and machine OODA loops can interact or operate together to support joint decision making. By conceptualizing a team as interacting OODA loops, we make explicit the ways in which different elements of the uniquely machine and uniquely human separate OODA loops impact each other. Where adversarial reasoning with OODA was about predicting and disrupting the other's process, we want teams to exhibit synchrony and supportive interactions that mutually benefit each team member to achieve the common goals for mission success. Hearkening back to Boyd's original theories, non-cooperative centers of gravity produce friction, which creates confusion and disorder. This in turn produces decision paralysis and system collapse, which is the desirable state when seeking to win in a fight (Boyd, 1987).

On the other hand, in dynamic operational environments, harmony and focus of direction are created by implicit communication and trust bonding a team together (Boyd, 1987). In humans, these evolve as a consequence of similar mental images or impressions (Boyd's orientations) that "…each individual creates and commits to memory by repeatedly sharping the same variety of experiences in the same ways" (Boyd, 1987, p. 18). That is, for human teams, repeated, shared experiences result in common mental models that allows us to effectively reason about the other person's reasoning. And the common Orient processes resulting from common mental models promotes teaming. A big (though not new) challenge is: how do we achieve this harmony and bonding between humans and machines?

The answer to this challenge lies squarely in the argument for what makes good human-human teams: evolving common orientations through shared operational experiences and understanding our allies OODA process. To do this with humans and machines, we must clearly define the ways in which the human and machine OODA processes interact to develop common orientations. Integrating the Human Models into the machine-Orient stage (Figure 3) is one step that will aid this process, because it helps to provide the machine with human-like reasoning and other ways of interpreting human inputs. However, we need to go further to look at how the decision-action behaviors of each system will influence the OODA process of the other. That is, we need to develop a conceptual model for joint decision making through interacting OODA loops. If we are able to successfully interact human and machine OODA loops, we may achieve the benefits, described by Boyd, of implicit orientation between team members: diminished intra-team friction and reduced time to decision making, improve ability to exploit variety/rapidity while maintaining harmony/initiative, get inside adversary OODA, magnify the adversary's friction and stretch out his time for mismatch in our favor, and deny the adversary opportunity to cope with events as they unfold (Boyd, 1987).

Different operational concepts for human-machine teaming can place different demands on the ways humans and machines need to interact. Thus, we can specify multiple interacting OODA loops frameworks. Figure 4 illustrates two conceptual models for interacting OODA loops that cover a large portion of military operational settings. Figure 4a shows the interactions as concurrent processes: humans and machines operating in parallel and influencing each other through communication actions. Figure 4b illustrates integrated OODA loops where the output of one loop becomes the input to the other. This concept models humans and machines coupling together. Each is discussed in more detail below.

4.1 Concurrent Processes

Figure 4a depicts human-machine teaming as concurrent processing, with the machine-OODA and human-OODA processes operating independently and in parallel with sequential communication between the (at least) two agents. Blaha, Jasper, and Cottam (2017) argued that human-machine teams are well characterized as concurrent processes. Concurrent processes are independent processes executing at the same time. Concurrency enables the construction of systems with different compositions of the component processes. Communication is necessary between processes to make systems work. Human-machine teams qualify as concurrent processes because they usually possess the following characteristics: (1) they operate as independent processes, (2) they share at least one common goal, (3) the set of tasks are shared and separately allocated, according to skills and resources, (4) they operate with noisy communication channels, (5) they do not have a shared memory, (6) they do not share processing resources, and (7) they use different processing languages to handle information. With these assumptions, Blaha et al. (2017) presented an approach leveraging process algebras (e.g., Hoare, 1978) to define different human-machine team compositions, such as sequential decision or merging decisions compositions.

For human and machine OODA loops acting as a team through concurrent processes, communication between the two processes is critical for each agent providing input to the other. Communication from the other agent constitutes one of the inputs to the human or machines Observe stage of processing. That is, it behaves as just one of the many potential input streams within the Outside Information observations. As such, its communications will be processed at the time the agent returns to the Observe stage and then applies the orientation process to integrating that communication with the other observations. There could be a delay between the communication being sent and being received, as well as between being received and integrated into the Decide-Act stages. This can facilitate the human and machine operating asynchronously to optimize their individual behaviors, but it can also introduce delays if one is waiting for response from the other. For example, operators and unmanned vehicles sent into contested air space can be separated for some time. In this case, the team may operate as concurrent processes with minimal communication until they can come back together for safe information transmission.

4.2 Integrated OODA Loops

A tight coupling of the human and machine OODA loops is also possible, as illustrated in Figure 4b. In this case, the Output of the Act process of one agent is the input of the Observe process of the other. This concept is consistent with a brief claim by Grant, Venter, and Eloff (2007) suggesting exactly this method of teaming, and it is also consistent with Keus' (2002) framework for dynamic decision making with multi-agent collaborative teams. An integrated OODA loop architecture is key in domains where the human is experiencing the world strictly through machine mediated processes (e.g., managing satellite assets or cyber security operations). Machine-mediated human observations means the machine-Observe process is only receiving inputs for Outside Information and Unfolding Circumstances through the machine. The human does not have direct sensory access to the information separate from the manner by which the machine represents raw data, data analytic outputs, and machine reasoning.

It is advantageous in integrated OODA loops for both the machine-Decide and machine-Act to be the inputs to the human-Observe process. This need has been recognized within the human-automation teaming literature under the call for transparent automation. Transparency, it is argued, is critical for people to have some understanding of what the automation is doing, to improve calibration of trust, and engender appropriate reliance on machine capabilities. More recently, the machine learning and artificial intelligence literatures have seen a push for explainability. Such "explainable AI" should provide a degree of explanation about the machine reasoning processes, similar to transparent automation. Note that in both these cases, as is shown in Figure 4b, the human and machine continue to operate concurrently, with independent Orient-Decide processes. This independence of processes is assumed in both transparent automation and explainable AI systems as well. The Act and Observe stages directly impact each other, though. Consequently, the machine will need to determine or be designed to provide a level of transparency or explanation appropriate for the Observe and Orient needs of the human. This can be achieved more effectively using Human Models integrated into the machine-Orient process in the augmented machine-OODA (Figure 3).

Likewise, the output of human-Act stage is a direct input to the machine OODA through user interface interactions. This relies on system designs leveraging interactions that directly reflect the human decision-action process. The Human Models module in the machine-Orient process again becomes important, because those models can provide rich interpretation of the human inputs based on consistency with the predictions of cognitive models, mathematical characterizations of real-time decision dynamics (e.g., Spivey & Dale, 2006), or semantic interactions (Endert, Fiaux & North, 2012). Any of these formalisms provide an interpretation of the human for the machine in a way (mathematical, computational) that the machine intelligence can integrate into its reasoning and orientation process.

Unlike the concurrent processes, however, if one of the agents is not available to the other, the integrated OODA activity will be terminated. If the system is designed with some degree of flexibility, the agents may be able to switch to behaving as concurrent processes until both are present and interacting again. This might occur, for example, during a shift change. As one operator leaves and another comes on, the machine might continue to run in the background. The drawback to this is the machine will lack any awareness of changes in the environment that are only understood through human input.

4.3 Implications for Being Out-of-the-Loop

Out-of-the-loop errors are the major cause of mistakes and failures of human-machine teams. This is often attributed to a lack of situation awareness on the part of the human operator. Out-of-the-loop errors are thought to be caused by complacency or lack of attention; when people are not paying attention to machine operations, they can lack an ability to re-engage the machine processes with enough time or understanding to effectively mitigate errors (Parasuraman & Riley, 1997). Equal responsibility is not yet given to the machine for failing to communicate or anticipate what the human operator will need. However, if machines are to be treated as equal teammates, such responsibility will be warranted (see, e.g., Geiselman, Johnson & Buck, 2013a; Geiselman, et al., 2013b). From the human-machine teaming OODA conceptual model perspective, we can begin to define and explore multiple ways that both the human *and the machine* could be out of the loop. The classic human out-of-the-loop error might arise because the human reasoned incorrectly about the machine OODA loop. Alternatively, errors might arise in the Observe stage if the human did not wait for appropriate communication between the concurrent processes. The machine could equally fail to receive communication from the concurrent human processes, subsequently executing Orient-Decide-Act processes on incomplete observations. Or, given our augmented machine-OODA concept in Figure 3, we could have failed to provide the machine with adequate models over which to reason about the human. This could result in the artificial intelligence making perfectly logical inferences about the wrong model or wrong human behaviors. It will be an important challenge in the development of human-machine systems to carefully consider how we can design the intelligent processes to be robust against out-of-the-loop errors arising from either the human or the machine teammates.

4.0 CONCLUSION

In this paper, I have used the high-level conceptual framework of OODA loops to capture the decision-action processes of both humans and machines, as well as human-machine teams, in a consistent language and taxonomy. This framework will enable the exploration of different system designs and team configurations through reasoning about the system capabilities in each of the Observe, Orient, Decide, and Act stages. We can reason about the impact that different choices in the modes of interactions have on joint human-machine decisions. We can also reason about cooperative and adversarial agents, human and machine, in this common language. In future work, we could improve this by integrating more explicit process models to capture specific operations and mission environments. We should also seek additional ways to further augment the machine's capabilities to reason about humans, as well as humans' ability to reason about artificial intelligence, and to improve the effectiveness of both agents to serve as effective warfighting team members.

5.0 ACKNOWLEDGMENTS

I thank Rob Jasper and Mark Greaves for early discussions of these concepts, and Corey Fallon for his helpful feedback on an early draft. This research was sponsored by the Analysis in Motion Initiative at the Pacific Northwest National Laboratory. 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 U.S. Government. PNNL-SA-134413

6.0 REFERENCES

[1] Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought*. Psychology Press.

[2] Azuma, R., Daily, M., & Furmanski, C. (2006). A review of time critical decision making models and human cognitive processes. In *2006 IEEE Aerospace Conference*. IEEE.

[3] Bilar, D. (2008, May). Noisy defenses: Subverting malware's OODA loop. In *Proceedings of the 4th Annual Workshop on Cyber Security and Information Intelligence Research: Developing Strategies to Meet the Cyber Security and Information Intelligence Challenges Ahead* (p. 9). ACM.

[4] Blackhurst, J., Gresham, J., & Stone, M. (2012). The quantified warrior: How DoD should lead human performance augmentation. *Armed Forces Journal*.

[5] Blaha, L. M., Jasper, R., & Cottam, J. (2017). Characterizing human-machine teams with process algebras. In *39th Annual Conference of the Cognitive Science Society*, London, England.

[6] Blasch, E. P., Breton, R., & Valin, P. (2011, May). Information fusion measures of effectiveness (MOE) for decision support. In *Signal Processing, Sensor Fusion, and Target Recognition* (Vol. 8050, p. 805011). International Society for Optics and Photonics.

[7] Blasch, E. P., Breton, R., Valin, P., & Bosse, E. (2011, July). User information fusion decision making analysis with the C-OODA model. In *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, (pp. 1-8). IEEE.

[8] Boyd, J. R. (1976). *Destruction and creation*. US Army Command and General Staff College.

[9] Boyd, J. R. (1987). Organic design for command and control. *A discourse on winning and losing*.

[10] Boyd, J. R. (1996). The essence of winning and losing. Unpublished lecture notes, 12(23), 123-125.

[11] Brehmer, B. (2005, June). The dynamic OODA loop: Amalgamating Boyd's OODA loop and the cybernetic approach to command and control. In *Proceedings of the 10th international command and control research technology symposium* (pp. 365-368).

[12] Breton, R., & Rousseau, R. (2005, June). The C-OODA: A cognitive version of the OODA loop to represent C2 activities. In *Proceedings of the 10th International Command and Control Research Technology Symposium*.

[13] Bryant, D. J. (2006). Rethinking OODA: Toward a modern cognitive framework of command decision making. *Military Psychology*, *18*(3), 183.

[14] Curts, R., & Campbell, D. (2001). Avoiding information overload through the understanding of OODA loops. In *Proceedings of the Command & Control Technology Research Symposium*.

[15] Demming, W. E. (1951). The New Way. *Elementary Principles of the Statistical Control of Quality*. Tokyo, Japan: Nippon Kaagaju Gijutsu Remmei.

[16] Dussault, H. M., & Maciag, C. J. (2004). Forensics, fighter pilots and the OODA loop: The role of digital forensics in cyber command and control. In *Digital Forensic Research Workshop*.

[17] Endert, A., Fiaux, P., & North, C. (2012, May). Semantic interaction for visual text analytics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 473-482). ACM.

[18] Endsley, M. R. (2015). Autonomous horizons: System autonomy in the Air Force: A path to the future. *Washington, DC: US Air Force Office of the Chief Scientist*.

[19] Fadok, D. S. (1995). *John Boyd and John Warden: Air Power's Quest for Strategic Paralysis*. Air University, Maxwell AFB AL, School Of Advanced Airpower Studies.

[20] Galster, S. M., & Johnson, E. M. (2013). *Sense-assess-augment: A taxonomy for human effectiveness* (No. AFRL-RH-WP-TM-2013-0002). Air Force Research Lab Wright-Patterson AFB OH Human Effectiveness Directorate.

[21] Geiselman, E. E., Johnson, C. M., & Buck, D. R. (2013a). Flight deck automation: Invaluable collaborator or insidious enabler?. *Ergonomics in Design*, *21*(3), 22-26.

[22] Geiselman, E. E., Johnson, C. M., Buck, D. R., & Patrick, T. (2013b). Flight deck automation: A call for context-aware logic to improve safety. *Ergonomics in Design*, *21*(4), 13-18.

[23] Giordano, J., & Maciag, C. (2002). Cyber forensics: A military operations perspective. *International Journal of Digital Evidence*, *1*(2), 1-13.

[24] Grant, T. (2005, July). Unifying planning and control using an OODA-based architecture. In *Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries* (pp. 159-170). South African Institute for Computer Scientists and Information Technologists.

[25] Grant, T., & Kooter, B. (2005). Comparing OODA & other models as operational view C2 architecture. *ICCRTS2005.*

[26] Grant, T. J., Venter, H. S., & Eloff, J. H. (2007, October). Simulating adversarial interactions between intruders and system administrators using OODA-RR. In *Proceedings of the 2007 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries* (pp. 46-55). ACM.

[27] Hoare, C. A. R. (1978). Communicating sequential processes. *Communications of the ACM 21(8),* 666-677.

[28] Keus, H. E. (2002). *A framework for analysis of decision processes in teams*. TNO Physics And Electronics Lab Hague (Netherlands).

[29] Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, N.J.: Prentice Hall.

[30] Olsson, F. (2006). Intrusion Management. Masters thesis, report 06117, School of Mathematics and Systems Engineering, Växjö University, Sweden.

[31] Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253.

[32] Révay, M., & Líška, M. (2017, October). OODA loop in command & control systems. In *Communication and Information Technologies (KIT), 2017* (pp. 1-4). IEEE.

[33] Rousseau, R., & Breton, R. (2004, June). The M-OODA: A model incorporating control functions and teamwork in the OODA loop. In *Proceedings of the 2004 Command and Control Research Technology Symposium* (pp. 14-16).

[34] Shewhart, W. A., & Deming, W. E. (1939). *Statistical method from the viewpoint of quality control*. Courier Corporation.

[35] Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, *15*(5), 207-211.

[36] Wohl, J. G. (1981). Force management decision requirements for air force tactical command and control. *IEEE Transactions on Systems, Man, and Cybernetics*, *11*(9), 618-639.

[37] Zager, R., & Zager, J. (2017, Oct. 21). OODA loops in cyberspace: A new cyber-defense model. *Small Wars Journal.* Retrieved from http://smallwarsjournal.com/jrnl/art/ooda-loops-cyberspace-new-cyber-defense-model