

Con conversationally Adept Virtual Humans Technologies: Autonomous Combat Team Members’ “Incarnation”

Mark C. Davis
Wood Duck Research, Inc.
Mooresville, North Carolina
mark@woodduckresearch.org

Ke-Thia Yao & Dan M. Davis
Univ. of Southern California
Marina del Rey, California
kyao@isi.edu & dmdavis@acm.org

David H. Barnhill
Naval Postgraduate School
Monterey, California
david.barnill@nps.edu

ABSTRACT

This paper addresses the technological opportunities and organizational needs for the Modeling and Simulation (M&S) community to facilitate the extension of emerging M&S capabilities into non-human warfighters. The authors posit that these non-humans might take various forms, the nature of which is left to the passage of time to reveal. The various existing forms of autonomous devices and the degree of their autonomy are surveyed. The next section briefly reviews the rapid and accelerating progress in virtual human behavior and adduces evaluative comments from users in the authors’ research projects into virtual humans. Based on their operational and research experience, the future of such technology trends is extrapolated. A scenario is advanced in which various non-human operational entities would be infused with doctrinal guidance and controlled locally by a human with back-up. The man/machine interfaces are suggested and analyzed in light of current on-going research in which some of the authors are participating. The benefits of using non-humans for dangerous tasks is considered, as are the ethical and societal risks of loosing machines with lethal weapons into the confusion and tumult of the modern, often urban, warfare. The beneficial interface impacts of making the non-human entities respond to human control and appear human, via augmented reality, to human controllers would be major human factors research goals, which the paper will outline, and pursue in the abstract. They will discuss the avenues of research in Artificial Intelligence that will surely be applicable and raise issues of checks on learned behaviors becoming violative of civilized standards or US humanitarian mandates. These issues are considered in a way that should be extensible into other areas with similarly conundrumal issues, *e.g.* valuable unique critical benefits, but significant risks. The paper closes with an articulation of the needed community support and contributions.

ABOUT THE AUTHORS

Mark C. Davis, Ph.D. is the Chief Technical Officer at Wood Duck Research, Inc, and is semi retired after careers in the US Navy and as a computer design engineer for both IBM and Lenovo. Rising to the level of Distinguished Engineer at Lenovo, he was responsible for the design of laptop computer cross-disciplinary technology, including PC architecture, embedded systems, open source and virtualization. Previous work was with IBM in the areas of software development and architecture involving security, storage and virtualization. Dr. Davis has been granted well over fifty patents that were filed during his service at both companies. He is a graduate of the Duke University NROTC program and was commissioned as an Ensign, attended nuclear power school, and served as a Submarine Officer for twelve years, including one duty tour as a classroom instructor. He left the active duty as a Lieutenant Commander to pursue a PhD. Mark holds a BSEE degree from Duke University and a PhD in Computer Science from the University of North Carolina, Chapel Hill, where his advisor was Professor Fredrick P. Books.

Ke-Thia Yao, PhD, is a Project Leader and Research Scientist in the Computational Systems and Technology Division of the University of Southern California (USC) Information Sciences Institute (ISI). His primary research interest is in understanding large complex systems and data sets. He teaches data management classes at USC’s Viterbi School of Engineering. He is part of the team investigating uses for USC’s 2,000 Qubit Quantum Annealer. Dr. Yao was one of the leads on the JESPP project where he analyzed the computational and behavioral properties of large-scale battlefield simulations. That project had the goal of supporting very large-scale distributed military simulations involving millions of virtual and constructive entities. He received his B.S. degree in EECS from the University of California Berkeley, and his M.S. and Ph.D. degrees in Computer Science from Rutgers University.

Dan M. Davis, CDR, USN-Ret, is active as a consultant at the Institute for Creative Technologies, University of Southern California (USC), focusing on large-scale DoD simulations and avatar implementations. Prior to retirement, he was the Director of the JESPP project at USC for a decade. As the Assistant Director of Advanced Computing Research at Caltech, he ran Synthetic Forces Express, bringing HPC to DoD simulations. He also served as a Director at the Maui High Performance Computing Center and in computer research roles at the Jet Propulsion

Laboratory and Martin Marietta. He was the Chairman of the Coalition of Academic Supercomputing Centers and has taught at the undergraduate and graduate levels. As early as 1971, Dan was writing programs in FORTRAN on one of Seymour Cray's CDC 6500's. While in the Marine Corps, he saw duty in Vietnam as a Cryptologist and retired \ as a Commander, U.S.N. He received B.A. and J.D. degrees from the University of Colorado in Boulder.

David H. Barnhill, LCDR, USN, is enrolled the US Naval Postgraduate School (NPS), in Monterey California. He is a student there in the Operations Research Department. He is particularly interested in the analysis of human behavior and command relationships of defense personnel. A topic of immediate concern is the imminent adoption of various levels of robotic and artificial intelligence-controlled weapons on unit cohesion and command functions. He is a Naval Aviator, and has flown rotary-wing aircraft from both land and vessel platforms. David has served in flying status leadership positions and has performed staff officer duties. The Navy is currently considering him for candidacy for a PhD and appointment as a Permanent Military Professor. He graduated from the US Naval Academy with a BS and is scheduled to complete an MS in Operations Research at NPS in the spring of 2021.

Mark C. Davis
Wood Duck Research, Inc.
Mooreville, North Carolina
mark@woodduckresearch.org

Ke-Thia Yao & Dan M. Davis
Univ. of Southern California
Marina del Rey, California
kyao@isi.edu & dmdavis@acm.org

David H. Barnhill
Naval Postgraduate School
Monterey, California
david.barnill@nps.edu

INTRODUCTION

Combat is by definition exceedingly difficult because it pits humans against human with ultimate stakes. Identifying and deploying technologies that can reduce the costs and improve the outcomes of combat have been a centuries-long quest from the improvement in metal weapons, through the introduction of the Gatling gun and on to the Manhattan Project (Keegan, 2011, Marcot 1862,). Advancements in technology have enabled new weapons systems with more capability. In particular various recent forms of artificial intelligence have enabled autonomous devices to perform existing and emerging combat roles. This technology has brought its own problems: notably information overflow and efficient fail-safe control. The solutions advanced below are offered as a way to improve the interface between a trusted human and multiple efficient mechanical henchmen. The authors offer both their research and operational experience involving the use of autonomous entities, followed by enumerating issues of control and ethics. This is followed by an outline of the paper's vision and current research that supports achievement of that vision. The paper concludes with a discussion of the implementation envisioned and the challenges that must be addressed.

BACKGROUND

Technology for autonomous warfighting entities is emerging. The art has advanced from early enhanced weapons to self-driving cars and remote piloted aircraft such as the MQ-1 Predator and MQ-9 Reaper. These Remotely Piloted Aircraft Systems (RPAS) are not passive intelligence gathering capabilities, but at manifestly lethal weapons with on-board capabilities to identify and actively target both facilities and human beings. (NY Times, 2020). Figure 1 shows a MQ-9 Reaper launching a surface to ground missile, an AGM-114 Hellfire. That missile can carry 25 pounds of explosive and is effective against personnel, armored vehicles and bunkers. It can be guided remotely or by local targeting using lasers.



Figure 1. MQ-9 Reaper (Military-Today, 2020)

Using this technology in combat situations faces many major obstacles. Two are important to the thesis of this paper. The first is that we, as a society, do not want to cede life-and-death control to non-human entities (Bieri, 2014). The second is that the information flows from the autonomous entities have now dramatically increased (Defense Systems, 2020).

Historical Progression of Combat Technology

Development and deployment of autonomous weapons has proceeded since WW II. The introduction of radar detonated anti-aircraft shells (Brennan, 1969) and self-guided torpedoes (Jolie, 1978) progressed to more sophisticated fire and forget weapons. These developments were straight forward because the target was designated before the weapon was fired. All of the ethical decision concerning correctness and rightness of the target were made before weapon launch. In this case information exchange between launcher and weapon was limited to minor adjustments in position. Even if information flow failed, the consequence was rarely more than the target was not destroyed.

Technology has progressed to enable more capable autonomous vehicles. Much of the work required for autonomous automobiles such as image recognition, route strategy and emergency response is directly applicable to combat machines. Current autonomous automobile technology has progressed to solve almost every driving

environment. Shipped products have the capability to drive from coast to coast, find a parking space and then pick up the driver at a nearby location. *Robotaxi* demonstrations are already being conducted in some locations <Krumwiede, K & Liou R., 2019>. The remaining problems in autonomous driving are under development.

Current Remote Control Experience

One the major differences in operating unmanned vehicles is the simple fact that the operator is not ‘present’ with the vehicle in its ‘real-time’ location. This illuminates a number of challenges that are perhaps manifest in manned operations but are now largely ancillary. (Callahan, 2014) Using manned aircraft as an example, the pilot is the ultimate arbiter or what they decide to act upon. The use of situational awareness to make informed decisions is a key tenet in the operation of any aircraft. However, in operating unmanned aircraft, at best a pilot’s situational awareness is muted, at worst it is significantly degraded (Figure 2). Throughout flight school, student pilots are told to rely on their instruments. They must not rely solely on their kinesthetic sense because a pilot’s body and mind are easily misled. However, the sensory indications that pilots feel as they fly aircraft are essential to build situational awareness in order to evaluate what the aircraft is “telling” them. A short example: a caution light illuminates; the pilot sees it and immediately references gauges associated with the light. Once the gauges are evaluated, the pilot may evaluate the feel of the aircraft (*i.e.* vibrations, sounds, smells). Ultimately, a decision is made. The senses help build the extra information necessary to make informed decisions. All of this happens, usually, in a matter of seconds.



Figure 2. US Reservists staff a Drone Control Console (US Air National Guard Photo)

The aforementioned example is largely absent in unmanned operations. The operator will get a visual indication on a problem, but the sensory indicators are absent. Evaluation is hampered due to the sensory deprivation and oftentimes, the only option is to direct the vehicle or drone to return to its base of operations. There is little troubleshooting that an operator can do to fix a remote problematic situation.

This is further exacerbated by the filter of technology. In most cases, an unmanned operator is operating two systems in series. They sit at a console or some sort of control station that takes inputs and then feeds the information to the vehicle. Likewise, the flow returns. While this is necessary, this ‘filter’ means that there are two separate pieces of equipment that can have problems and, again, detract from the operator’s ability to properly and efficiently evaluate situations whether it be mission or aircraft-related. This “filtering” idea is further compounded by the communications latencies associated with unmanned operations. Again, using aviation as an example, if a pilot moves his hands to manipulate aircraft controls, the aircraft responds almost instantaneously. By contrast, the operator at a control station makes a control input, the control station sends the input via communication link and the unmanned system responds. The idea of “real time” action is not yet possible in the unmanned or autonomous systems.

This leads inevitably to further discussion on what unmanned aircraft should be “allowed” to do. In a military context, decisions using drones are not taken at face-value. Drone usage has often been confined to environments that allow them to be there. There is no threat to their being destroyed or shot down. To put it simply, there is no appreciable threat to their destruction. However, implementation of drones in a real-time combat situation will require careful evaluation of the Law of Armed Conflict (LOAC) (Green, 2018) and associated governing directives. If a force is not actively putting its own personnel at risk (*i.e.* death, maiming), are they adhering to the tenets of LOAC if the other force is? This idea of mutual risk has been the major tenet of LOAC and made it acceptable. If we risk nothing but machinery, is it still armed conflict? Or is it something else? Also, the issue of the emotional impact on the operator may be different than that of the live pilot in a cockpit.

On a less philosophical level, consideration must be given to how much do humans trust unmanned assets to help them make time-critical informed decisions. Regarding the idea of latencies, the information received via drone (in the current construct) will always be less timely than if a human passes that information (Hoppe *et al.*, 2019). Part

of this distrust is resultant from the fact that a decision maker cannot easily ask further questions of a drone like they might of a human. Basically, the limited sensor data one sees from a drone is often constrained. Analysis is left up to those at the base who may be significantly detached emotionally, for good or for ill. A more human interface might make these issues more amenable to current training, skills, and practice. In the time it took for the information to get from the drone to the operator to the staff to the decision maker, things may have changed. Having a virtual human as the communications interface might reduce human reaction latencies enough to counter communication latencies. If this were to be possible with humans-in-the-loop, there might be faster updates and more clarification.

In order for unmanned systems to execute more than rudimentary functional military tasks, the interaction between operator and remote vehicle must resemble that of a typical human experience. Interfaces must allow for greater situational awareness that allows the operator to “feel” the drone as it operates (Knierim et al., 2017). While technology is a force multiplier, its complexity can prevent the operator from effectively and efficiently using their vehicle. The filter, while likely ever-present, must be reduced in scope such that time-lag is as small as possible. Information passage must be of high fidelity such that real-time analysis is possible and answers inevitable questions of clarification. This list of necessary changes is not exhaustive and simply represents a starting place for better human-machine interactions. To make progress this is where one must begin.

Ethical Control of Combat

To maintain the ethical control required by society, the autonomous entity capability must be limited, and a human supervisor must control the lethal actions of the entity. For a remotely piloted aircraft, the interface can parallel an existing aircraft and few additional problems are introduced. But a combat infantry drone is another matter. Although the nature of the controlling interface for such a drone is not clear, the dramatic increase in information flow is clear and will certainly be a problem.

Information Overflow

Information flow and specifically too much information has been a major and well-studied problem. It is widely acknowledged that information overflow can be deadly to warfighters (Shankar & Richtel, 2011)>, <,(Bateman, 1998). In high technology areas such as piloting fighter aircraft, getting the right information has been a problem under study for many years (Hettinger & Haas, 2000)

VISION OF THE CONCEPT

Throughout history, militaries have hone command and control to assure that combat meets the strategic and ethical objectives of organizational leadership as described centuries ago by Sun Zu (Griffiths, 1963)>. Rather than create a totally new approach to warfare, a logical development is to integrate the autonomous entity into existing military command relationships with minimal changes. By limiting the ethical decision making to well-trained, trusted warriors in control. The relationship between this trusted human and the autonomous entity would then resemble the relationship between a Corporal and Private.

Arguably, an autonomous private should not have all of the ethical responsibilities of a human private. This change in role is not a dramatic and the ethical requirements of a fighter pilot so the modification in role should be achievable. Similarly, in many combat teams’ minor revisions to duties, responsibilities and ethics of the roles would enable autonomous entities to server in that capacity.

One advantage of this approach is that it leverages all of the existing military training and thought. All levels of command would continue to use the commands and concepts they have already learned to use this new technology. The impact to training organizations and operations would thus be minimized.

The vision is to place autonomous entities into the existing military hierarchy with minimal changes. In particular, the command and control language and semantics should be as close to existing practice to take advantage of the thousands of years of experience represented.

Elements of the Concept

The elements of this concept include

1. A mechanical device capable of physical tasks necessary for a combat mission
2. An artificial intelligence to operate the device and support the interface with the command structure
3. A command structure utilizing the function of the device including humans designated to control the device
4. An optimal user interface to efficiently use the abilities of the device and the human controller

This paper will not address the technologies of the mechanical devices. The existing technologies discussed above will logically be extended to provide the capabilities required for autonomous warfighting devices. The best command structure is the existing command structure that has evolved over centuries that has already adapted to changes in technology and human society. The interface for this task may be Virtual Humans. This technology is introduced below. It should be noted that advancements in artificial intelligence will enhance both the interface and the capability of the autonomous entities. Quantum computing is one of these technologies and is discussed below in relation to Virtual Humans.

Current Advances and Emerging Technologies in Virtual Humans

One accepted, but emerging, technology is that of Virtual Humans. As that term is used in this paper, a wide range of virtuality is accepted, including using computer-selected video clips of a live human. It a creation in virtual reality portrayed by an avatar which attempts to recreate the appearance, voice, feel, and interaction that a live human would produce. Using the advances in several new technologies, including but not limited to natural language processing (NLP), virtual reality (VR), computer-generated imagery (CGI), machine learning, and virtual learning, live teachers can be presented. The uses, as well as the limits of these tools are becoming evident. Researchers at ICT have developed programs that have been shown to be effective, *e.g.* SimCoach, New Dimensions in Testimony (NDT), PAL3, and others (ICT, 2019). These are generalized under learning sciences, medical VR, mixed reality, narrative, social stimulation, virtual humans, and vision and graphics. The field of knowledge available here is immense, and countless resources and expertise are now becoming available in this promising field. The user interface can take many forms, as in Figures **Figure 3**, **Figure 4** & **Figure 5** below.



Figure 3, SimCoach uses CGI to advise PTSD patients



Figure 4. A 3D holographic display makes a Holocaust Survivor “real”

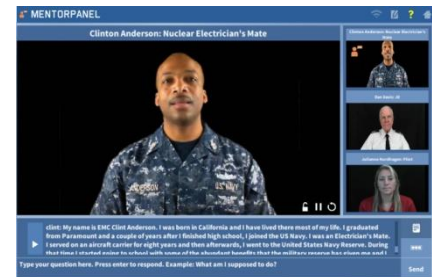


Figure 5.– Mentor Pal uses recorded video clips to help STEM Student

Natural language processing composes “an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things” (Chowdhury, 2003). Using this definition within the context of virtual environments, NLP tools allow computer technology to recognize voice input, analyze voice tone, provide lifelike conversation, retrieve information, and many other applications in combination with machine learning. Recent developments in NLP have made amazing bounds like “a single convolutional neural network architecture that, given a sentence, outputs a host of language processing predictions: part-of-speech tags, chunks, named entity tags, semantic roles, semantically similar words and the likelihood that the sentence makes sense (grammatically and semantically) using a language model” (Collobert & Weston, 2008).

The MentorPal project at ICT was designed to produce a virtual conversational computer agent to high school students select a career. In order for the project to gain some insight into the “conversationality” of the mentors, they made the product available in three settings: an engineering fair on campus, a class room setting at a private tutorial organization, and in a computer lab at the Naval Postgraduate School’s summer program for high school students. For students using the system individually, they were content to ask questions from MentorPal for fairly long periods of time, approaching fifteen minutes. However, it should be noted that these students were not randomly selected and the data does not presume to be probative of a formal thesis.

As the intent of the project was to create a conversational easiness with the student that would encourage the mentor-mentee relationship, one of the areas in which the team was most interested was the students' casual impression of their dialogue with the mentor. Since this effort was to help guide the programmers, rather than as a research study, no formal survey was done of all of their views. However, to help benefit from suggestions about how to improve the system and collect visitor reactions, a straw poll was offered to students and adults who spent significant time with the mentors. The data from this straw poll are summarized in Table 1. A smaller preliminary subset of this data was reported in an earlier paper by the authors (Beck *et al.*, 2018). The most dramatic factor was the high level of "conversationality" perceived by the students and the few adults who also opted to assess it.

Young Mentee Responses from Three Trial Cases*					
N = 91	Gender	Male: 55	Female: 36	Age:	Mean = 17.3 Median = 16.1
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think or worry about my future Life almost every day	10	15	43	15	8
Having easy access to better career advice would be useful	0	1	34	37	19
I liked the on-line mentors and learned from them	0	0	15	53	23
The mentors' answers were on-target and useful	0	0	37	33	21
The responses seemed "conversational" and real	0	0	20	32	39

*plus half a dozen adults; their ages were estimated

Table 1. Likert-style Straw Poll on "Conversationality"

Applications of interest in cross reality, as a synthesis of several vitalities and reality, can use developments in the interpretation of language which have been recently enhanced by the many researchers focusing their studies on the field. At ICT alone, progress has been made in training/learning environments (Kenny *et al.*, 2007), multi-party dialogue (Traum & Rickel, 2002), ethics and cooperation (Allwood & *et al.*, 2000), health applications (Rizzo *et al.*, 2011), and representation and reasoning (Swartout *et al.*, 2006). Although automated speech recognition (ASR) is far from perfect, some of the best software, "Google achieved 73.3% of exact recognized phrases with a 15.8% [Word Error Rate]" (Kudryavtsev *et al.*, 2008), and the technology will continue to improve. It is also important to note that many of the errors in ASR are caused by slurred speech, cultural slang, and context, stemming from a "lack of consistent units of speech that are trainable and relatively insensitive to context" (Lee, 1988). Although this causes problems when comparing these transcriptions with global data, models can be trained locally using software such as CMU Sphinx from Carnegie Mellon University as applied to languages such as Arabic (Satori, *et al.*, 2011). All of these will continue to improve with time, as hardware, software, and data storage and availability are areas within which research is eminent and directly valuable. Specifically, the speed and application of quantum computing will engender significant advancement in NLP and its applications: one of which is the impactful area of virtual humans.

Effective Conversational Interfaces

The SimCoach system (see Figure 3 above) experience aimed to "motivate users to take the first step – to empower themselves to seek advice and information regarding their healthcare". These virtual systems have been shown to have more success, generating deeper levels of confidence with patients than even live healthcare interactions (Rizzo, *et al.*, 2011). The successes of these operations so far foreshadow abundant uses in the near future.

A closer analog may be found in the area of storytelling and gaming. The New Dimensions in Testimony project "allows people to have an interactive conversation with a human storyteller (a Holocaust survivor) who has recorded a number of dialogue contributions (Figure 6), including many compelling narratives of his experiences and thoughts" (Traum *et al.*, 2015). The project's important mission of preserving the stories of Holocaust survivors can be applied to endless amounts of other important persons and tales. Advances in gaming allow for utilization in learning, entertainment, healthcare, and lifelike training for things like combat and critical thinking. Critical thinking training in the English Language Arts area will require the use of virtual environments that make the most skilled instructors and exciting resources available. Some of the current negatives of virtual learning are the difficulties of personalizing learning, handling live question, accommodating different learning styles, and similar issues which are often mirrored in the classroom instructors' lack of time for thirty individuals in a class.



Figure 6. Subject receiving direction in light stage used for generating 3-D imagery for holographic display

Future fields of focus will come in increased sophistication in such virtual environments and interaction. For example, imagine a system which takes in a live recorded interview with a president, or other prominent or knowledgeable figure, and generates an interactive environment in which a user can experience the interview by asking the questions themselves. Such things are already in existence, and they will only become more prominent and powerful. However, the need for extensive processing power, efficiency, and data storage and transfer remains a limiting factor for such developments. We believe that the aforementioned applications call for considerable further exploration and research.

Computational Advances will Enhance the Experience.

Many of the looming hurdles to advancing these new technologies seem to be beyond the reach of current technologies, both in hardware capabilities and software sophistication. (Yao *et al.*, 2018). The advent of promising emerging technologies and approaches is worthy of consideration here. To have defense utility, training a “virtual human” should be orders of magnitude faster than that necessary to develop and live human. The major impediment here is the long training times for so complex a behavior. The amount of material to be input and then recursively looped through a training algorithm presents a daunting obstacle.

Deep learning, which uses layers of computational units to learn data representations at multiple levels of abstraction, has emerged as the leading methodology for developing NLP applications (Young *et al.*, 2018, Pouyanfar *et al.*, 2018). These networks are called Convolutional Neural Networks (CNN). Word vector representation or word embeddings learned using neural networks capture the meaning of words by embedding similar words closer to each other in vector space (Mikolov, 2013). Distance and relative direction of words can capture semantic meaning. The word for *queen* can be approximated using vector arithmetic: *queen* = *king* – *man* + *women*. These word embeddings form the basis for input to deep CNN networks for a variety of NLP applications, including topic classification and sentiment analysis. In question answering (QA) applications CNN are used to select semantically similar answers from an existing knowledge base.

CNN processes sequential data, such as sentences, by using windowing, but the window size is fixed at training time. Word context outside the window is ignored. Recurrent neural networks (RNNs) are designed to process sequence data. RNN adds feedback loops into the neural network, which enables it to retain state information and process variable length of input sentences. A popular recurrent network is long short-term memory (LSTM), which introduced the notion of gates (input, output and forget gates) to regulate the flow of information (Hochreiter, 1997). This ability to remember long distance information enables LSTM to perform well in applications like natural language translation and dialogue systems.

Stochastic neural networks, or Boltzmann machines, are a type of neural network that introduces randomness into the computation units. Instead of neurons firing at fixed input thresholds, Boltzmann machines fire stochastically based on the input.

These neural networks are sometimes referred to as generative networks because they learn the probability distribution of the input training data and they are able to stochastically regenerate the data. Popular generative networks, include the autoencoder, deep belief network (DBN), deep Boltzmann machine (DBM) and generative adversarial network (GAN) (Goodfellow *et al.*, 2014). GAN pairs a deep generative network with regular deep discriminative network. These two networks are evolved simultaneously, where the generative network creates artifacts and the discriminative network critiques the artifacts. GANs are capable of generating photo realistic faces (Karras *et al.*, 2019) as well as novel paintings (Elgammal *et al.*, 2017).

Boltzmann machines allows for arbitrary network topology, but most generative networks forbid connection within a network layer, *i.e.* using restricted Boltzmann machines (RBM) as layers. The reason for the restriction is to support computational efficiency. Intralayer edges introduce loops in the network which dramatically increase the training time in an undesirable manner.

At the University of southern California’s Information Sciences Institute (USC’s ISI) the quantum computing center is in possession of a D-Wave open-system adiabatic quantum annealer that is capable of sampling from Boltzmann machine network with loops, specifically chimera graphs. Although the current generation has limited number of qubits and intralayer connection, we have shown that the extra representational power afforded by these extra edges can decrease the training time and improve learning (Yao *et al.*, 2018, Liu *et al.*, 2020). This advance should enable more realistic and “human” virtual computer agents.

One of the issues facing both the previous work on MentorPal and the projected future work on the incarnation of the robotic warfighter is the identification, assessment and ranking of various scenarios and environments. The prior experience of the selection of issues to be addressed will be critical as we face the problems of effectively controlling the virtual battlefield superimposed on the real battlefield. In previous work, the authors had to ascertain the corpus of questions to be asked. They developed useful techniques for ascertaining this corpus.

These advances are significantly interesting and have attracted inquiries for various agencies of the government (**Error! Reference source not found.**), who also often have questions about the size of current quantum machines (Figure 8).



**Figure 7. Dr.s Yao and Lucas briefing
Asst. Dir Biometric identity Management, DHS**



**Figure 8. Demonstrating the scale of
Quantum Annealer Case**

The utility of the Incarnation suggested herein would have dual use attractions to those outside of the Department of Defense, and wide promulgation of research efforts would be advisable..

New display and sensor technologies are expanding rapidly. The holographic displays in use by the New Dimensions in Testimony project cited above might find significant applicable utility in the provision of a “virtual humanizing” of battlefield robotic entities. The pervasive use of virtual, augmented, mixed, and cinematic reality is the core of the definition of cross reality, also called XR. This paper’s major thesis not only will require, but may be the paradigm case for the emergence of this new discipline. This cross-disciplinary approach is central to the success of the proposed effort to support the warfighter in ways that will be effective in high-stress situations. Such a development strategy will also assist in preventing untoward actions by a rogue autonomous entity that would violate the Laws of Armed Conflict.

Extensibility

It is the authors experience that not only are there an unfathomable plethora of military situations, technologies (Burmaogla & Saritas, 2017), and organizational hierarchies, but these are becoming more dynamic, more complex and more transitory (Kott & Perconti, 2018.) In response to this ostensibly inevitable flow of new technologies over the course of a career, it seems most important to focus analyses on successful approaches more than specific platforms and units. The insights from a system to facilitate a Squad Leader’s optimal control over a mixed unit made up of human and non-human entities may illuminate how to best approach the control of a UAV pilot over a flight of UAS aircraft. Interoperability may not be universally feasible, but experience dictates a cautionary admonition not to become to insular and remain open to other services’ advances in the human/non-human interfaces (Gong *et al.*, 2001).

Another dimension of extensibility is the dual-use interest held by DoD research organizations. Clearly the issues discussed in this paper would be immediately extensible into civilian first responder contexts, but that may only be the beginning. The authors hold that providing a more human-like interface, critical in high stress situations, may have similar benefits in other important functions. One of these functions could be that of counseling, such as demonstrated in the SimCoach project (Rizzo *et al.*, 2016), another may be that of assessment (Shaw *et al.*, 2019) and yet another may be that of instructional interfaces, (Elstad & Davis, 2017). Additional implementations will no doubt occur to the reader. This paper asserts that it is incumbent upon the researchers to abstract, identify, define and communicate their insights and approaches in this field, eschewing the proclivities to focus solely on the task before them.

Inculcating Appropriate Behavioral and Communication Styles

One area of potential is the area of computer agent personality, whether the agent is a pure virtual human, e.g. an animated agent, or a live human, e.g. the use of a library of video-clips as in the MentorPal project (Burns *et al.*, 2018). The authors have, among them, more than half a century experience in leading personnel in dangerous military situations. They know that a commander's primary concern is knowing their people, understanding their needs and providing the guidance they require. Some need their confidence bolstered, some are foolhardy and require constraints, and some need constant "motivation" in order to achieve mission goals. It seems clear that the passionless voice of HAL in the film 2001 (Kubrick, 1968) which was characterized by a lack of affect, would not be compelling or optimal in a highly stressed situation such as combat. However, it might be beneficial to modify both the computer agent's basic personality and their emotional tone, as may be warranted by the situation. This would hold for situations calling for a human/non-human interface both up and down the Chain of Command. This paper's vision was the issues arising when a live human is controlling computer-driven entities, but it is within the authors' *Weltanschauung* that there may be times when an intermediary computer agent will be relaying direction from a senior commander to the warfighter in the field (Khalid *et al.*, 2018).

The application to combat would require some general customization of the personality of the devices based on the roles of the device. Study will be required to determine customizing personalities for individual devices is useful.

Scenarios

We showed in an earlier paper (Davis, et al., 2018) that this interface could assist information flow. Some examples of this concept are:

1. Privates in an infantry unit could be replaced by autonomous privates.
2. Various functional teams on a warship could be replaced by entities.
3. A flight of autonomous attack aircraft could be controlled by a single human flight commander.

For the example of an infantry squad, the human leader would take the role of Corporal leading his privates. The squad would stay in close proximity autonomously, but tactics and use of lethal fire would be employed only at the verbal command of the squad leader. The autonomous privates would communicate orally and via the use of computer enabled avatars. along with real time data channels such as video and audio.

An entire submarine could be run by a single Captain, with AI enabled computer agents operating the propulsion, weapons, sensors and fire control systems. They would communicate with the using Natural Language Processing capabilities shown effective in virtual conversational capabilities cited above. The Captain's communications interfaces would be similar to those functions experienced with typical human crews.

A flight of autonomous aircraft could be controlled by a single human flight commander flying in concert with the swarm of drones. An example of the type of drone and the possible tactics are described in (Levik, 2019). The issue we raise here is that of the best way to make his control a similar to the way it is now. A couple of potential paths of implementation would be based on the technologies propounded above, based on current successes and emerging capabilities in NLP. One issue to be addressed would be that of verbal garbles. Like so many things that are ostensibly easy for humans, understanding human speech is often a matter of knowing what to expect, when to "fill in the gaps" and when to ask for a clarifying repeat. Lethality, ethics and the Law of Armed Conflict mandate a set of redundant fail safes to prevent, or at least minimize, catastrophic failures. Carefully crafted and explicitly delimited specifications would be required and the authors' experience is having a full-time team member who has actual combat operations experience is invaluable. Seeing a Subject Matter Expert once a month is no substitute. Such a capability is envisioned in Figure 9 by the Boeing Corporation.



Figure 9. Artist's Rendition of a Mixed Drone and Manned Aircraft Flight (Boeing Photo)

CONCLUSIONS

Using more technology to fight will have advantages in efficiency and safety for the warriors. As more autonomous weapons are introduced, the ethical challenges will emphasize the criticality of the interface between the human commanders and the incarnation of the autonomous warfighting entities. The authors think that the best interface is the centuries' proven human to human interaction of natural language. The interactions of the classic chain of command have been developed over the millennia to optimize the sensing of situations, personal capabilities, fatigue and morale, these translated into effective control and assessment of appropriate delegation of responsibility. AI can assist in these functions if appropriate emerging technologies are enhanced and adopted effectively. The technologies exists to support this type of interfaces utilizing virtual humans and exploiting emerging technologies such as quantum computing, Natural Language Processing, Deep Learning, advanced sensors and sophisticated display and haptics.

Further research studying these scenarios in detail will lead to better solutions.

ACKNOWLEDGEMENTS

The authors would like to thank the work done by the ICT team on virtual humans showing the applicability of this technology to a wide variety of problems. Most especially the Co-Principal Investigators of the project the Doctors Benjamin Nye and William Swartout.

REFERENCES

- Allen, I. E., & Seaman, C. A. (2007). Likert Scales and Data Analyses. *Quality Progress*, 40(7), 64-65.
- Allwood, J., Traum, D., & Jokinen, K. (2000). Cooperation, Dialogue and Ethics. *International Journal of Human-Computer Studies*, 53(6), 871-914.
- Artstein, R., Gainer, A., Georgila, K., Leuski, A., Shapiro, A., & Traum, D. (2016). New dimensions in testimony demonstration. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations* (pp. 32-36).
- Bateman III, R. L. (1998). Avoiding information overload. *Military Review*, 78(4), 53.
- Beck, S., Carr, K., Davis, D. M., Nordhagen, J. N., and Nye, B. D., (2018). Virtual Mentors in a Real STEM Fair: Experiences, Challenges, and Opportunities .In *Third International Workshop on Intelligent Mentoring Systems (IMS 2018) Proceedings*.
- Bieri, M., & Dickow, M. (2014). Lethal Autonomous Weapons Systems: Future Challenges. *CSS Analyses in Security Policy*, 164.
- Brennan, James W. (September 1968), The proximity fuse whose brainchild?, *United States Naval Institute Proceedings*
- Burmaoglu, S., & Saritas, O. (2017). Changing characteristics of warfare and the future of Military R&D. *Technological Forecasting and Social Change*, 116, 151-161.
- Burns, D.P., Davis, D.M., & Nordhagen, J. N. (2018). Systems Engineering: Optimizing Creation of Virtual Conversational Human. In the *Proceedings of the ModSim World Conference*. Norfolk, Virginia.
- Callahan, A. L. (2014). Reinventing The Drone, Reinventing The Navy. *Naval War College Review*, 67(3), 98-122.
- Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology*, 37(1), 51-89.
- Collobert, R., & Weston, J. (2008, July). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th International Conference on Machine Learning*

- Craig, S. D., Chiou, E. K., & Schroeder, N. L. (2019, November). The Impact of Virtual Human Voice on Learner Trust. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 63, No. 1, pp. 2272-2276). Sage CA: Los Angeles, CA: SAGE Publications.
- Davis, D. M., Predovich, K.B., Stassi, F.J., Spaulding, H., Shaw, K & Nye, B.D. (2018). Enhancing Menteeship: Improving Career Selection for Potential DoD Personnel. In the *Proceedings of the SISO Fall Simulation Innovation Workshop*, Orlando, Florida: SISO
- Davis, M.C., Stassi, F.J. & Davis, D.M. (2018). "Innovative Education Technologies: Optimizing Future Information Comprehension Capabilities". In the *Proceedings of the SISO Fall Simulation Innovation Workshop*, Orlando, Florida: SISO
- Defense Systems (2020, May 15). *Sensory overload: military is dealing with a data deluge*. Retrieved on 15May20 from: <https://defensesystems.com/articles/2010/02/08/home-page-defense-military-sensors.aspx>
- Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. *arXiv preprint arXiv:1706.07068*.
- Elstad, E.C., & Davis, D.M. (2017). Implementing Innovative Constructivism: An Architected Approach to Enhancing STEM Education. In the *Proceedings of the Interservice/Industry Simulation, Training and Education Conference*. Orlando, Florida, 2017
- Gervits, F., Thurston, D., Thielstrom, R., Fong, T., Pham, Q., & Scheutz, M. (2020) Toward Genuine Robot Teammates: Improving Human-Robot Team Performance Using Robot Shared Mental Models. In the *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*. (AAMAS 2020, Auckland New Zealand),
- Green, L. C. (2018). *The contemporary law of armed conflict*. Manchester University Press; 3 edition (May 1, 2008)
- Gong, L., Nass, C., Simard, C., & Takhteyev, Y. (2001). When non-human is better than semihuman: Consistency in speech interfaces. *Usability evaluation and interface design: Cognitive engineering, intelligent agents, and virtual reality*, 1558-1562.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).
- Griffith, S. B. (1963). *Sun Tzu: The art of war*. London: Oxford University Press.
- Haas, M. W., & Hettinger, L. J. (2001). Current research in adaptive interfaces. *The international journal of aviation psychology*, 11(2), 119-121.
- Hochreiter, S., & Schmidhuber, J. (1997). LSTM can solve hard long time lag problems. In *Advances in neural information processing systems* (pp. 473-479).
- Hoppe, M., Kosch, T., Knierim, P., Funk, M., & Schmidt, A. (2019, May). Are Drones Ready for Takeoff? Reflecting on Challenges and Opportunities in Human-Drone Interfaces. *1st International Workshop on Human-Drone Interaction, Ecole Nationale de l'Aviation Civile [ENAC]*, May 2019, Glasgow, United Kingdom
- ICT, (2019). *Virtual Humans*, Retrieved 17 February 2019 from: <http://ict.usc.edu/groups/virtual-humans/>
- Jolie, E. W., (1978). *A brief history of U.S. Navy torpedo development*. Navy Underwater Systems Center Report 5436. Newport, Rhode Island
- Juergens, J. M. (2010). *CPO leadership: unique and innovative leadership characteristics of senior enlisted that sustain Naval operations*. Naval Postgraduate School, Monterey, California.
- Kaimakis, N.J., Davis, D.M., Breck, S. & Nye, B.D. (2018). Domain-Specific Reduction of Language Model Databases: Overcoming Chatbot Implementation Obstacles. In the *Proceedings of the MODSIM World Conference*, Norfolk, Virginia.

- Kaimakis, N.J., Davis, D.M., Breck, S. & Nye, B.D. (2018). Domain-Specific Reduction of Language Model Databases: Overcoming Chatbot Implementation Obstacles. In the *Proceedings of the MODSIM World Conference*, Norfolk, Virginia.
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4401-4410).
- Keegan, J. (2011). *A history of warfare*. Random House. New York, New York.
- Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., & Piepol, D. (2007, November). Building interactive virtual humans for training environments. In *Proceedings of Interservice/Industry Training, Simulation and Education Conference* (Vol. 174).
- Khalid, H. M., Shiung, L. W., Sheng, V. B., & Helander, M. G. (2018). Trust of Virtual Agent in Multi Actor Interactions. *Journal of Robotics, Networking and Artificial Life*, 4(4), 295-298.
- Knierim, P., Kosch, T., Schwind, V., Funk, M., Kiss, F., Schneegass, S., & Henze, N. (2017, May). Tactile drones-providing immersive tactile feedback in virtual reality through quadcopters. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 433-436).
- Kott, A., & Perconti, P. (2018). Long-term forecasts of military technologies for a 20–30 year horizon: An empirical assessment of accuracy. *Technological Forecasting and Social Change*, 137, 272-279.
- Krumwiede, K. & Liou, R., (2019). The China electric vehicle industry, *Strategic Finance; Montvale* 10:4,
- Kubrick, S. (1968). 2001: a space odyssey.
- La Falce, W. (2017). The officer, NCO Relationship, *The NCO Journal*, US Army University Press. Retrieved 23 February, 2020 from: <https://www.armyupress.army.mil/Portals/7/nco-journal/docs/Officer-NCO-Relationship.pdf>
- Lee, K. F. (1988). Automatic speech recognition: the development of the SPHINX system. (Vol. 62). *Springer Science & Business Media*.
- Levick, E. (2019). A robot is my Wingman. *IEEE Spectrum*, 57(01), 32-56.
- Liu, J., Mohan, A., Kalia, R. K., Nakano, A., Nomura, K. I., Vashishta, P., & Yao, K. T. (2020). Boltzmann machine modeling of layered MoS₂ synthesis on a quantum annealer. *Computational Materials Science*, 173, 109429.
- Mann, S., Leonard, B., Brin, D., Serrano, A., Ingle, R., Nickerson, K., ... & Janzen, R. (2016). Code of ethics on human augmentation. *VRTO Virtual & Augmented Reality World Conference+ Expo*, June, 25-27.
- Marcot, R. M. (2020, May 26). *The Development of The Gatling Gun*. Retrieved from the internet 26May20 from: <https://www.scribd.com/document/353815818/Gatling-Gun-Development-of-B065-Marcot>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Multisensory Interface Design for Complex Task Domains: Replacing Information Overload With Meaning in Tactical Crew Stations. In *The International Journal of Aviation Psychology*
- Nye, B.D., Swartout, W., Campbell, J., Krishnamachari, M., Kaimakis, N. and Davis, D.M. MentorPal: Interactive Virtual Mentors Based on Real -Life STEM Professionals. In the *Proceedings of the Interservice/Industry Simulation, Training and Education Conference*. 2017.
- NY Times, (2020 May 27), *Whitehouse confirms killing of Terrorist Leader in Yemen*, Retrieved on 27May20 from: <https://www.nytimes.com/2020/02/06/world/middleeast/terror-yemen-drone-attack.html> .
- Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., ... & Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5), 1-36.

- Rizzo, A., Forbell, E., Lange, B., Galen Buckwalter, J., Williams, J., Sagae, K., & Traum, D. (2012). Simcoach: an Online Intelligent Virtual Human Agent System for Breaking Down Barriers to Care for Service Members and Veterans. *Healing War Trauma A Handbook of Creative Approaches*. Taylor & Francis
- Rizzo, A., Lange, B., Buckwalter, J. G., Forbell, E., Kim, J., Sagae, K., ... & Parsons, T. (2011). SimCoach: an intelligent virtual human system for providing healthcare information and support. *International Journal on Disability and Human Development*, 10(4), 277-281
- Rizzo, A., Shilling, R., Forbell, E., Scherer, S., Gratch, J., & Morency, L. P. (2016). Autonomous virtual human agents for healthcare information support and clinical interviewing. In *Artificial intelligence in behavioral and mental health care* (pp. 53-79). Academic Press.
- Satori, H., Harti, M., & Chenfour, N. (2007). Introduction to Arabic speech recognition using CMU Sphinx system. *arXiv preprint arXiv:0704.2083*.
- Shanker, Thom and Richtel, Matt (2011). In New Military, Data Overload Can Be Deadly. *New York Times*, January 16, 2011. P 16
- Shaw, K., Davis, D.M., Rizvi, S.Z., & Davis, M.C. (2019). Quantum Computing: Evaluating Potential Quantification of Projective Psychological Test Scoring. In the *Proceedings of the ModSim World Conference*. Norfolk, Virginia
- Swartout, W. R., Nye, B. D., Hartholt, A., Reilly, A., Graesser, A. C., VanLehn, K., ... & Wang, L. (2016, March). Designing a Personal Assistant for Life-Long Learning (PAL3). In *The Twenty-Ninth International Flairs Conference*.
- Traum, D., & Rickel, J. (2002, July). Embodied agents for multi-party dialogue in immersive virtual worlds. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: part 2* (pp. 766-773). ACM.
- Traum, D., Jones, A., Hays, K., Maio, H., Alexander, O., Artstein, R., ... & Jungblut, K. (2015, November). New Dimensions in Testimony: Digitally preserving a Holocaust survivor's interactive storytelling. In *International Conference on Interactive Digital Storytelling* (pp. 269-281). Springer International Publishing.
- Yao, K-T., Davis, D. M., Liu, J. J., & Kaimakis, N. J. (2018). New Technologies to Enhance Computer Generated Interactive Virtual Humans. In the *Proceedings of the SISO Fall Simulation Innovation Workshop*. Orlando, Florida:SISO
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55-75.