Advancing Virtual Conversation Efficacy: Technologies to Enable Predictive Language Initiation

Dan M. Davis & Benjamin D. Nye
Univ. of Southern California
Los Angeles, California
dmdavis@acm.org & nye@ict.usc.edu

Brian L. Morgan
Naval Post Graduate School
Monterey, California
blmorgan@nps.edu

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

ABSTRACT

The thrust of this paper is to analyze the emerging level of conversational interactions by virtual humans. The authors relate their experience with several projects using Natural Language Processing (NLP) in pure virtual humans or by using videotaped conversational answers from live humans on the topic to be addressed. They present data on "conversationality" of these computer managed interfaces that supports the increasing level of acceptance of the users as they relate to computers as if they were live human beings. Nevertheless, it has been observed that there are a number of instances when the conversation indicated a dramatic divergence from live human behavior due to the lack of the computer's ability to initiate new conversational topics in a natural and conversational way. Several efficacious aspects of virtual humans could profit greatly from the extensibility of NLP into a more proactive interface with users, The paper discusses a few of these opportunities as being critical in the face of the lack of good alternative approaches to vital concerns. What follows is a description of the overall goals of this research area and the impediments currently hampering progress. Then, there is a presentation of advances in high performance computing, *e.g.* new implementations of quantum computing, and breakthroughs in data management and artificial intelligence, *e.g*. Deep Learning developments, that will enable virtual human conversational improvements which were heretofore impossible. Quantification is offered and data is reported to support these contentions. The paper closes with a discussion, based on the authors' experience over decades, of ways in which this community can accelerate the emergence of the necessary capabilities and enhance the early adoption of the new implements and methods. The paper concludes with guided and guarded speculation about potential extensibilities of these insights into other related fields.

ABOUT THE AUTHORS

Dan M. Davis is active as a consultant at the Institute for Creative Technologies, University of Southern California (USC), focusing on large-scale DoD simulations and avatar uses. 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 in 2002 as a Commander, U.S.N. He received B.A. and J.D. degrees from the University of Colorado in Boulder.

Benjamin D. Nye, Ph.D. is the Director of Learning Science at the University of Southern California, Institute of Creative Technologies (USC-ICT). Ben's research tries to remove barriers development and adoption of adaptive and interactive learning technology so that they can reach larger numbers of learners. Dr. Nye's research has been observed for excellence in intelligent tutoring systems (1rst Place ONR ITS STEM Grand Challenge; Nye, Windsor, et al., 2015; Nye et al., 2014), cognitive agents (BRIMS 2012 best paper; Nye & Silverman, 2013; Nye, 2012), and realistic behavior in training simulations (Federal Virtual Worlds Challenge; Silverman et al., 2012). His research is on scalable learning technologies (Nye et al., 2014) and design principles that promote learning (Nye, Graesser, & Hu, 2014; Nye, 2014; Nye, Morrison, & Samei, 2015). This research has led to 20 peer-reviewed papers, 11 book chapters, and 5 open-source projects. Ben is the membership chair for the International of Artificial Intelligence in Education (IAIED) Society and holds memberships in Educational Data Mining Society (EDM), and Association for the Advancement of Artificial Intelligence (AAAI) He also co-chairs the FLAIRS Learning Technologies track (2015-2017).

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.

Brian L. Morgan, CAPT, USN, Ret. is currently listed as a member of the Faculty at the Naval Postgraduate School in Monterey California. He has over six years of experience in the practical application of operations research at the Service headquarter level while assigned to the Assessment Division (OPNAV N81) in the Pentagon. He is a Naval Flight Officer with over 3,700 hours in the E-2C Hawkeye and served as Commanding Officer, VAW-117. His military assignments also include service on the Joint Staff as Chief of Operations Test and Evaluation Branch, National Defense Command Center. Professionally, he is President-Elect, Defense Operations Research Society and a Council Member, INFORMS Defense and Security Society. Brian received a BS in Aeronautical Engineering from the University of Virginia and an MS in Operations Research from the Naval Postgraduate School.

Advancing Virtual Conversation Efficacy: Technologies to Enable Predictive Language Initiation

Dan M. Davis & Benjamin D. Nye
Univ. of Southern California
Los Angeles, California
dmdavis@acm.org & nye@ict.usc.edu

Brian L. Morgan
Naval Post Graduate School
Monterey, California
blmorgan@nps.edu

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

INTRODUCTION

The advances needed and capabilities emerging will allow the human avatar computer agents to emulate more spontaneous generation in their virtual conversations with human users. Great strides have been made in Artificial Intelligence (A/I) and Natural Language Procession (NLP) that now enable digital computers to respond to users in a conversational way that emulates a human discourse. Their ability to initiate conversations is much more severely limited. The authors will present insights from their design, implementation and evaluation of systems designed to be a surrogate for human-to-human conversations. These response-oriented systems have shown results from human testing. The paper outlines and emphasizes why a computer-enabled spontaneous generation may be important. Further the authors will discuss current software, hardware and system constraints that make solving these problems difficult, it not virtually impossible. Then, there follows discussion of a few emerging techniques and technologies that the authors will conjecture how these new capabilities may enable of more human-like interaction with the users. There is an outline on how to implement a virtual equivalent, several alternative algorithms and approaches with the varying features of each being evaluated. The paper considers the use of Artificial Intelligence (A/I), as well as the potential hardware suites and long-haul digital communications necessary for the success of the system. The authors envision potential misfeasance by the A/I functions and an Ethical A/I approach is recommended. Following that is an outline and justification for the appropriate metrics for such an effort and the best way to measure and evaluate the data against goals of the project. Additional research and probable impacts are to be expected and should be studied and resolved. The authors support the implementation and exploitation of emerging hardware capabilities, *e.g.* quantum computing and software advances. The paper proposes several future research possibilities. Next is the analysis of the needs for these emerging capabilities and the opportunities for extensibility into other spontaneous generations and implementation implications are discussed to benefit the readers who may want to pursue these approaches.

Current state of the art

The capabilities of the human-computer interface have advanced significantly since the advent of common accessibility to computers beginning half a century ago. Today, virtually every family has several devices that respond to the human voice: "Hey Siri, what's the weather going to be tomorrow?". In 1968, when "2001: A Space Odyssey" was released, the viewers were fascinated by "HAL" who engaged the human astronauts in very human-like dialogues. Some of the portrayed technology was hard to implement in real life, but is common today. Nor is it done with a huge room-sized computer, but on a hand-held device.. Even though the state of the art has yet to create a threatening presence like HAL, it is working, but is still missing a human-like spontaneous conversation generation. The authors assert that many of the above issues can be soon be addressed by the efficacious application of advanced computation. Some of these computer assets may necessarily be other than digital computers (Figure 1). Research into both Quantum computing, and a reawakened interest into analog and hybrid analog/digital computers may be useful to provide the desired and required computation power. Several operational

Figure 1. A Production Quantum Annealer

quantum computers and quantum annealers are now available to researchers and techniques for their use are now being devised. As now envisioned within quantum computing theory, these machines will not replace digital computers, as they are probabilistic and stochastic by their very natures, while many computational tasks require deterministic precision.

Defense Context

The paper asserts the thesis that there is a significant interest within the defense community concerning communication issues, but they are not the same as the non-defense issues that are analogous. These issues may be addressed alone, but they are mutually intertwined in their impact on the issues of national defense. The authors urge caution in any attempt to generalize from one community to another.

One issue is the articulation and documentary formalization of the need for security. Another difference is the criticality of the duties assigned. The defense agencies are distinguished by a very special understanding of and mutual adherence to job boundaries. The military group's component part identities are much more ingrained and internalized than in civilian life. Each person's status within the organization is clearly marked on their uniform and reports of rebellions against those rules of seniority are rarely reported. Another observation from the authors experience is that the defense personnel often come with a much more deeply ingrained sense of community oneness than the nation as a whole, in academia or in other civilian contexts. In the services, membership in a unit is more a commitment than a status. The last issue is that of the number of lives at stake and the personal experience of loss; one in 20 Marines perished in their one year tour in Vietnam (National Archives, 2020). As opposed to many civilian industries, DoD personnel are there to fight and that life threatening set of environments is understood by most people to be significantly more stressful.

Another issue with which most DoD personnel would be very familiar is the recognition of and adherence to the combat job training and assignment principles. The authors have all seen how personnel trained and certified in one occupational specialty, but manifesting a much greater skill in another needed capability, will be informally and exclusively assigned to duties for which they are neither trained not certified. That tradition even finds its analog in the command structure. There are times in the service when, while retaining the position of the titular head of a unit, a wise commander will designate much of the day-to-day operation of the unit to a more compelling and efficacious leader. Again, the widespread regard for mutual respect for position and hierarchy is invariably reflected in the junior's strict adherence to deference to the senior, albeit the exercise of the command functions delegated to him from that senior. All of these issues keep the Tuckman insights, Forming, Storming, Norming and Performing, (Mind Tools, 2019) as to the formation of teams in play, even in the more formal, defense community context.

In addition to that, in the Defense , when a new unit is formed, it is very likely that all of the personnel will not have ever met the others in the unit. In the civilian experience, it is not infrequently the case that new teams are made up of a majority of members who have been on similar teams with the members of this new team, which is true in academic teams. Again referring to the motion picture industry, a producer usually has a reasonably free rein on the professionals he employs and some experience with the personalities of those candidates, *e.g.* it was well-known in the California film industry that hiring David Niven would entice other actors to join the cast, as he was famous for his civil congeniality that made every movie in which he participated a better production experience (Lord, 2019). Armed forces commanders have virtually none of that freedom in selecting their own crews.

One major issue is the geographical separation of prospective team members. The first several stages of the Tuckman analysis require personal contact and face-to-face experience. The authors maintain that this is also best accomplished with a mediating senior being present. Falling once again on the motion picture industry, there are often protracted face-to-face negotiations as to team composition and duties prior to the first day of team formation.

The most salient difference could be the differences in goals. A slow or misbegotten team in civilian life may lead to loss of money or market advantage. In a combat context, the same delay or ineffective effort in team formation may very well result in, not only equipment and monetary losses, but in mission failure or loss of life.

COMPUTER AGENT CONVERSATIONS RESEARCH

One area of rapid expansion in the computational sciences is the area of Virtual Humans and Natural Language Processing. As the world has observed, early predictions of virtual human simulations and other advances in this arena have not matched the predictions. Some areas have sprung ahead in ways not envisioned. Dick Tracey's wrist watch phone has been put to shame by current models of smart watches that not only provide instant communications anywhere in the world, but also sense an unexpected fall and summon medical care on their own (Technokogizer, 2019). On the other hand The "HAL 9000" fully conversational, sentient and self-aware computer from the movie Space Odyssey is still a significant ways off (Wikipedia, 2019). However, great strides have been made at USC's Institute for Creative Technologies and elsewhere.

Current Advances

Artificial Intelligence (A/I) and Natural Language Processing have now achieved the stage in which personal conversational interfaces are possible. Acceptance of these means that large-scale implementation issues are increasingly seen as imminent challenges (Tram, *et al.*, 2015). That, coupled with the popularity of hand-held devices such as tablets, MP3 players, and smart-phones has brought to the fore the issue of not over-burdening these limited resources with data storage and data transfer requirements that would tax their memory storage limits and bandwidth constrictions. The Institute for Creative Technologies (ICT) of the University of Southern California has been engaged by the Office of Naval Research to conduct research on MentorPal, a computer-generated mentor capable of sustaining a conversational series of responses to secondary school students who are considering careers, especially careers in one of the Science, Technology, Engineering, and Math (STEM) disciplines. An early version is shown in Figure 2. The prototype for this interactive mentor targeted the Microsoft Surface® line of tablet computers. The limited power, RAM and secondary memory of these small portable devices mandates optimal sizing of all processes and data storage and transfer (Nye *et al.*, 2017).

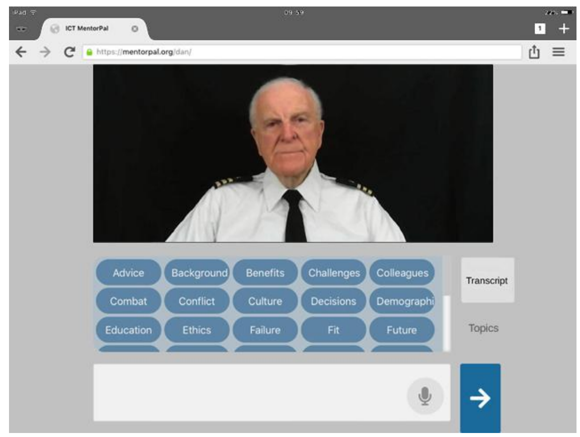

Figure 2. A virtual conversation Interface

The insights gained and the techniques employed apply to model reductions in practically any specific domain, rather than only to chatbots. Part of this assertion is recognizing the need to prepare for operations under conditions of limited computational power and data capacity, while at the same time designing code to make it amenable to scaling if a less constricted environment is available. This would allow computer assistance for team formation, no matter where the team members were stationed prior to the teams converging. These constrained environment issues are known to be common in much of the work currently being conducted at ICT (ICT, 2018).

The basic objective of the Mentor PAL project was to provide a proof-of-concept version of career mentoring to students who may otherwise have no, or severely limited, access to advice or mentoring as they face the daunting task of selecting a career. The ultimate goal is to increase the number of technically trained personnel available. They might be used in the uniformed services or as civilian researchers for the DoD. To accomplish this, prospective career candidates may make other choices if they do not have someone to whom to speak who is experienced in and feels positively toward the desired careers. The authors envision a similar mentoring process in the team formation scenarios of the future. Users, who do not otherwise, have access to a mentor, may benefit from a virtual mentor. The reason for such lack of access may be due to a number of reasons, the two of which are: geographic remoteness and poor network connectivity (Crisp *et al.*, 2009).

The solution advanced by ICT was to produce a computer-generated mentor to be available via any system with the appropriate software installed. The user would be presented with a computer screen interface via which he could engage a mentor in a conversation-like exchange about issues of concern. Input from the user would be by text entry or audio speech recognition software. This mentor was envisioned as being both compelling and engaging. ICT advanced the position that the mentors would be most compelling and engaging if they were real people (Traum *et al.* 2015) recorded live and exhibiting a demonstrated screen presence.

Implementing a Conversational Agent with Spontaneous Generation

Initially, a list of germane questions would be drafted. Depending on the topic, this might total a thousand or so items. The creative process requires a broad knowledge of both the questions that are a part of the mentor's experience and represent the questions that the target users are likely to ask. In the previous projects, the team studied input from several members who had professional and Navy experience in order to generate a thorough list of questions. A previous paper (Nye *et al.*, 2017) discussed this process at greater length.

The next step is to have the mentors record a varying number of video clips relating to their experience in the service and their duty in their own profession. These clips are then stored in a standardized database. When the user poses a question the program selects the most appropriate clip to play. This must be done very rapidly to sustain the desired "conversational" effect of the exchange. Based on the team's experience with briefing users, the clips were designed to be on the order of 90 seconds or less. This technology rests on previous research into the use of Virtual Humans acting as program interfaces in counseling (Morbini, *et al.*, 2012) and in history capture (Artstein *et al.*, 2014). The team found that these videotaped mentors are difficult to classify easily: sometimes being referred to as Virtual Humans, but they are actually intelligent agents that sequence and present recordings of real people. This has been termed "time-offset interaction" in some related work. There still needs to be a substantial corpus of carefully collected questions and answers available to the system that can be answered within the latency limits that circumscribe the feel of conversational discourse.

In Figure 3, the flow begins at the top left and proceeds clockwise through the process. The icons are notional only and should not deceive the reader as to their importance, weight or size, *e.g.* the storage icons approximately the same image-on-page size, whereas, in physical fact they differ by three orders of magnitude in data size. The important steps are the data flows from the storage locations into and out of the computational functions.

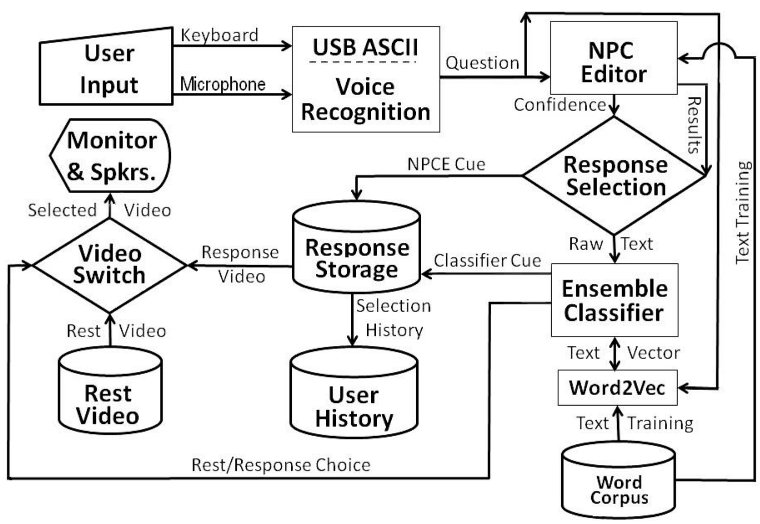

Figure 3. MentorPal Notional Flow Chart

OPPORTUNITY FOR ENHANCED UTILIZATION

Based on what the authors have observed in their professional careers and on what has developed with his research and several decades of further observation, the authors now turn their attention to problem areas that may be amenable to computer-enabled intervention. There are basically four areas in want of additional assistance:

- Providing guidance, counsel, and stress relief
- Articulating acceptable "field expedients" featuring a more free-form skills assessment and assignment;
- Helping to bridge the communications gap for personnel assigned around the globe.
- Reiterating the imperatives mandated by the national goals and the lives that are in the balance.

24 Hour Mentor Access

Using the technologies described above, it is possible to design, implement and field a set of mentors who are focused on assisting the prospective team members get through the roughest stages of new team formation. Some enhancements to current capabilities would be useful, but just implementing what is already been proven will have a markedly beneficial impact.

Implementing a Virtual Conversation with Spontaneous Generation

It is now putatively possible to create a Virtual Conversation with spontaneous generation. A number of choices need to be made, but the decision is just a search of optimum effect and a wrong choice would impact efficacy, but would not be fatal to the concept. And example of such a choice would be an animated literal virtual human, speaking with synthesized voice driven by text. This is what was done in the ICT project called SimCoach; as shown in Figure 4. The other option, which was discussed above, is the utilization of a thousand or so video clips of live humans. This technique is often also called a "virtual human," but the purest point out the subject is not "virtual" at all, but a real, live human. This terminology difference will be addressed later.

Figure 4 - SimCoach for Advising VA Patients

There are two known hurdles to this vision. First: studying, evaluating, and collating the appropriate question data set. The authors have submitted a paper on this topic to another conference on how such a data set was created for the current project, called PAL3, A/I. (Davis *et al.*, 2020). Second, it has been the authors' observation that much of good mentoring is asking the salient questions. The second is that NLP is very good at answering direct inquiries, but it would be a new investigative effort to identify, draft, implement and validate a series of germane questions or probes to get the users to discuss and analyze their issues on their own, as well as a bank of well-formulated suggestion to resolve common issues that are highly likely to arise.

It is envisioned that part of the solution will be to urge the participants to set up and schedule regular video conferencing meetings. There are now a range of video conferencing programs available to the public at no direct cost. Most of them now support multi-user sessions and various forms of document sharing. The one issue that I think has yet to be resolved would be the participation of those relying in satellite connections: remote sites in areas not served by cable and ships at sea The latencies imposed by the round trip to geostationary satellites can result in 900 to 1500 milliseconds, which makes conversation problematic, and just the speed of light limits adds at least 550 milliseconds to the response time (V-Sat Systems, 2019) and that is just enough delay to cause some minor disruptive interference to the normal flow of conversation. Even so, it could be a very efficacious tool for resolving some of the counter-productive impacts of the Tuckman “Storming” stage.

Standardizing Terminology

In an emerging technology such as Virtual Humans, one of the issues that continually hampers the achievement of consensus as to progress and the discovery of the optimal path for future research is the lack of consistency in both the terms used to describe the various techniques and the absence of carefully considered quantification of performance. As mentioned above, the term Virtual Humans is used in the technical vernacular to refer to any human avatar, be the image a true animation object portraying a Virtual Human or a really just a computerized delivery of a real live human. Comparative studies are needed to show the relative costs, time constraints, performance, and impacts of both implementations.

The call for more studies on performance brings to light the need to better quantify many factors now left to subjective assessments like face validity and anecdotal reports. The human behavior disciplines, e.g. social psychology, systems engineering and operation research rely on quantified measures of merit in order to produce the analyses that would facilitate future analyses and enable more optimal pursuit of goals that would result in actual benefit to the Warfighter.

The simulation community may be able to make good use of the experience of the standards discipline in managing the standards for various physical and process operation. Just this professional expertise could help the simulators in creating and promulgating and new *lingua franca* for Virtual Human research.

NEW TECHNOLOGIES

Nearing the end of the Moore's Law expansion of digital computing, many simulation professionals are concerned with solving the grand challenges set forth above. One of the alternatives frequently mentioned is Quantum computing. It has hopefully been considered an extension of computational capability since the Nobel Laureate Richard Feynman presented the seminal paper in 1982. In that paper he held that: "… with a suitable class of quantum machines you could imitate any quantum system, including the physical world." (Feynman, 1982). The authors have assiduously followed the development of such a machine and those devices still are, as near as can be ascertained, almost entirely at the test-bench phase. There seem to be no such "general purpose" quantum computer that is even nearing operation. There is one operational design in the quantum world: while not a general purpose quantum computer, it relies on very cold temperatures (15 milliKelvin) to create a useable quantum effect. (Lucas, 2013)

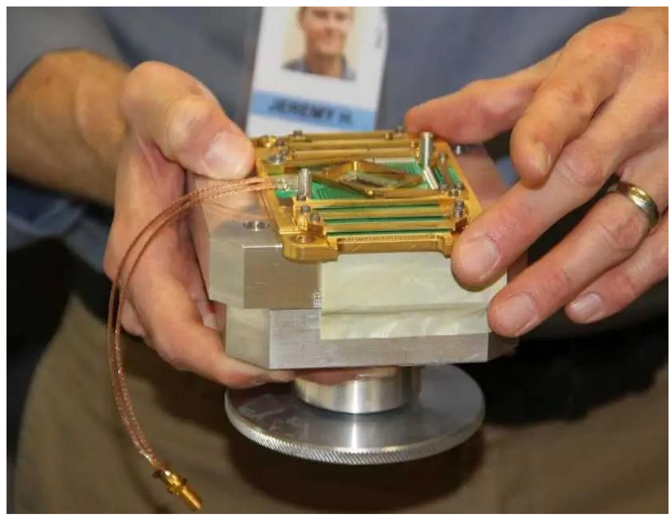

Figure 5 – Quantum processor of 1,025 Qubits before being cooled to 0.015 Kelvin.

This early production version of a Quantum Computing deveice has been conceived, designed, produced, and delivered to the University of Southern California. It has been in operation since 2012. In its current configuration, D-Wave computers have a design providing approximately 2,000 qubits, but need to be cooled to very close to absolute zero. Figure 5 gives an idea of scale, but Figure 1 shows the size of the case required, but does not show the size of the two story tall tank of refrigerant required. The "approximate" figure is required as some delivered machines have some fraction of these qubits turned off and a small number of the qubits (~ 1%) are not stable after the processor reaches the target 15 mK. Figure 5 shows the D-Wave Two, as installed in the USC-Lockheed Martin Quantum Computing Center (QCC) at the Information Sciences Institute (ISI) in Marina del Rey. There is another in the San Francisco Bay area in a joint Google and NASA project, and one has putatively been ordered by the DoD's High Performance Computing Modernization Program. Others are in varying stages of procurement and delivery.

Recently, other authors have touted quantum computing's ability to produce more power, using terms like "magic" to stir the imagination and whet the appetites of the user community. They point out that the capability of quantum computers arises from the different way they encode information. Digital computers represent information with transistor-based switches having a state of 0 or 1, labeled as a bit. In contrast, the basic unit of quantum computer operation, the quantum bit or qubit, can exist simultaneously as 0 and 1. A quantum bit, called a qubit, might be represented by an atom in one of two different states, which can also be denoted as 0 or 1. Two qubits, like two c1assical bits, can attain four different well-defined states (0 and 0, 0 and 1, 1 and 0, or 1 and 1). But unlike classical bits, qubits can exist simultaneously as 0 and 1, with the probability for each state given by a numerical coefficient. Data in a two-qubit quantum computer thus requires four coefficients. In general, n qubits demand 2^n numbers, which rapidly becomes a sizable set for larger values of n. For example, if n equals 50, about 10^{15} numbers are required to describe all the probabilities for all the possible states inside the quantum machine. That large number exceeds the capacity of the largest digital computer. A quantum computer should demonstrate incredible computational power because it can be in multiple states at once, a condition called "superposition." Also, perhaps more importantly, it can act on all its possible states simultaneously. The quantum computer can evaluate a series of optima that would be beyond the power of the largest digital cluster (Gershenhorn, 1998). Evaluating a large number of parameters to "learn" patterns and behaviors is at the heart of Virtual Human enhancement.

The authors have witnessed and participated in the development of high performance computing for several decades and have developed a significant body of experience with newly introduced technologies. They were engaged in the very early introduction of parallel computing and aware of its rivalry with sequential computing and with vector computing. They heard the detractors of parallel computing argue the limits of parallelism and the proponents who argued that it could be used more universally. While acknowledging there are many problems that have remained outside of the easily parallelized arena, it is evident that the majority of all large-scale computational problems are now run in parallel. This is due to the application of new techniques to decompose both data and computation in efficacious ways. Such technology has proven very useful to the simulation community, which has many issues identical to the test and evaluation environment. By using super-cold processors, the D-Wave has been able to demonstrate accepted quantum computing, see Figure 6 to the right. Even if the projected speed-ups are not realized on this design, it is a workable and verifiable quantum computing device. (Albash & Lidar, 2018)

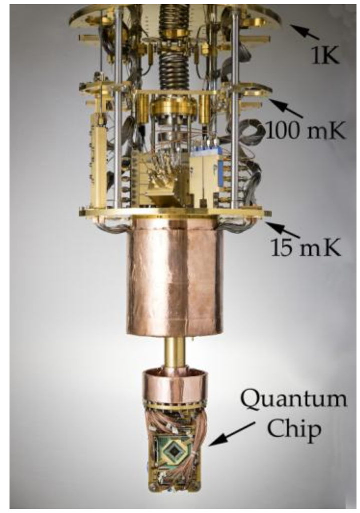

Figure 6. D-Wave Cooling Tower

The D-Wave machine in its developmental adolescence; it is no longer helpless, but its potential remains to be seen and characterized. This phenomenon was described by the late Professor Clayton Christensen in his series of books on disruptive technologies (Christiansen, 2011), *i.e.* a period in which the new technology is operational, but not yet competitive with technologies that have had decades of development and optimization. The criticisms and skeptical comments are familiar to all who have witnessed the displacement of older technologies. However, this community has also seen the claims of new capabilities that were then quickly set aside when old technologies improved faster than the new technology could prove itself. These issues are all part of this community's history and provide insights for analyzing the value of new technologies.

Several things need to happen before that D-Wave becomes a tool rather than a proof of concept device. The hard-wired connectivity paths need to be increased to represent a viable representation of the experimental environment sought to be simulated. A huge code base, both theoretical underpinnings and practicable software approaches, must be conceived and implemented. Perhaps decades of developmental advances will be needed to bring even this limited capability to its promise of revolutionary change. But this community will also remember the naysayers of the early days of personal computing and the skeptics of the first half of the 90's when the internet was first becoming available to the public at large.

Current plans at D-Wave include: more connectivity pathways, more qubits and, perhaps most critically, improved software in both the applications and the digital-to-quantum interface. The programmers at USC have used all of the D-Wave programs and become familiar with their Application Program Interface (API). D-Wave currently supports C/C++, Python, and MATLAB, three very common research languages. The company provides both on-line and live tutorial sessions to bring new users on-board. Much of the USC work focused on the treatment of large databases, known today as "big data.".

Deep Learning

Deep learning is another term that is used to describe an emerging technique in machine learning and artificial intelligence. The authors take it to be an extension of earlier work (DeMuth, 2014) in the areas of neural networks, evolutionary computing, and data mining. But in deep learning, the refinement is done by several layers of convolution processing, in which some evaluation process sends along only such information that the layer values as beneficial. These successions of processing steps are call hidden layers and are the distinguishing feature of deep learning. Deep learning advocates note that this hierarchical approach to sifting large amounts of data is more in keeping with real word issues, albeit it's need for more compute power.

Deep learning has achieved state-of-art in performance in a wide range of problems, including hand writing recognition, object recognition, classification in images, speech recognition, understanding natural language text, and adversarial games, (LeCun, *et al.*, 2015). A key characteristic of being able to tackle these problems is the need to find complex structures in high-dimensional data. Deep learning with its multiple layered structure and distributed representation turns out to be well suited for these problems as compared with other machine learning approaches and other manual knowledge engineering approaches.

The current interest in deep learning can be viewed as the third wave of neural network development. In the late 1950's the Perceptron ® captured the public imagination as a new type of electronic brain (Bernstein, *et al.*, 1981). But, the interest in Perceptron waned when Minsky and Papert pointed out the limitations of a single Perceptron, which can only capture simple linearly-separable structures. In the 1980's, interest renewed as researchers constructed much more capable neural network models with multiple layers with multiple neurons per layer. But, again interest waned, because multiple layered networks are difficult to train and the computation power was not sufficient at that time. The current interest in neural network started in 2006 as Hinton *et al.* introduced a novel method to pre-train the network evaluation weights for deep belief networks using Restricted Boltzmann Machines (RBM) (Hinton, *et al.*, 2002). Moreover, the advent of graphical processing units (GPUs) with their Single Instruction Multiple Data (SIMD) capabilities reduced the computational bottleneck, often reducing weeks of computation to a few hours. GPUs are increasing in their power, but also face physical limits to their computational expansion.

Of the three major deep learning approaches (feed-forward neural networks, recurrent neural networks, and Boltzmann machines), the Boltzmann machine approach is the most amenable to quantum annealing. Boltzmann machines are probabilistic neural networks that implicitly define probability distribution over the activation states of the neurons in the deep learning network. Training a Boltzmann machine requires being able to repeatedly sample from the distribution of activation states. However, sampling for Boltzmann network with loops can be computationally expensive.

The reason restricted Boltzmann machines can be trained efficiently is because they avoid edges within a layer, which enables layer-based approximate sampling of the network. With quantum annealing, there is the potential of efficiently sampling networks with loops. This enables the creation of a broader range of networks, with more complex topology. The current generation of quantum annealers does support complete graphs; they also impose limits on the topology of the intra-layer edges. This is the definition of limited Boltzmann machines (LBM), which are strict supersets of RBMs, and demonstrate the efficaciousness of the additional edges.

The rise of deep learning is related to the rise of big data. Deep learning models require very large datasets to properly train the neural network weights. The activation of a neuron depends on the activation of its neuron neighbors mediated by the weights of the connection, see Figure 7.

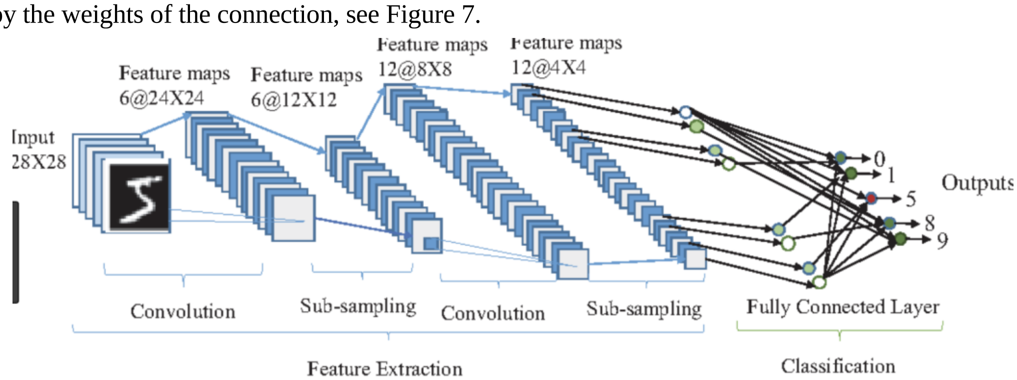

Figure 7. Diagram Showing Successive Layers of Deep Learning Analysis (Medium, 2019)

The winner of the 2012 ImageNet Large-Scale Visual Recognition Challenge AlexNet (Krizhevsky *et al.*, 2010) has over 60 million weights with 5 convolutional layers, max pool layers, and 3 fully connected layers. The network was trained on over 15 million images with over 22 thousand labeled categories. In 2014, Simonyan and Zisserman developed VGG net, which has 19 convolutional layers, and Microsoft ResNet (Jaderberg *et al.*, 2015) has 152 layers.

Deep learning does have limitations and there are recent research efforts that are addressing these limitations. One limitation is brittleness. Deep learning models have demonstrated remarkable ability to perform well in noisy data, for example the ability to observe objects in photos with busy backgrounds. But, paradoxically it is often quite fragile. Given new classes of objects to observe, the model must be extensively retained with many instances of the new objects as well as the previous know objects. In the games area, a deep learning-based Alpha Go program can beat top ranked human players, but the number of handicap stones is hardwired. Changing from the default handicap of 7.5 stones requires extensive retraining. An approach to address this limitation is to develop less monolithic deep networks, such as developing visual attention models to incrementally decode images and combining deep networks with artificial reasoning capabilities. Another limitation is the need for large labeled training dataset. Large datasets are often easy to find, but labeling is often a labor intensive process, *cf.* discussion of "sarcasm" recognition below.

Researchers are developing unsupervised learning techniques, such as using Boltzmann machine-based auto-encoders; and incorporating artificial intelligence techniques (like reinforcement learning in the AlphGo) to self-generate labeled data.

ETHICAL CONSIDERATIONS

In their work on developing on-line mentors, the authors have recorded two instances of sensitivities not otherwise anticipated, despite the ICT teams having had one prior service officer observing all sessions. One was not particularly serious, but involved a Naval Academy graduate who refused to say one of the “repeat after me” words that were used as segue terms if an answer were not to be found. Such terms as “I am not sure I can answer that.” and you’ll have to ask me something else.” The term at which he balked: “Maybe.” He protested that commissioned officers never say maybe and was appently offended the team would suggest that he do so.

Another was junior officer who responded, when asked to differentiate the differences between officer and enlisted ranks, opined that there was no difference. When this response was reported to several flag and general officers, the emotionality of the response indicated that this too was a very sensitive area. All those contacted responded with vociferous and emphatic analyses explicating why and how there was a difference and justifying their responses with voluminous historical examples.

A computer, not recognizing what had happened, might pursue a sensitivity to the point of severely injuring the user. There are examples of human insensitivity resulting in suicides by those offended and humiliated. The ethics of both humans and computer programs must be carefully examined and appropriate limits articulated and validated.

As one might expect, both of those sensitive issues were quickly observed by the humans who were in the loop, but it is unclear how they may have been observed by a computer. The authors propose two, not mutually exclusive approaches:

1. Using a Quantum Computing-enabled Deep Learning program, the computer could be given recorded clips of offensive statements from the entertainment industry (cinema and television video files) and the Deep Learning programs set to isolating those indicia of an offended person, e.g. certain words or tones, body posture, physical attack,

2. Emerging sensor technology designed to assess user emotional states in body language, facial expression, and voice dynamics could be employed to help the computer both “learn” and to observe angst in the user

Closely related to these issues are the issues of confidentiality. To be efficacious, the computers need to keep a database of approaches and efficacy. But the users need to have solid assurance that their confidences are as safe or safer with the computer than they would be with a human therapist.

As a way to assess the issues at hand, the authors suggest using a process not unlike the well-established requirements of an Internal Review Board and concomitantly set up an external ethics panel made up of ethicists, theologians, therapists and clinical psychologists. The goal would be to assure that any user could feel a peace with the steps taken to avoid the danger cited above, along with the more sensational dangers portrayed in movies such as 2001: A Space Odyssey and The Terminator.

CONCLUSIONS

While there is yet much to be accomplished, the early progress in Virtual Human implementations has shown both the promise and the problems in this simulation discipline. This paper has attempted to lay out these by using the work of Professor Tuckman as an exemplar of the issues before the researchers. But the major issue is whether this technology will enable more human-like conversation, including the ability to gracefully “strike up” a conversation.

Issues that Require Virtual Human Conversation with Spontaneous generation

There are several areas of potential use of virtual human conversation that require the computer agent initiate and guide the conversation and not just be responsive to the user. Some of these are team creation coaching, PTSD patient therapy, career counseling, and academic tutoring. The experience of the authors supports the proposition that team formation in large organizations is essential to the achievement of organizational goals. The fusion of disparate personalities, often initially located in site remote from each other, is critical to success. There appears to be a significant lag in team efficacy and this lag may be significantly different in civilian and in defense environments.

Emerging Technologies and Capabilities in Computer Agent Conversation

In virtually every human endeavor, there is a need for mentoring and mediation in order to optimize both the timely achievement of goals and the optimal efficiency in that achievement. This is especially true in combat where goals have national survival implications and costs are not just financial, but human lives. Operations schedules and funding constraints may make the use of computer-generated human avatars more feasible, especially in the light of geographic dispersion of personnel that will form the envisioned team. Advances currently are hampered by a lack of community awareness.

For future research to be efficacious and rapid, a set of community standards needs to be defined and adopted. Much of the work centers on human behavior and a new area "Virtual Human Behavior." These are areas in which there will be a need to compare capabilities and progress. This situation bespeaks a need for both community-accepted terminology and rigorously-defined quantifications. The existing standards professionals could be of invaluable assistance in this process.

Emerging Technologies have Potentials as Enablers of Spontaneous Conversation Generation

Virtual Conversation with Spontaneous generation is physically possible with current technology, but not practicable due to latencies imposed by digital computing limitations. Advances in Deep Learning and Quantum Computing hold promise for breaking the digital barriers and enabling a truly human-like conversationalist. Many benefits from this implementation would accrue to the warfighters who are dispersed and time constrained in their interpersonal relationships with professionals for whom they have a need.

Ethical Issues that May Arise

Several ethical issue that will likely arise with the advent of a virtual human with spontaneous generation One of the more obvious and potentially problematic is the issue of sensitivity. Humans spend their lives generating carefully drawn sensitivities and a quick interview with a person rapidly gives a good sense of that person's awareness of topics that can be destructive to other individuals. There may be A/I approaches a not yet implemented that would address these issues, but the authors in no way represent this is a closed issue. Even the use of a quantum computer and Deep Learning techniques may take a long time to develop a trustworthy program that would sense the apprehension in a user's face and steer clear of objectionable comments.

Perhaps, some researchers will raise the issue that there are some similarly insensitive humans and it is not outside of the realm of possibility that humans will be less offended by a machine mistake that is ostensibly not malicious and a human mistake that may have been intended to injure. The problems are real but the promise of new hardware and more sophisticated A/I techniques create a hopeful environment for the 21st Century.

ACKNOWLEDGEMENTS

The various projects mentioned were supported by a range of DoD organizations and, while this paper reports insights gained during that work that was not the focus of the research. Nevertheless, the positions taken in the paper are the authors' own and do not represent in any way the views of the Department of Defense or the US Government.

REFERENCES

Albash, T., & Lidar, D. A. (2018). Demonstration of a scaling advantage for a quantum annealer over simulated annealing. *Physical Review X*, 8(3), 031016.

Artstein, R., Traum, D., Alexander, O., Leuski, A., Jones, A., Georgila…& Smith, S. (2014). Time-offset Interaction with a Holocaust Survivor. In *Proceedings of the 19th international conference on Intelligent User Interfaces* 163-168ACMFebruary,

Bernstein, Jeremy, (1981). A. I.. *New Yorker Magazine*, New York, New York,

Christensen, Clayton. (2013). The Innovator's Dilemma: when new Technologies Cause Great Firms to Fail. *Harvard Business Review* Press, 2013.

Crisp, G., Nora, A., & Taggart, A (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Education Research Journal*, pp 924-942.

Davis, M.C., Nye, B.D& Davis, D.M. (2020). Populating a Virtual Conversation Database: Community Data Collection and Analysis In the *Proceedings of the ModSim World Conference*, Norfolk, Virginia, (Pending).

Demuth, Howard B., (2014). *Neural Network Design*. Martin Hagan,.

Feynman, R., (1982), Simulating Physics with Computers. *International Journal of Theoretical Physics* 21 (6–7): 467–488.

Gershenfeld, Neil, and Isaac L. Chuang. (1998) Quantum computing with molecules. *Scientific American* 278.6 (1998): 66-71

Hinton, Geoffrey E. (2002). Training Products of Experts by Minimizing Contrastive Divergence. *Neural Computation* 14.8: 1771-1800.

Hornsby, JS. (2018). Critical Elements of Team Formation to Enhance Organizational Innovation In *The Challenges of Corporate Entrepreneurship in the Disruptive Age* (pp123-140). Emerald Publishing Limited.

ICT (2018). *Research Project One-Sheets* Institute for Creative Technologies, University of Southern California. Retrieved on 04 (2019 from: http://ict.usc.edu/research/pdf-overviews/ ..

Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. (2015). Spatial transformer networks. *Advances in Neural Information Processing Systems*.

Krizhevsky, Alex, and Geoffrey E. Hinton. (2010). Factored 3-way Restricted Boltzmann Machines for Modeling Natural Images. *International conference on artificial intelligence and Statistics*.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). Deep learning. *Nature* 521.7553: 436-444.

Lord, G. (2019). *In Thespian Praise of: David Niven*, Paulburgin.blogspot.com25 January 2006Retrieved from the Internet on 02 Dec..

Lucas, Robert F., Gene Wagenbreth, John J. Tran, David R. Pratt, and Dan M. Davis. (2013) Practical Adiabatic Quantum Computing: Implications for the Simulation Community. in the *Proceedings of the Interservice/Industry Simulation, Training and Education Conference*, Orlando, Florida, 2013

Medium, (2019). *Deep Learning*. Downloaded from the Internet in November of 2019 from URL: https://medium.com/tag/deep-learning/latest

Mind Tools (2019). *Forming, Storming, Norming, and Performing: Understanding the Stages of Team Formation*. Retrieved on 21 January, from https://www.mindtools.com/pages/article/newLDR_86.htm.

Morbini, F., Forbell, E., DeVault, D., Sagae, K., Traum, DR., & Rizzo, AA. (2012). A Mixed-spontaneous generation Conversational Discourse System for Healthcare. In Proceedings of the *13th Annual Meeting of the Special Interest Group on Discourse and Discourse* pp 137-139.

National Archives, (2020). *Vietnam War U.S. Military Fatal Casualty Statistics*. Retrieved on 15 April 2020 from: https://www.archives.gov/research/military/vietnam-war/casualty-statistics

Norris, PE., O'Rourke, M., Mayer, AS., & Halvorsen, KE. (2016). Managing the Wicked Problem of Transdisciplinary Team Formation in Socio-Ecological systems. *Landscape and Urban Planning,* Vol 154, pp 115-122.

Nye, B.D., Swartout, W., Campbell, J., Krishnamachari, M., Kaimakis, N and Davis, D.M. (2017). MentorPal: Interactive Virtual Mentors Based on Real -Life STEM Professionals. In the *Proceedings of the Interservice/Industry Simulation, Training and Education Conference*.

Pelton, JN., & Jacqué, B. (2016). Distributed Internet-Optimized Services via Satellite Constellations. *Handbook of Satellite Applications*, pp 1-21.

Takai, S & Esterman, M. (2017). Towards a Better Design Team Formation: A Review of Team Efficaciousness Models and Possible Measurements of Design-Team Inputs, Processes, and Outputs In *ASME (2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.

Technologizer, (2019). ; Time.com *Dick Tracy's Watch: The Most Indestructible Meme in Tech Journalism* http://techland.time.com/2013/02/11/dick-tracys-watch-the-most-indestructible-meme-in-tech-journalism/ .Retrieved from the Internet on 03 December 2019.

Traum, D., Jones, A., Hays, K., Maio, H., Alexander, O., Artstein, R., ..& Jungblut, K., (2015). New Dimensions in Testimony: Digitally Preserving a Holocaust Survivor's Interactive Storytelling. In *International Conference on Interactive Digital Storytelling*, pp269-281. Cham, Switzerland: Springer.

Tsai, J., Bowring, E., Marsella, S., Wood, W., & Tambe, M. (2012). A Study of Emotional Contagion with Virtual Characters. In *International Conference on Intelligent Virtual Agents* (pp81-88). Springer, Berlin, Heidelberg. September,.

V-Sat Systems, (2019). *Latency- why is it a big deal for Satellite Internet?.* Retrieved from the Internet from: https://www.vsat-systems.com/satellite-internet-explained/latency.html on 02 December 2019

Wikipedia, (2019), *HAL 9000.* Retrieved from URL: https://en.wikipedia.org/wiki/HAL_9000 on 02 December 2019b