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 Carolina
mark@woodduckresearch.org

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 support the increasing level of acceptance by 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 Carolina
mark@woodduckresearch.org

INTRODUCTION

Access to computer agents that respond to human inquiries are common, but a truly human interaction has two additional features that are required: spontaneous initiation and content guidance. The advances needed to produce those features depend on capabilities now emerging that will allow the human avatar computer agents to emulate more spontaneous generation in their virtual conversations with human users and to guide the topics discussed. Great strides have been made in Artificial Intelligence (A/I) and Natural Language Procession (NLP) that already 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. Insights from the design, implementation and evaluation of systems intended to be surrogates for human-to-human conversations will be reviewed. These response-oriented systems have shown results verified by human testing. The paper outlines and emphasizes why a computer-enabled spontaneous generation may be important. Further it will discuss current software, hardware and system constraints that make solving these problems difficult, if not virtually impossible. There follows discussion of a few emerging techniques and technologies that may enable more human-like interaction with the users. There is an outline on how to implement a virtual equivalent conversation. There are several alternative approaches with the varying features being evaluated. The potential damage caused by the A/I functions is envisioned and an Ethical A/I approach (Jobin, 2019) is considered.. Some additional research topics are identified and probable technology impacts are considered. These are to be expected and should be studied and resolved. There is support forthcoming for 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. 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 computers became commonplace 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 then, but is common today. Nor is it done with a huge, room-sized computer. It is accomplished on a hand-held device. Even though the state of the art has yet to create a threatening presence like HAL, it is working toward a human-like spontaneous conversation. The ability to generate that is still missing. Many of the above issues can be soon be addressed by the efficacious application of advanced computation. Some of these computer assets may necessarily more advanced and capable than digital computers, such as quantum machines (Figure 1). Research into both quantum computing, and analog and hybrid analog/digital computers may provide the desired and required computation power. Several operational quantum computers and quantum annealers are now

Figure 1. A Production Quantum Annealer

available to researchers, and techniques for their use are now being devised. As quantum computing theory is now understood, these machines' analytic capabilities will be probabilistic and stochastic. They will not replace digital computers which will be needed for computational tasks that require repeatable and stable deterministic precision.

Defense Context

The paper asserts that there is a significant interest within the defense community concerning communication issues, and that they are analogous, but not the same as the non-defense situations. These issues may be addressed by separate groups, but they are mutually intertwined in their impact on national defense. It would be prudent to caution any attempt to generalize from one community to another, but folly to ignore advances in either.

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 difference experienced is that the defense personnel often come with a much more deeply ingrained sense of community 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. These issues make compelling conversational agents more critical in the military context.

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. When missions require a new team, all of these issues keep the Tuckman insights as to team formation in play: Forming, Storming, Norming and Performing, (Mind Tools, 2019). Using a virtual human to address these issues is quite likely to be an issue of future interest, but again, one wanting of the initiation and guidance function.

One major issue is the geographical separation of prospective military 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 senior mediator present. This may be different from in the civilian environments where there are often protracted face-to-face negotiations as to team composition and duties. The needs for and goals of virtual humans may be quite different. 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. In a combat context, the same delay or ineffective effort in team formation may result in, not only equipment and monetary losses, but in mission failure or loss of life. An A/I-enabled virtual human might well be able to better facilitate this process, with less drain on personnel resources.

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 often not matched the expectations. 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 distance off (Wikipedia, 2019). However, great strides have been made in the germane disciplines.

Current Advances

Artificial Intelligence (A/I) and Natural Language Processing can now provide personal conversational interfaces. Acceptance of these at large-scale will generate implementation issues. These are increasingly seen as imminent challenges (Tram, *et al.*, 2015). That, coupled with the popularity of hand-held devices such as smart-phones has brought to the fore the issue of not over-burdening these constrained resources with data storage and transfer demands that would tax their memory storage and bandwidth capabilities. The Office of Naval Research is conducting research on a computer-generated mentor. This mentor is capable of sustaining a conversational using a series of responses to secondary school students who are considering technical careers. An early version is shown in Figure 2. The prototype for this interactive mentor targeted tablet computers. The limited power, RAM and secondary memory of these small portable devices mandated optimal efficiency in all processes and cautious sizing of data storage and transfer requirements (Nye *et al.*, 2017).

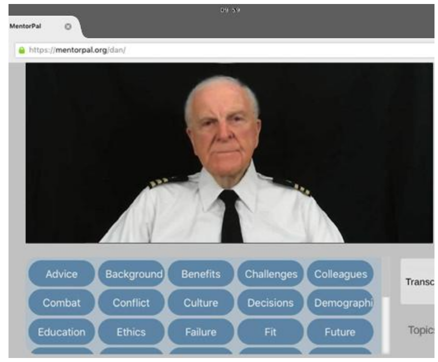

Figure 2. A virtual conversation Interface

The insights gained and the techniques employed apply to virtual conversations in general, but these were limited to responses to student initiated inquiries. Part of this restriction was imposed by conditions of limited computational power and data capacity, designing the code to accommodate the constricted local platform environment. This still allowed computer generated conversations, no matter where the service member was stationed due to operational requirements. These constrained environment issues are known to be common in much of the defense work currently being conducted,

The basic objective of the mentor 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 was to increase the number of technically trained personnel available. They might be employed in the uniformed services or as DoD civilian researchers. Prospective career candidates may make non-technical choices if they do not have a professional to speak to who is experienced and positive. One can envision a similar mentoring process in future team formation or other scenarios. The two main reasons for lacking access to a mentor are geographic remoteness and limited personal networking connectivity (Crisp *et al.*, 2009). These characteristics and goals are reflected in many of the issues facing the DoD in the areas of mentoring, training, retention, and behavioral problem responses.

The solution advanced in the mentor project was to produce a computer-generated conversational interface available via any system with the appropriate software installed. The user would be presented with a computer screen interface to engage a mentor, a counselor, a tutor or an information source in a conversation-like exchange. Input from the user would be by text entry or through audio speech recognition software. Both should be provided, as high noise areas often preclude the reliable use of voice recognition interfaces. Data has shown that the agents are most compelling and engaging if they were real people (Traum *et al.* 2015) recorded live and exhibiting a demonstrable "screen presence".

Implementing a Conversational Agent with Spontaneous Generation

Initially, for a conversation limited to responses to user questions, a list of germane questions would be drafted. Depending on the topic, this might total several thousand questions. The list creative process requires a broad knowledge of both the questions that should be a part of the computer agent'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 to generate a thorough list of questions. Those include irrelevant questions that the target audience has been known to ask. A previous paper (Nye *et al.*, 2017) discussed this process at greater length. For more directed conversation initiation, consider how humans initiate conversations with a new person and the indicia observed before saying anything. Some specific observations and concomitant analyses of the person might be: age, gender, dress, facial expression, apparent socio-economic-status and demeanor. Then the human conversationalist would do a quick search of their own cultural and personal opening gambits and apply the one that seemed most likely to achieve the desired effect.

The next step is to have the live mentors record a variety of video clips of whatever utterances their germane experience in the service and in their professional duties would indicate as appropriate. These clips would then be stored in a standardized database. When the user provides any input, 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 experience with debriefing test 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 oral history capture (Artstein *et al.*, 2014). The team found that these videotaped mentors are sometimes referred to as Virtual Humans, when they are actually intelligent agents that sequence and present recordings of real people. A substantial corpus of carefully collected questions and answers is still needed that can be answered within the latency norms of live human conversational discourse. Initial openings would require a more extensive list of possible openings or topic-guiding segues, which would typically be much shorter than 90 seconds, *e.g.* "You're looking a little down today; what's up in your life?" or "If you've got a few minutes, let's talk about your professional development.".

In Figure 3, the flow begins at the top left, drops down into the initiation analysis section, and is delivered over to the monitor and speaker, and then proceeds clockwise through the processes. The flowchart symbols are notional only and do not indicate importance, weight or size, *e.g*. the storage symbols are all approximately the same image-on-page size, whereas, operationally they differ by three orders of magnitude in data storage size. The important steps are the analytical functions that could be advanced by quantum computing and deep learning data analyses to enhance neural net training of the evaluation functions, and consequently affect enhanced human-like behaviors.

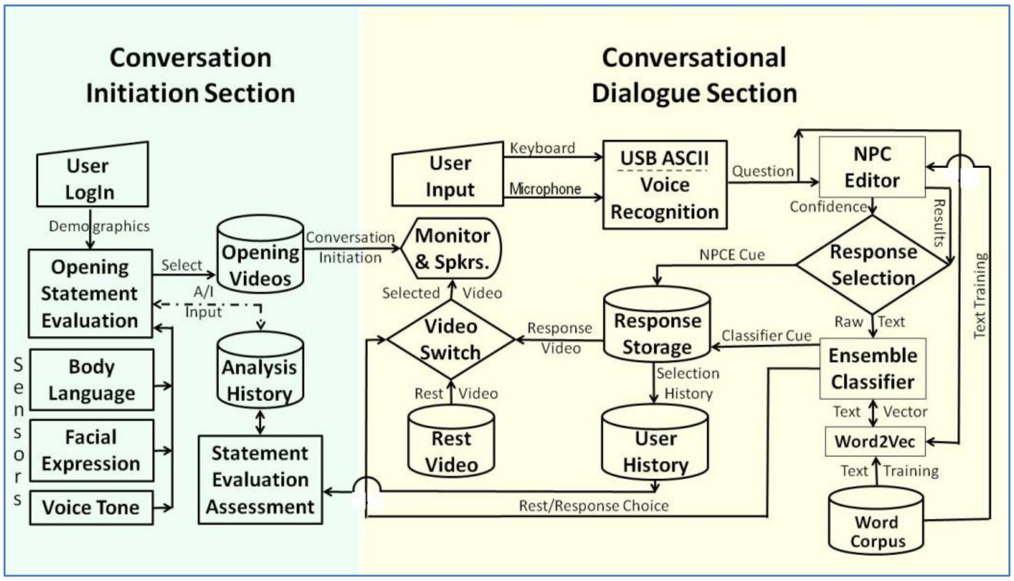

Figure 3. Initiation of Conversation: Notional Flowchart

OPPORTUNITY FOR ENHANCED UTILIZATION

There are basically four potential areas where additional emphasis within the following subsections is warranted:

- Conversationally initiating sessions for guidance, counsel, and stress relief "Hello there,…"
- Articulating acceptable "field expedients," featuring a more free-form topic guidance; "Let's talk about …"
- Helping to bridge the communications gap for personnel assigned around the globe. "Have you heard …"
- Reiterating the national goals and mandates. ever mindful of lives in the balance. "Our new mission is .."

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 decisions are just a search for optimum impact. Any wrong choice would impact efficacy, but may not be fatal to the concept. An 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 project called SimCoach; as shown in Figure 4 (ICT, 2020). The other option, which was discussed above, is the utilization of a thousand or so video clips made by live humans.

Figure 4 - SimCoach for VA Patients who might benefit from conversations (ICT, 2020)

There are two known hurdles to this concept of spontaneous generation. First is studying, evaluating, and collating the appropriate utterance 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, good counseling, tutoring and mentoring is often focused on asking salient questions. NLP is very good at answering direct inquiries, but it will require more research to identify, draft, implement and validate a series of germane probing questions or suggestive scenarios for the agent. These will be needed to get the users to reveal and analyze their issues on their own. Also, a bank of well-formulated suggestions will be required to resolve common issues that are highly likely to arise.

To serve the DoD personnel, it is envisioned that part of the solution will be to urge the participants to set up and schedule regular video conferencing meetings with live and virtual attendees. 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 has yet to be resolved would be the participation of those relying in satellite connections such as remote sites in areas not served by cable, as well as 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 add at least 550 milliseconds to the response time (V-Sat Systems, 2019). That is just enough delay to cause some minor disruption to the normal flow of conversation. Even so, it could be a very effective tool for resolving some of the counter-productive impacts of remote military service.

24 Hour Mentor Access

Using the technologies described above, it is possible to design, implement and field a set of conversational computer agents who are focused on assisting DoD personnel. Operational exigencies create very erratic and taxing schedules. Having 24 hour, seven day-a-week access is a high value aspect of computer agent interfaces. Content should also be designed to admit of random and frequent needs to abort and recontinue sessions later. Some enhancements to current scheduling capabilities would be useful, but the computer's initiating and guiding functions if sensitive to signs of distress or fatigue in the human user, would be a dramatic improvement in the interfaces.

Standardizing Terminology

In an emerging technology such as Virtual Humans, one of the issues that continually hampers progress and the discovery of the optimal path for future research is the lack of consistency in both the terms used and the paucity of actual performance metrics or standards. The term Virtual Humans is often used in the technical vernacular to refer to any human avatar, whether the image is a true animation portraying a human or a computerized delivery of a real human. Additional comparative studies may be beneficial in better showing the relative costs, time constraints, performance, and impacts of both animated and videotaped implementations after actual field use.

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 insight and enable an improved pursuit of goals. This is needed for 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 metrics for various physical and process operations. Just this professional expertise could help the simulators in creating and promulgating a new *lingua franca* for Virtual Human research.

NEW TECHNOLOGIES

The "Law" propounded by Gordon Moore suggested limits to the power of digital computing, *cf.* (Shalf, 2020). One of the alternatives to solving the grand challenges faced by the simulation community that is frequently mentioned is quantum computing. It has wishfully been considered an extension of computational capability since the Nobel Laureate Richard Feynman presented his 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 development of such a machine and those devices still are, as near as can be ascertained, still largely 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 milikelvins) to create a useable quantum effect. (Lucas, 2013)

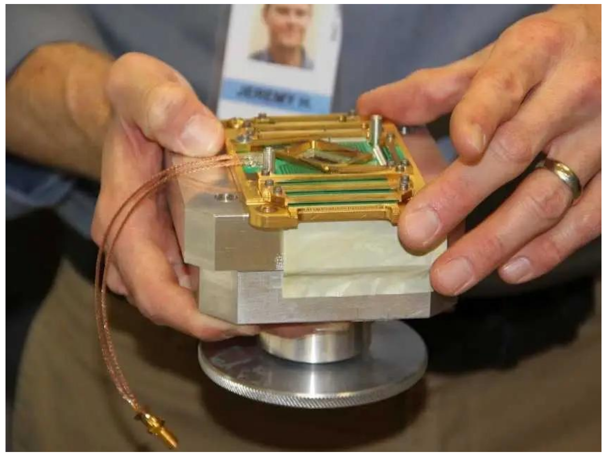

Figure 5 – Quantum processor of 1,025 Qubits before being cooled to 0.015 Kelvin.

This early production version of this quantum annealer has been conceived, designed, produced, and delivered to two sites in California. This technology has been in operation since 2012. In its current configuration, these quantum computers provide approximately 2,000 qubits, but need to be cooled to very close to absolute zero to be stable. Figure 5 gives an idea of scale of the actual processing unit, whereas Figure 1 above 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 for number of qubits is required as some delivered machines have a few 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. One machine is installed in a USC/Lockheed Martin facility in Marina del Rey. There is another in the San Francisco Bay area in a joint Google and NASA project, and one has been ordered by the DoD's High Performance Computing Modernization Program. Others are in varying stages of procurement and delivery. They are special purpose machines, limited to annealing. Nevertheless, annealing is a very useful function for virtual humans and would greatly enhance conversation initiation.

Recently, other authors (Arute,2019) 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. Consider again how many instances of both real and fictional conversation initiations a live human experiences before they are deemed capable of civil behaviors like starting a conversation gracefully.

For several decades the development of high performance computing has resulted in a significant body of tumult with newly introduced technologies. Many examples of contention happened early at the very inception of parallel computing, particularly during its rivalry with sequential computing and with vector computing. The detractors of parallel computing argued that the limits of parallelism fought the proponents who held that parallel computers 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. This evolution of programming design and system engineering, leading to the resultant increase in acceptance should be expected in the future with both quantum computing and deep-learning and subsequent A/I data exploitation techniques. There are physical enhancements already awaiting release and, as in all new technologies, there are even more dramatic advances that should be expected at any moment. Caveat: one should not be too certain any new technology will become main-stream. Many new technologies become obsolescent before they are even introduced, but we can only report upon that which is rationally foreseeable. There may be other hardware advances that will occur instead or clever and innovative software approaches may provide the capabilities to otherwise enable conversation initiation and topic guidance.

Assuming quantum computing may be a mainstay of future computational science, it would be useful to note that the current technology operates by using super-cold processors, see Figure 6. The quantum annealer has been able to demonstrate accepted quantum computing and entanglement. Even if the projected speed-ups are not realized on this design, it is a workable and verifiable quantum computing device on which to experiment the possibilities of its uses in the simulation community, as has been reported in several papers.

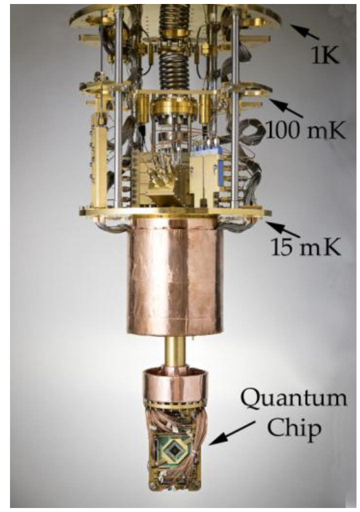

Figure 6. Cooling Mechanism
(D-Wave photo)

This machine is 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 (Christensen, 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 quickly set aside when old technologies improved faster than the new technology could be fully developed and proven. These issues are all part of the Modeling and Simulation (M&S) community's history and provide insights for analyzing the value of new technologies.

Several things need to happen before quantum computing becomes a tool rather than just a proof of concept device. The connectivity paths need to be increased to represent a viable simulated experimental environment. A code base, with 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 implementable change. But the M&S community has known 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 needs for quantum computing acceptance include: more connectivity pathways, more qubits, and, perhaps most critically, improved software in both the applications and the digital-to-quantum interface. The quantum annealer line of machines currently supports C/C++, Python, and MATLAB, three very common research languages. There are both on-line and live tutorial sessions provided to bring new users on-board. Much of the work at USC has been focused on the treatment of large databases, known today as "big data." (Liu, 2020)

Deep Learning

Deep learning is another term that is used to describe an emerging technique in machine learning and artificial intelligence. It is similar to 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 processes only send information that the layer values as beneficial. These successions of processing steps are called 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-world issues, albeit its need for more computing power, which leads back to the advances in quantum computing mentioned above.

Deep learning has achieved state-of-the-art in a wide range of problems, including hand writing recognition, object recognition, classification of 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 finding 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 following paragraphs discuss the theoretical foundation of some of these concepts.

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). This interest faded when it was found that there were limitations on a single Perceptron, which can only capture simple linearly-separable structures. In the 1980's, new enthusiasm arose as researchers constructed 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 computers extant were not powerful enough. The current interest in neural network can be said to have begun in 2006, introducing a novel method to pre-train the deep belief network network evaluation weights, using Restricted Boltzmann Machines (RBM) (Hinton, *et al.*, 2002).

The Boltzmann machine approach is the most amenable to quantum annealing compared to the feed-forward neural networks and recurrent neural networks. The Boltzmann machines are probabilistic neural networks that implicitly define probability distribution over the activation states of the neurons (Figure 7) in a deep learning network. Training a Boltzmann machine calls for repeated sampling from the distribution of activation states. However, sampling for Boltzmann network with loops can be computationally expensive, so another reason for quantum annealing.

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 topologies. 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 (Liu *et al.*, 2020a).

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.

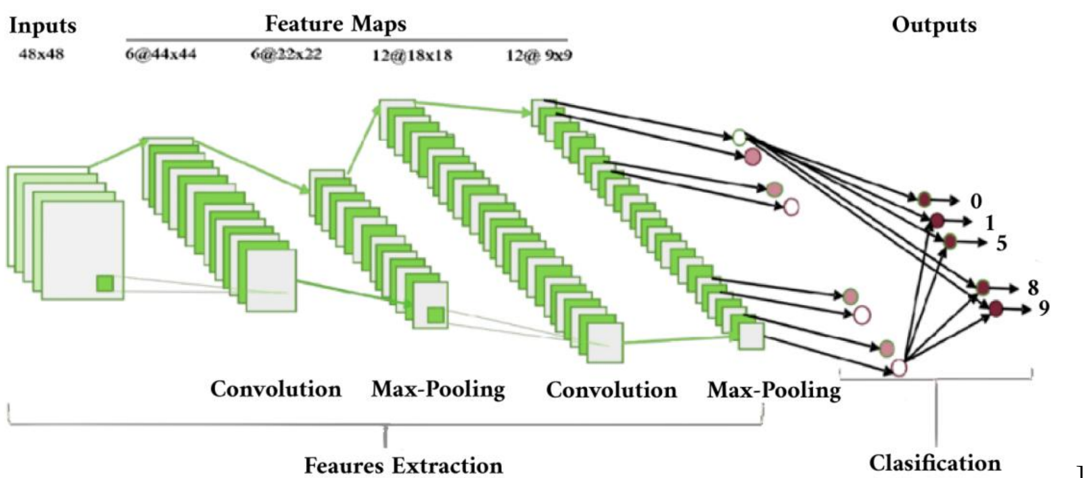

Figure 7. Diagram Showing Successive Layers of Deep Learning Analysis (after Medium, 2019)

To show the order of magnitude of deep learning implementations, the following are illustrative of the scales to be expected. The 2012 ImageNet Large-Scale Visual Recognition Challenge AlexNet winner (Krizhevsky *et al.*, 2010) has over 60,000,000 weights with five convolutional layers, max pool layers, and three fully connected layers. The network was trained on over 15,000,000 images with over 22,000 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. Consider the need for and applicability of the power of a quantum annealer.

Deep learning does have limitations and there are recent research efforts 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 images with busy backgrounds. This ability also induces fragility. With new classes of objects to observe, the model must be retained with many instances of the new objects as well as the previous known objects. In the games area, a deep learning-based program has beaten top ranked human players, but some parameters were manually fixed. Changing from a default parameter by only a few units required extensive retraining. An approach to address this limitation may be to develop less monolithic deep networks, such as 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 to self-generate labeled data.

Work is progressing in the research into the best algorithms to implement on the quantum annealer. One of the figures of merit in this optimization is the number of training epoch required to generate the desired data. As the initiation and guidance of conversational virtual humans is assumed to require truly unfathomably large data sets, this line of inquiry is being closely followed by the M & S community. A recent paper from colleagues at the University Southern California quantum annealer center (Liu *et al*., 2020b) showed a modest, but significant 25% advantage by using the LBM approach over RBMs. These were described above and are more formally explicated in the paper. Further optimizations are anticipated, but the data now available is conducive to making a better-informed choice of conversation opening and guiding training functions (Figure 8). Close coordination with the quantum computing center is vital to early adoption of this technology.

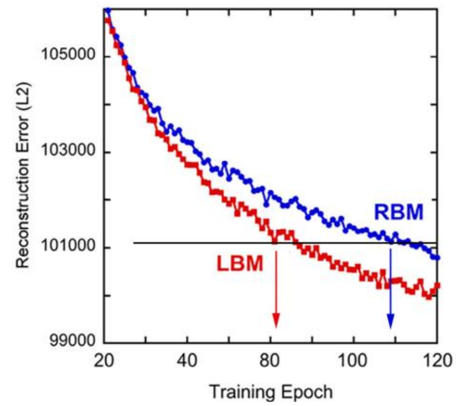

Figure 8. LBM trains faster than RBM (after Liu *et al.,* 2020b)

ETHICAL CONSIDERATIONS

Most humans learn over their lifetimes to recognize those issues that might arouse a negative reaction due to personal sensitivities. If unaware of those sensitivities, they also develop skills at recognizing discomfort or dismay in the person to whom they are speaking. 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, implemented and validated. As one might expect, 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. Using digital computing and standards neural net training to enable a sensitivity avoidance function appears improbable. Two, approaches could be envisioned; they are not mutually exclusive:

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, and physical attack; and

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 useful, 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 confident or therapist. As a way to assess the issues at hand, the authors suggest using a process like 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 at 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.

There are Many 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 users' stated questions. Some of these are team creation coaching, PTSD patient therapy, career counseling, leadership mentoring, and academic tutoring. Another critical areas is suicide prevention. Experience has shown that young leaders are very aware of the need to intervene in suicides and are well-trained to recognize the signs of a suicidal risk. They are, however, still hesitant to approach someone in whom they observe suicidal characteristic behaviors. A virtual human counselor would have no such hesitancies and could be programmed to trigger outside intervention if appropriate.

Emerging Technologies and Capabilities in Simulation and Virtual Humans may be Transformative

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

For future research to be constructive 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 quantified metrics. The existing standards community professionals could be invaluable in this process.

Emerging Technologies have Potentials as Enablers of Virtual Humans with Spontaneous 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 overcoming current barriers and enabling a truly human-like conversation. Many benefits from this implementation would accrue to the warfighters who are time-constrained and often isolated from others.

There are Ethical Issues that arise from a Computer that has Spontaneous generation

There are several ethical issues 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 sensitivity matrices and a quick live interview with another person can provide a good sense of topics that can be destructive to that individual. There may be A/I approaches and sensor capabilities not yet implemented that would address these issues, but this paper in no way represents that 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 Another issue is whether such an A/I programs might develop computational pathologies such as the fictional HAL.

Perhaps, some researchers will raise the issue that there are some humans who are similarly insensitive and it is not outside of the realm of possibility that humans will be less offended by a machine mistake that is apparently not being malicious than by a human mistake that seems to have been intended to injure. The problems are real but the promise of new hardware and more sophisticated A/I techniques creates a hopeful M&S future 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. We would be remiss were we not to recognize the contributions to our paper by Dr. M. Beth Pettitt, who made many helpful suggestions and offered constructive, much appreciated advice.

REFERENCES

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, Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., ... & Burkett, B. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, *574*(7779), 505-510.

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 (2020). *Prototypes, SimCoach, Images.* Institute for Creative Technologies, University of Southern California. Retrieved on 04 (2019 from: https://ict.usc.edu/prototypes/simcoach/ .

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

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399.

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.

Liu, J., Lucas, R., Kalia, R., Nakano, A., Spedalieri, F., & Yao, K. (2020a). Dynamic Topology Reconfiguration of Boltzmann Machines on Quantum Annealers. *Computational Materials Science.* Volume 173

Liu, J., Mohan, A., Kalia, R. K., Nakano, A., Nomura, K. I., Vashishta, P., & Yao, K. T. (2020b). Boltzmann machine modeling of layered MoS2 synthesis on a quantum annealer. *Computational Materials Science*, 173,

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.

Shalf, J. (2020). The future of computing beyond Moore's law. *Philosophical Transactions of the Royal Society A*, *378*(2166), 20190061.

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