On the need for Artificial Intelligence and Advanced Test and Evaluation Methods for Space Exploration.

D. H. Scheidt, C. A. Hibbitts, M. H. Chen, D. L. Bekker, L. J. Paxton, John Hopkins University Applied Physics Laboratory 11100 Johns Hopkins Road, Laurel, MD. david.scheidt@jhuapl.edu.

Introduction: The romantic view of artificial intelligence (AI) is the study of thinking machines, specifically machines that are equivalent to, and think like, humans. Robots embodied with AI have been a romantic element of space exploration through influential films including Fred Wilcox's Forbidden Planet (1956) and, using a much darker tone, in Stanley Kubrick's 2001 a Space Odyssey (1968). While it is true that many of the reasoning methods being pursued by the AI community are anthropomorphic, artificial intelligence has a more pragmatic side, which is the study of algorithms that allow machines to produce answers to complex, multifaceted problems. To understand the difference between an artificial intelligence system and "unintelligent" computer systems, we should first consider how software is normally developed. The software development process begins with the specification of requirements, followed by design where software engineers rigorously determine the appropriate machine response to each and every combination of inputs that might be encountered. Sometimes the appropriate responses to system stimuli are defined mathematically by using control theory, and sometimes software engineers use brute force by exhaustively enumerating all possible stimulus-response pairings. Traditional software engineering works well for the vast majority of software systems; however, when system requirements demand that machines make decisions in a world that is too complex or too uncertain for engineers to solve during software design, AI is required.

The design process for AI, necessitated by the need to have machines address intractable problems during execution, is a radically different approach to developing control software. When developing artificial intelligence, software engineers do not program explicit responses to situations encountered by the machine, but rather write software that provides machines with an ability to solve problems during program execution, allowing the AI software to produce a decision that was not explicitly encoded in the software. All major forms of AI research, which includes deductive, inductive and abductive reasoning, search-based algorithms, machine learning and neural networks, exhibit the property that the machine's answers to specific problems are not explicitly encoded within the AI software; rather methods for devising answers to the problem are encoded. This subtle distinction between AI and unintelligent controller provides the power and promise of AI and AI's greatest risk. The promise of AI is an ability to solve important problems that cannot be solved through traditional programming means. The risk of AI is the potential to produce unvetted responses to situations that run counter to the designer's wishes.

On the use of Artificial Intelligence: When is it useful to have a machine use AI to make a decision? After all, after millions of years of evolution and roughly 10,000 years of civilization, humans are (usually) quite good at making decisions in complex, uncertain environments. Through our research in AI-enabled systems, Johns Hopkins University's Applied Physics laboratory has identified three general use cases for AI: first, for some tasks AI is more cost effective than humans; second, AI is better suited than humans at solving some, but not all, problems; third, AI allows us to develop machines that are capable of responding faster than when a human is in the decision loop. [1]. Each of these three strengths is potentially relevant to future NASA mission architectures.

AI's ability to allow machines to respond more rapidly than when a human is inserted into the decision chain is significant for NASA because communications between Earth and extraterrestrial spacecraft or rovers dramatically lengthen operational response times. The speed at which AI-enabled machines can react has several benefits. First, it allows diagnosing and repairing faults within complex systems to prevent measurement (instrument) or mission failure through timely diagnosis and management of unexpected anomalies. Secondly, it allows responding to unexpected, ephemeral science opportunities and the exploration of high-temporal phenomena. Finally, AI can accelerate the cadence at which science is conducted.

The use of AI to enable science by observing the pace of rapidly evolving phenomena was demonstrated spectacularly with the Jet Propulsion Laboratory's 2006 discovery of dust devils and clouds on Mars. [2] Both JPL [3] and APL [4] have demonstrated that AI can accelerate the pace of science by more effectively coordinating and utilizing space-based sensors. As an example, the intelligent fault management system *Livingstone* was developed by NASA AMES and flown on NASA's Deep Space One mission in 1998. [5]

The current risks of AI: Today, AI is immature and requires further development to reach its potential. For instance, the AI algorithms that detected the dust devils could not have identified whether the Martian weather represented a threat to the rover. Also, AI can not yet use instrument input to determine what, where, and how to autonomously make the next science measurement. An equally important factor limiting AI's deployment is that we lack the methodology and technology to effectively test AI. We have explored emerging requirements for testing aspects of terrestrial autonomous systems [6] and these needs are reflected in space-based AI testing.

Path to AI in 2050, Testing: The greatest risk associated with AI is the risk of undesirable detrimental, consequences from decisions emerging from unintended combinations of legitimate rules and/or patterns. Traditionally, system test and evaluation requirements define the desired system response for all anticipated operating conditions. Requirements-driven design is problematic for AI-enabled systems because the size of the condition-response matrix is intractably large, preventing test engineers from fully enumerating system requirements. In addition, autonomous systems, by their very nature, determine responses at run time, a control technique that is itself antithetical to an a priori system response matrix. The first challenge with testing AI-enabled systems is: how can AI performance be measured? The National Institute of Standards and Technology (NIST) autonomy levels for unmanned systems ALFUS [7] codify the degree to which a system is, or is not, autonomous; standard metrics on autonomous system performance are not codified. It is clear that AI-enabled system metrics must include measurements of the decision made by the system; and a subset of any autonomous system metrics should be derived from mission performance. Relying solely on mission-based metrics for AI-enabled systems can be problematic as the decisions made by the AI may have unintended consequences that are unrelated to mission objectives, yet very detrimental to the larger objectives of the operator. How can the test team provide performance assurances given that it is impossible to test all circumstances?

Because AI-enabled system performance is dependent upon the complex interactions between the AI-enabled system and artifacts in the real world, it is vital that tests be conducted in environments that mimic the complex interactions between actors in the real world and the AI-enabled system. This need presents us with our final challenge to testing AI-enabled systems: how can we provide a complex, *interactive* test environment that, from the point of view of the AI, mirrors the diverse interactions experienced in real-world operations.

Path to AI in 2050, Technology Advancements and Distributed Systems: Complex algorithms are often synonymous with power hungry electronics. As technology advances, the realization of low power computing becomes viable in space. An example of this new technology is neuromorphic computing. Influenced by how the mammalian brain processes and communicates data, neuromorphic computing is a new class of non von Neumann machines that is showing excellent performance in neural network applications. One such system, the IBM TrueNorth, has packed a million spiking neurons into a chip consuming on average less than 100 mW [8] and recent work has been demonstrated on deep learning datasets [9].

The combination of distributed autonomous systems [10] with low power yet high computing resources provides a bridge to fully autonomous mission concepts. Imagine a swarm of dispensable autonomous explorers that can intelligently investigate large areas and provide reconnaissance and surface exploration prior to the main spacecraft arrival. Mission success is robust against individual unit failure, as the aggregate is more capable than the sum of the parts.

Advances in algorithms, testing, sensor technologies and packaging is beginning to make possible the concept of complete autonomous space systems. The realization of fully autonomous systems enables new solutions for planetary exploration.

References:

[1] D. Scheidt, Unmanned Air Vehicle Command and Control, Handbook of Unmanned Air Vehicles, Springer-Verlag, 2014. [2] A. Castano et al. "Automatic detection of dust devils and clouds at Mars," Machine Vision and Applications, October 2008, vol. 19, No. 5-6, pp. 467-482. [3] A.G. Davies et al. (2005) Monitoring Active Volcanism with the Autonomous Sciencecraft Experiment (ASE). Remote Sensing of Environment, Vol. 101, Issue 4, pp. 427-446.[4] T.H. Choo et al. "SciBox: A Software Library for Rapid Development of Science Operation Simulation, Planning, and Command Tools", Johns Hopkins APL Technical Digest, Vol 25, No 2 (2004). [5] N. Muscettola et al. 1998. Remote agent: to boldly go where no AI system has gone before. Artificial Intelligence 103:5–47. [6] B. D'Amico, R. Lutz, D. Scheidt, Testing of Autonomy in Complex, Interactive Environments, International Test and Evaluation Association Journal, December, 2014. [7] Huang, H., "Autonomy Levels for Unmanned Systems (ALFUS) – Version 1.1", National Institute of Standards and Technology (NIST), September 2004. [8] http://science.sciencemag.org/content/345/6197/668.full [9] http://www.pnas.org/content/113/41/11441.full [10] R. Chalmers, D. Scheidt, et al.,"Cooperating unmanned vehicles," in AIAA 1st Intelligent Systems Technical Conference, 2004, pp. 1–8.