Not for Public Release until Approved by the
Senate Armed Services Committee

Statement of Dr. Peter Highnam
Deputy Director, Defense Advanced Research Projects Agency (DARPA)

Department of Defense Artificial Intelligence Initiatives

Before the
Emerging Threats and Capabilities Subcommittee
Committee on Armed Services
United States Senate

March 12, 2019

DARPA's Seminal Role in the Field of Artificial Intelligence

Seventy years ago, when early electronic computers ran on vacuum tubes and filled entire rooms, researchers already were striving to enable machines to think as people do. Only a few years after its start in 1958, DARPA began playing a central role in realizing this ambition by laying some of the groundwork for the field of artificial intelligence (AI). Early work in AI emphasized handcrafted knowledge, and computer scientists constructed so-called expert systems that captured the rules that the system could then apply to situations of interest. Such "first wave" AI technologies were quite successful – tax preparation software is a good example of an expert system – but the need to handcraft rules is costly and time-consuming and therefore limits the applicability of rules-based AI technologies.

The past few years have seen an explosion of interest in a sub-field of AI dubbed "machine learning" that applies statistical and probabilistic methods to large data sets to create generalized representations that can be applied to future samples. Foremost among these approaches are deep learning (artificial) neural networks trained to perform a variety of classification and prediction tasks when adequate historical data is available. Therein lies the rub, however, as the task of collecting, labelling, and vetting data on which to train such "second wave" AI techniques is prohibitively costly and time-consuming.

DARPA envisions a future in which machines are more than just tools that execute human-programmed rules or generalize from human-curated data sets. Rather, the machines DARPA envisions will function more as colleagues than as tools. Towards this end, DARPA is focusing its investments on a "third wave" of AI technologies that brings forth machines that can reason in context. Incorporating these technologies in military systems that collaborate with warfighters will facilitate better decisions in complex, time-critical, battlefield environments; enable a shared understanding of massive, incomplete, and contradictory information; and empower unmanned systems to perform critical missions safely and with high degrees of autonomy.

Today, DARPA is funding more than 24 programs exploring ways to advance the state of the art in AI, pushing beyond second wave machine learning towards contextual reasoning capabilities. This is in addition to more than 55 active programs that are leveraging machine learning or AI technologies in some capacity–from managing the electromagnetic spectrum to detecting and patching cyber vulnerabilities.

This level of investment has been years in the making and will define scientific and technical exploration, as well as resulting military capabilities, for decades to come.

Current Programs

DARPA's **Lifelong Learning Machines (L2M)** program is exploring ways to enable machines to learn while doing without catastrophic forgetting. Such a capability would enable systems to improve on the fly, recover from surprises, and keep them from drifting out of sync with the world.

First announced in 2017, L2M research teams are developing complete systems and their components, as well as exploring learning mechanisms in biological organisms with the goal of translating them into computational processes. Discoveries in both technical areas are expected to generate new methodologies that will allow AI systems to learn and improve during tasks, apply previous skills and knowledge to new situations, incorporate innate system limits, and enhance safety in automated assignments. While the program is still in its early stages, L2M researchers already have identified and solved challenges associated with building and training a self-reproducing neural network.

DARPA is also currently running a program called **Explainable AI** or **XAI** to develop new machine-learning architectures that can produce accurate explanations of their decisions in a form that makes sense to humans. As AI algorithms become more widely used, reasonable self-explanation will help users understand how these systems work, and how much to trust them in various situations. XAI specifically aims to create a suite of machine learning techniques that produce explainable models – while maintaining a high level of prediction accuracy so human users understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners. Enabling computing systems in this manner is critical because sensor, information, and communication systems generate data at rates beyond what humans can assimilate, understand, and act upon.

The real breakthrough for artificial intelligence, however, will not come until researchers figure out a way for machines to learn or otherwise acquire common sense. Without common sense, AI systems will be powerful but limited tools that require human inputs to function. With common sense, an AI could become a partner in problem solving. Common sense knowledge is so pervasive in our lives that it can be hard to recognize. For example, in conflict and warzone situations, people tend to make snap decisions about the cause of the problem and ignore evidence that does not support their point of view. To act as a valued partner in such situations, the AI system will need sufficient common sense to know when to speak and what to say, which will require that it have a good idea of what each

person knows. Interrupting to state the obvious would quickly result in its deactivation, particularly under stressful conditions.

In order to find answers to the common sense problem, DARPA launched in October of last year the **Machine Common Sense (MCS)** program, which will explore recent advances in cognitive understanding, natural language processing, deep learning, and other areas of AI research. MCS is pursuing two approaches for developing and evaluating different machine common sense services. The first approach seeks to create computational models that learn from experience and mimic the core domains of cognition as defined by developmental psychology. This includes the domains of objects (intuitive physics), places (spatial navigation), and agents (intentional actors). Researchers will develop systems that think and learn as humans do in the very early stages of development, leveraging advances in the field of cognitive development to provide empirical and theoretical guidance.

To assess the progress and success of the first strategy's computational models, researchers will explore developmental psychology research studies and literature to create evaluation criteria. DARPA will use the resulting set of cognitive development milestones to determine how well the models are able to learn against three levels of performance: prediction/expectation, experience learning, and problem solving.

The second MCS approach will construct a common sense knowledge repository capable of answering natural language and image-based queries about common sense phenomena by reading from the Web. DARPA expects that researchers will use a combination of manual construction, information extraction, machine learning, crowdsourcing techniques, and other computational approaches to develop the repository. The resulting capability will be measured against the Allen Institute for Artificial Intelligence (AI2) Common sense benchmark tests, which are constructed through an extensive crowdsourcing process to represent and measure the broad common sense knowledge of an average adult.

AI Next Campaign

DARPA announced in September 2018, a multi-year investment of more than $2 billion in new and existing programs called the "AI Next" campaign. Campaign key areas include providing robust foundations for second wave technologies, aggressively applying second wave AI technologies into appropriate systems, and exploring and creating third wave AI science and technologies.

AI Next builds on DARPA‘s five decades of AI technology creation to define and to shape the future, always with the Department’s hardest problems in mind. Accordingly, DARPA will create powerful capabilities for the DoD by attending specifically to the following areas:

New Capabilities: AI technologies are applied routinely to enable DARPA R&D projects, including more than 60 ongoing programs, such as the Electronic Resurgence Initiative, and other programs related to real-time analysis of sophisticated cyber attacks, detection of fraudulent imagery, construction of dynamic kill-chains for all-domain warfare, human language technologies, multi-modality automatic target recognition, biomedical advances, and control of prosthetic limbs. DARPA will advance AI technologies to enable automation of critical Department business processes. One such process is the lengthy accreditation of software systems prior to operational deployment. Automating this accreditation process with known AI and other technologies now appears possible.

Robust AI: AI technologies have demonstrated great value to missions as diverse as space-based imagery analysis, cyber attack warning, supply chain logistics and analysis of microbiologic systems. At the same time, the failure modes of AI technologies are poorly understood. DARPA is working to address this shortfall, with focused R&D, both analytic and empirical. DARPA’s success is essential for the Department to deploy AI technologies, particularly to the tactical edge, where reliable performance is required.

Adversarial AI: The most powerful AI tool today is machine learning. Machine learning systems are easily duped by changes to inputs that would never fool a human. The data used to train such systems can be corrupted, and the software itself is vulnerable to cyber attack. These areas, and more, must be addressed at scale as more AI-enabled systems are operationally deployed.

High Performance AI: Computer performance increases over the last decade have enabled the success of machine learning, in combination with large data sets, and software libraries. More performance at lower electrical power is essential to allow both data center and tactical deployments. DARPA has demonstrated analog processing of AI algorithms with 1000 times speedup and 1000 times power efficiency over state-of-the-art digital processors, and is researching AI-specific hardware designs. DARPA is also attacking the current inefficiency of machine learning, by researching methods to drastically reduce requirements for labeled training data.

Next Generation AI: The machine learning algorithms that enable face recognition and self-driving vehicles were invented over 20 years ago. DARPA has taken the lead in pioneering research to

develop the next generation of AI algorithms, which will transform computers from tools into problem-solving partners. DARPA research aims to enable AI systems to explain their actions, and to acquire and reason with common sense knowledge. DARPA R&D produced the first AI successes, such as expert systems and search, and more recently has advanced machine learning tools and hardware.

In addition to new and ongoing DARPA research, a key component of the AI Next campaign will be DARPA's **Artificial Intelligence Exploration** (AIE) program, first announced in July 2018. AIE constitutes a series of high-risk, high payoff projects where researchers work to establish the feasibility of new AI concepts within 18 months of award. Leveraging streamlined contracting procedures and funding mechanisms enables these efforts to move from proposal to project kick-off within 3 months of an opportunity announcement.

Conclusion

Over its 60-year history, DARPA has made significant investments in the creation and advancement of artificial intelligence technologies that have produced game-changing capabilities for the Department of Defense and beyond. DARPA's AI Next effort is simply a continuing part of its historic investment in the exploration and advancement of AI technologies.

Current R&D investment around the world is largely focused on second wave AI or machine learning, which is very good in finding patterns in voice and imagery and has many commercial applications. The difference is, in the United States, DARPA is aggressively pursuing programs that will make second wave AI more robust for defense and security applications, all while helping realize the third wave of AI, or contextual reasoning. DARPA has unique access to the United States' world-class science and technology community, comprised of leading universities, government labs, and industry partners – this mix cannot be found or replicated anywhere else in the world. Marshalling those unique resources, the Agency's third wave research efforts will forge new theories and methods that will make it possible for machines to adapt contextually to changing situations, advancing computers from tools to true collaborative partners. Going forward, the agency will be fearless about exploring these new technologies and their capabilities – DARPA's core function – pushing critical frontiers ahead of our nation's adversaries.