ARL-SR-0397 • May 2018

US Army Research Laboratory

Reasoning about Complex and Uncertain Worlds in Physical, Social, and Virtual Realms: A Report of the Army Science Planning and Strategy Meetings Held in Fiscal Year 2018 at the US Army Research Laboratory

by Bryan Glaz, Brett Piekarski, Brian Sadler, Lisa Troyer, Peter Khooshabehadeh, Brian J Henz, Purush Iyer, Tien Pham, Ernest Chin, Jaroslaw Knap, Brent Kraczek, MaryAnn Fields, Brian Jalaian, and Alexander Kott

ARL-SR-0397 ● MAY 2018

US Army Research Laboratory

Reasoning about Complex and Uncertain Worlds in Physical, Social, and Virtual Realms: A Report of the Army Science Planning and Strategy Meetings Held in Fiscal Year 2018 at the US Army Research Laboratory

Bryan Glaz, Brett Piekarski, Brian Sadler, and MaryAnn Fields
Vehicle Technology Directorate, ARL

Lisa Troyer and Purush Iyer
Army Research Office, Durham, NC

Peter Khooshabehadeh
Human Research and Engineering Directorate, ARL

Brian J Henz, Tien Pham, Ernest Chin, Jaroslaw Knap, Brent Kraczek, and Brian Jalaian
Computational and Information Sciences Directorate, ARL

Alexander Kott
Office of the Director, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED (From - To)**
May 2018	Special Report	September 2017–May 2018

4. TITLE AND SUBTITLE Reasoning about Complex and Uncertain Worlds in Physical, Social, and Virtual Realms: A Report of the Army Science Planning and Strategy Meetings Held in Fiscal Year 2018 at the US Army Research Laboratory	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Bryan Glaz, Brett Piekarski, Brian Sadler, Lisa Troyer, Peter Khooshabehadeh, Brian J Henz, Purush Iyer, Tien Pham, Ernest Chin, Jaroslaw Knap, Brent Kraczek, MaryAnn Fields, Brian Jalaian, and Alexander Kott	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory ATTN: RDRL-VTV Aberdeen Proving Ground, MD 21005-5066	**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-SR-0397
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Since 2013, responding to a request from the Assistant Secretary of the Army for Acquisition, Logistics, and Technology, the US Army Research Laboratory has been conducting an annual series of meetings. The intent of these meetings is to explore novel scientific opportunities that may lead to providing the Army with an advantage in future conflicts. The temporal scope of these explorations is strategic in nature, with time horizon being on the order of 20 to 30 years. The meetings pay particular attention to the identification of research gaps and barriers that may hinder the achievement of the potential novel capabilities; and of possible approaches of overcoming these gaps and barriers. These meetings are called the Army Science Planning and Strategy Meetings. This report covers the findings and recommendations developed in 4 meetings held in the first half of fiscal year 2018, covering the following topics: Reasoning About and Interacting with the Physical World, Learning and Reasoning in Complex Data Environment, Sensing and Modeling Social Dynamics, and Uncertainty Quantification.

15. SUBJECT TERMS

artificial intelligence, machine learning, physics modeling, social sensing, cultural dynamics, trust in AI, human-agent teaming, training data, uncertainty quantification

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Alexander Kott
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	54	**19b. TELEPHONE NUMBER (Include area code)**
Unclassified	Unclassified	Unclassified			(301) 394-1507

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

Acknowledgments

The authors gratefully acknowledge the encouragement of the Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA(ALT)), the support of the US Army Research Laboratory (ARL) Director's Office, and the many participants, all of whom contributed to a thought-provoking and highly productive series of meetings.

The authors thank Jessica Schultheis for organizing and editing the manuscript of this report.

The authors are also indebted to numerous individuals who helped to structure the foci of the meetings, identify and recruit the participating experts, and enable the meetings with their logistical and administrative expertise: Joseph Myers (Army Research Office [ARO]), Bruce West (ARO), Stuart Young (Computational and Information Sciences Directorate [CISD]), Ethan Stump (CISD), David Han (CISD), Jon Fink (CISD), Samuel Stanton (ARO), Jason Pusey (Vehicle Technology Directorate [VTD]), Chad Kessens (VTD), Purush Iyer (ARO), Latasha Solomon (ARL/CISD), David Lind, Ed Habtour, Terrance O'Regan, and Tamara Christenson.

1. Introduction

Author: Alexander Kott

Since 2013, responding to a request from the Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA[ALT]), the US Army Research Laboratory (ARL) has been conducting an annual series of meetings. The intent of these meetings is to explore novel scientific opportunities that may lead to providing the Army with an advantage in future conflicts. The temporal scope of these explorations is strategic in nature, with time horizon being on the order of 20 to 30 years. The meetings pay particular attention to the identification of research gaps and barriers that may hinder the achievement of the potential novel capabilities; and of possible approaches of overcoming these gaps and barriers. These meetings are called the Army Science Planning and Strategy Meetings (ASPSMs).

The annual series of ASPSMs have been highly influential in shaping the Army investments in science and technology. Numerous research efforts—in-house, collaborative, and extramural—have been initiated or revectored based on the insights developed within the ASPSMs. This report covers the findings and recommendations developed in 4 meetings held in the first half of fiscal year (FY) 2018.

As discussed in previous ASPSM reports, modern warfare unfolds in 3 realms:

- physical, the domain of activities defined in space and time by the laws of physics;
- societal (cultural and human), the domain of activities defined by the interaction of people and societies; and
- informational, the domain of activities defined by thought and perception.

Accordingly, 3 of the 4 meetings covered in this report focused directly on each of the 3 domains, as follows:

- The meeting titled "Reasoning About and Interacting with the Physical World" focused on the physical realm;
- The meeting "Sensing and Modeling Social Dynamics" focused on societal realm; and
- The meeting "Learning and Reasoning in Complex Data Environment" focused on the informational realm.

The fourth meeting, “Uncertainty Quantification”, explored a meta-challenge that transcends all 3 realms: quantification of uncertainty that is inherent in our ability to understand and reason about the phenomena in all 3 realms.

The meetings in this series were conducted in 1 or 2 days, either at the ARL headquarters in Adelphi, Maryland, or at the ARL-West location in Playa Vista, California. Each meeting brought together 30–40 topic experts, drawn by invitation only, primarily from academia and industry. Each meeting addressed such questions as, What capability can this area of science or technology deliver to the military 25 years from now? What technical hurdles exist that limit our ability to realize this capability? What are the fundamental scientific questions that underlie the area? What research does the Army need to support now to overcome these hurdles and enable the desired capability? The participants of the meetings were encouraged to offer a variety of perspectives, with emphasis on a long-term, broad view of the specific area and its trends.

1.1 Highlights of Findings for Each Meeting

The meeting titled “Reasoning About and Interacting with the Physical World” found, inter alia, the following:

- The Army faces unique challenges in ground-based operations, which hinder and slow the adoption of autonomy into ground platforms. Navigation is dramatically more challenging in ground environments than in obstacle-free air and sea environments. Current and projected commercial technology will not reliably extend into new Army-relevant tactical environments and operations in dense urban areas, subterranean, and jungle environments.

- The science of robotic physical interaction is not well developed, except in static fixed environments with infrastructure. Physical reasoning and the laws of physics are not yet captured in artificial intelligence (AI) suitable for sufficiently general purposes, and common-sense reasoning, going beyond physical reasoning, is not yet possible.

- Online learning is limited to short time durations and is not well integrated with physics-based models or simulations. Physics-based simulations are highly scenario specific and not easily generalized or scaled. There is a lack of experiments and benchmarks in real-world environments.

The meeting titled "Sensing and Modeling Social Dynamics" found the following:

- Social phenomena like violent protests, large-scale social influence, violent extremism, coalition formation, mass movement fragmentation, and sudden population (im)migration pose challenges for the Warfighter in the Army operational environment. However, current scientific foundations are lacking with respect to the capability to 1) objectively measure the transitions of large-scale collectives from one state (e.g., peaceful protest) to another state (e.g., violent mass protest); and 2) causally and predictively model these transitions across cultures.

- New sensing capabilities such as geospatial sensors, biometrics, tracking of utility use, and multimedia data are enabling more objective ways to measure social action compared to traditional approaches such as observation and survey methods, which are often fraught with bias. There is an ever-growing array of opportunities to sense shifts at micro (individual and small-group) levels and macro (large-scale collective and population) levels. To date, however, there have been only relatively coarse attempts to integrate multiple types of sensors owing to vast temporal and spatial differences in their capabilities.

- Complex modeling approaches and new statistical techniques are being developed that more accurately represent the distributions of large-scale action and the often punctuated shifts from one collective state to another. There is a growing recognition that collective-level action is often not normally distributed; and it often abruptly shifts from one state to another. Often it is the rare events/actions in a distribution of social dynamics that are of both the greatest interest and the greatest risk.

- Research with greatest promise is that which moves away from traditional social science statistical approaches toward modeling approaches drawn from physics, natural sciences, and computer sciences that are emerging to capture complex crowd behavior. Yet, validation seemed missing from the normal research cycle.

The key findings of the "Learning and Reasoning in Complex Data Environment" meeting include the following:

- "Trust in AI" is a major technical gap identified in both focus areas explored in the meeting. This includes a number of open questions for research, including validation and verification (V&V) of algorithms (particularly adaptive or online learning), providing realistic expectations for AI, and learning individual preferences for cooperation and risk analysis. Loss of trust in deployed AI will result in resistance to use in the field. Closely related to the question of trust is the continued validation of deployed adaptive AI algorithms on learned tasks and verification of expected behavior.

- A significant number of existing data sets are available within Army organizations; these derive from field tests, laboratory tests, and deployed systems. A majority of the Army attendees at the ASPSMs provided lists of data sets that exist within their respective facilities. The challenge is that these data sets are not collated in central locations, often contain unlabeled data, are in many different formats, and are of unknown or limited quality and value. Whereas academia and industry rely heavily upon labeled, voluminous, clean data sets for learning, the Army faces a real gap in this area.

- A major technical gap is seen in deploying AI at the edge—at the tactical locations where AI, computing, and communications have to reside. Sensors are proliferating on the battlefield while the ability of the network to transport this data to the point of need is not improving commensurately. Although always recognized as a challenge, the meeting re-emphasized the extent of the issue and the current growth in demand for deploying AI as close to the point of data collection as possible.

The meeting titled "Uncertainty Quantification" yielded the following findings:

- Incorporation of uncertainty quantification (UQ) practices into data collection will allow researchers to track the quality of these data (e.g., fidelity, resolution, limitations) and update the confidence/trust on decision making with additional data; UQ is necessary to address confidence in the machine learning from such data. UQ guides where additional data is necessary to increase the output certainty for artificial decision or prediction. It is an approach to detect anomalies in the data, possibly deceptive in nature, and also to track rare events that machine learning from transferred knowledge (prior trained data) would most likely neglect.

- Tools have emerged that are used to evaluate specific designs, courses of action, and public danger. These include conditional value-at-risk (CVaR, also known as superquantile risk), distributionally robust (data-driven) optimization, risk quadrangle (risk-informed modeling), and capacity-based UQ. These methods are particularly well suited to problems for which a single objective can be readily quantified, the quantity of interest (QoI). Adaptation of these methodologies for anomaly detection leading to adversarial mitigation schemes and predicting low-probability/high-cost events is not trivial but is very promising.

- Recognizing the opportunities from physics-aware modeling-to-machine learning-to-decisions, a holistic framework for integrating uncertainty with risk that balances parameters such as information, knowledge, resources, and trust is necessary. This demands quantifiable metrics for uncertainty and risk representations that complement models of human–machine interaction, including trust. Continuous learning and evaluation of this framework's uncertainty are necessary to account for information degradation, malicious information insertion, dense versus sparse information, and processing tradeoffs (overload/underload).

1.2 Summary of Recommendations

Based on the output of the meetings, the authors of this report developed the following recommendations:

- A long-term integrated program of research should be undertaken to support a physical reasoning architecture that incorporates and balances the physical world and knowledge world constructs; learning that complements modeling and is incorporated into physical platforms and is self-guided and curious; abstractions that are scalable, can be shared across agents, and facilitate fast reasoning and action; general methods for characterizing families of affordances; and physical reasoning that is linked with the wisdom in the body. More on this in Section 2.

- There is a need to develop a truly multidisciplinary integrated research program that equally engages computational sciences, social scientists, and computational scientists to advance analytics, but also avoid reinventing the wheel when it comes to social science theory and research. Its foci should include research on how different sensors at both micro and macro levels work cross-culturally and within technologies such as the "Internet of Things", how different sensors are used, what factors drive users to different social sensing platforms, and on cultural interpretations of them, as well as the nature of tipping points and how they relate to institutional and cultural dynamics. More on this in Section 3.

- An Army-led research program is needed that would provide a quantifiable V&V of adaptive, adversarially robust AI algorithms and communicate the AI-reasoned solutions to a Soldier with trust and understanding. The close interaction of humans and AI agents in stressful environments is uniquely characteristic of military application of AI. A focused effort is needed to capture and maintain data from across the Army to develop AI algorithms and software. This data must be in an accessible location for Army collaboration and development efforts. Fundamental questions of suitability of data for training must be addressed; industry and much of the academic AI development are dependent upon labeled, voluminous, clean data sets for learning. Success in this effort will help bring in external collaborators through data sharing; generate synthetic data with transfer learning to the real world; and provide a central location for Army researchers to share common data and environments. More on this in Section 4.

- Both opportunities and needs exist for initiating a substantial portfolio of efforts in UQ that focus on integration of Bayesian frameworks and generative adversarial networks for anomaly detection and nonlinear autoregressive schemes with dynamical systems theory for rejection of deception; combining risk measures and uncertainty sources by including uncertainty in the objective functions and constrains; comparing stochastic outputs using objective measures (e.g., superquantiles). More on this in Section 5.

2. Reasoning About and Interacting with the Physical World

Authors: Bryan Glaz, Brett Piekarski, and Brian M Sadler

2.1 Vision

The Army seeks to develop and deploy intelligent physical agents that will team with and assist Soldiers, manned combat vehicles, and other intelligent agents in complex operational conditions across the multidomain battlespace. These future systems will not only need to efficiently and autonomously move through diverse environments (dense urban, subterranean, jungle, etc.), they will be required to perform complex and adaptive missions in the face of contested and adversarial technologies. This teaming with Soldiers and next-generation ground and air systems will require machine intelligence that goes well beyond current standoff perception methods, state-awareness techniques, and emerging autonomous systems technologies, such as driverless cars. Intelligent agents will need to interact with the physical world to reason about and perform complex actions such as moving obstacles to aid navigation, manipulating objects to accomplish mission objectives such as a breach operation, physically interrogating structures and dynamic systems to improve perception, and physically interacting with Soldiers or other agents to achieve complex multi–degree of freedom collaborative actions such as jointly carrying an object through complex terrain.

2.2 Objective and Scope

The focus of this ASPSM was to explore and discuss the underpinning science needed for future intelligent systems that can reason about, interact with, and manipulate the physical world around them to achieve this vision for complex military-relevant actions in the future multidomain battlespace.

The meeting brought together a unique mix of research leaders in AI and reasoning, autonomy, learning, robotics, and mechanics. The meeting was organized around 2 focus areas:

1) *Reasoning about the physical world:* Advances in AI, sensing, and perception are leading to enhanced understanding of autonomous reasoning as applied in robotics. This area considers the trends and future of reasoning about and manipulating the physical environment.

2) *Real-time learning and manipulating the environment:* This area focuses on methods for learning and validating perceptions of the physical world through real-time manipulation of and interaction with objects, humans, other intelligent agents, and complex dynamic systems.

2.3 Background

Historically, research in autonomous systems has focused primarily on obstacle detection and avoidance through the use of topological maps when available, simultaneous localization and mapping (SLAM) that used depth sensors to generate local maps and navigate through them, or reactive strategies that used depth sensors or vision-based techniques, like optic flow, to detect and avoid obstacles while navigating to a goal. Relatively little work has considered reasoning about the physical properties of the obstacles or the environment. As such, prediction of future events has been limited to physics-based models and simulations that are environment-specific, not generalizable, and not robust to mobility.

Recent advances in AI and machine learning (ML), especially based on learning from massive image data sets, have yielded progress on commercial driverless car navigation on marked roadways with GPS assistance. AI/ML is also leading to advances in unmanned aerial vehicle (UAV) navigation in known and unknown indoor/outdoor environments (especially for hovering vehicles such as quadrotors). However, significant challenges remain for vehicles traversing arbitrary and unknown terrain, and for physical reasoning and interaction generally. Current research seeks to go beyond big-data-driven learning, progressing to online learning, including learning through interaction with the environment. Approaches are sought that combine control and AI/ML—for example, incorporating model predictive control, learning by demonstrated example, and learning by failure in realistic environments and against adversaries.

Mechanics and manipulation research is also following similar trends, incorporating AI/ML and learning, for example, to improve grasping, and enable rapid retasking for manufacturing, reasoning about objects, and assisting during medical procedures. AI/ML methods are also being explored to reason about and predict physical events from video, predicting motion from applied force or gravity, and learning object dynamics and interactions with complex materials (e.g., the interaction of a robot leg with granular media to generate an efficient gait).

Even with these significant advances, state-of-the-art techniques still fall far short of the general intelligent agent physical reasoning and interaction capabilities needed to interact with objects and environments, and the ability to make adaptive decisions in the course of a mission based on that reasoning. To address this, intuitive physics and commonsense reasoning are beginning to emerge to explore how humans infer physical properties of objects and predict the outcomes of dynamic events, with the goal of applying these to robotics. Progress to date is limited to specific niche areas such as taxonomic and temporal reasoning or physics in a fixed and repeatable setting.

In a variety of future Army operational constructs involving robotic teammates, control theories and methodologies to enable reflexive adaptations to the environment need to be developed. An emerging school of thought is embodied intelligence, which is often closely linked with morphological computation/control. This field advocates for the seemingly obvious idea that a body's morphology is critical to physical reasoning and adaptability (as opposed to the brain or central processor being responsible for all reasoning and adaptation no matter the embodiment). Although the basic idea is somewhat intuitive, theoretical and mathematical formalizations of this construct into robotic design and control methods are open and significant challenges. The value of pursuing advances in this field can be exemplified in the example of running on uneven terrain: the timescale does not allow continuous neural control. Instead, the body mechanics are coupled with reflexive control with much faster feedback, while neural higher reasoning works at a slower timescale.

There is significant optimism with respect to the continuing advancement of AI and component technology for autonomous system navigation, reasoning and decision making, and manipulation of the environment. Technology convergence continues to unite networking, processing, sensing, and control onto portable devices and robotics. This follows general mass production trends in cellular networking, robotic, and sensor technologies. *However, the Army faces unique challenges in ground-based operations, which hinder and slow the adoption of autonomy into ground platforms.* Current and projected commercial technology will not reliably extend into new Army-relevant tactical environments and operations in dense urban areas, subterranean, and jungle environments.

2.4 Gaps and Recommendations

Several observations, technical gaps, and recommendations came out of the ASPSM discussions and breakout sessions. Some general observations were the following:

- Navigation is dramatically more challenging in ground environments than in obstacle-free air and sea environments.
- The science of robotic physical interaction is not well developed, except in static fixed environments with infrastructure.
- Physical reasoning and the laws of physics are not yet captured in general-purpose AI.
- Common-sense reasoning, going beyond physical reasoning, is not yet possible.
- Reasoning about general manipulation is very immature, and manipulation on the move has not been achieved.
- Control that incorporates perception and physical interaction is in early-stage research.
- Robotic physical partnering with humans is limited to fixed, repetitive tasks.
- Ad hoc robotic tasking and human-robot dialog are in early-stage research.
- Physics-based simulations are highly scenario-specific and not easily generalized or scaled.
- Online learning is limited to short time durations and not is well integrated with physics-based models or simulations.
- Experiments and benchmarks are needed in real-world environments.

To try and capture what makes the Army problem challenging, the difficulty of autonomous physical reasoning and action was characterized to increase along the following *axes of complexity*, all of which can be at the extreme end of the scale for a relevant Army scenario:

- Context and environment
- Task complexity
- Task variety
- Physical variety
- Degree of human assistance and interaction
- Operational tempo coupled to amount of prior knowledge, onboard versus offboard processing, and networking

To advance the science of physical interaction, there are many broad issues that must be addressed from an interdisciplinary perspective, including the following:

- Physical scene and context understanding
- Reasoning about material properties and object physics
- Dynamic physical interaction
- Combining semantic and physical perception
- Embodied introspection and lifelong learning
- Physical agility and reactive control
- Human–robot communication and task sharing
- Failure prediction, safety, and trust
- Progressing from tool to teammate

These broad research goals are not mutually exclusive and contain several critical Army elements that are unlikely to be addressed by commercial enterprise. Tactical application is reliant on heterogeneous architectures across Army platforms, networks, sensors, and processors. Distributed operation is essential and must be resilient to attack, mobility, and network failure.

Based on the ASPSM discussions on the emerging research in AI/ML, intuitive physics and commonsense reasoning, and embodied intelligence for application to intelligent systems that can reason about and perform complex actions in Army-relevant environments, it is recommended that long-term research be undertaken to support the following:

- *A physical reasoning architecture should incorporate and balance physical-world and knowledge-world constructs.* Physics-based reasoning allows for mathematical frameworks and physical law models but is domain-specific and difficult to generalize. Knowledge-world reasoning relies on logic and knowledge-based mining to provide semantic intelligence, which can be learned online and extended over many environments. An overall modular architecture is desired that incorporates both forms of reasoning to provide a general-purpose, physical-reasoning AI.

- *Learning should complement modeling.* Data-driven learned representations such as deep learners (neural networks) are potentially very powerful but quickly become abstract and lose the connection with physical interpretability, whereas physical models may be easily interpreted and can significantly reduce learning time. Physical-reasoning architectures are needed that incorporate both, and that are modular to facilitate engineering development, testing, and validation.

- *Abstractions are needed that are scalable, can be shared across agents, and facilitate fast reasoning and action.* These abstractions are needed to facilitate different levels of interaction, with different levels of fidelity and speed. Intuitive physics—reasoning based on physical modeling—may be learned, mathematically modeled, or a hybrid mixture of these. More generally, and a much more difficult long-range research direction, common-sense reasoning should be developed that spans physical and semantic abstractions.

- *General methods for characterizing families of affordances should be developed.* An *affordance* is a relation between an object (or an environment) and an organism that, through a collection of stimuli, affords the opportunity for that organism to perform an action. For example, a knob affords twisting, and perhaps pushing, while a cord affords pulling. Methods for characterizing families of affordances and generalizations are needed. These should be linked with goal-driven behavior, learning, and rewards.

- *Learning should be incorporated into physical platforms that are self-guided and curious.* The context of the physical task plays an important role in cognition and physical reasoning. Agents should have the ability to explore and learn that mimics curiosity, incorporates frustration and self-goals, links with perception and scene understanding, and combines physical and semantic reasoning.

- *Physical reasoning should be linked with the wisdom in the body.* Animal studies reveal the links between physical reasoning and interaction that exploit the specific animal morphology, sensing, and intelligence. Robot morphology, materials, and sensing should be combined with AI and reasoning in synergistic designs.

- *Control methods must seamlessly span fine to gross control, reflexive to carefully planned, and be tightly coupled with the platform physics*. Control architectures must accommodate different timescales, contexts, and tasks. They must be adaptable to platform degradation and robust to a seemingly endless set of conditions in an operational environment. They also must be tunable with respect to risk, tempo, and multiagent teaming.
- *Human–machine dialog should be developed to enable rapid learning and teaming.* Natural-language processing and other forms of human–machine dialog and communications should be developed to enable lifelong learning, task refinement, and seamless operations. This may include wearable sensors, gestures, and other unconventional methods.

2.5 Some Key Research Questions

1. How can learning methods (e.g., deep learning) be linked with physics-based models to provide general physical-reasoning frameworks?
2. How can families of affordances be generalized and formulated, and linked with goals and behaviors?
3. How can we link perception, context, and goals? How to build a learning framework that spans these? What goes beyond current reinforcement-learning paradigms?
4. How to do perception that is driven by the needs of physical reasoning? What sensing modalities are needed?
5. How to develop an analytical and design framework for commonsense reasoning that connects control and semantic learning?
6. How to build a long-term learner that uses exploration, frustration, curiosity, and self-goals?
7. What are appropriate metrics, benchmarks, and behaviors that can drive long-term physical reasoning research and development?

2.6 Conclusions

The Army has a unique set of challenges in robotics, autonomy, and physical interaction in complex operational environments. Long-term fundamental research is needed to address the critical issues outlined in this report, combining AI, learning, physics-based and common-sense reasoning, embodied intelligence, and human–machine teaming. Successful integration into physical agents will provide autonomous teammates that will transform future Army operations.

2.7 Acknowledgments

ARL organizers gratefully acknowledge the substantial contributions of many ARL, academic, and industry colleagues that participated in the ASPSM discussions. The organizers would like to specifically acknowledge and thank the 4 invited speakers: Prof Ernest Davis from New York University, Prof Jason Fischer from Johns Hopkins University, Prof Shai Revzen from the University of Michigan, and Prof Abhinav Gupta from Carnegie Mellon University.

3. Sensing and Modeling Social Dynamics

Authors: Lisa Troyer and Peter Khooshabehadeh

3.1 Summary

Social phenomena like violent protest, large-scale social influence, violent extremism, coalition formation, mass movement fragmentation, and sudden population (im)migration pose challenges for the Warfighter in the Army operational environment. Rarely can such large-scale dynamics be detected, much less predicted, through an aggregation of individual motivations, cognitions, and behaviors. Rather, these dynamics often reflect unexpected shifts that are not necessarily a response to majority tendencies. Moreover, they sometimes emerge slowly over time and sometimes suddenly erupt in surprising and even unimagined alterations of the status quo. Our inability to predict these transitions poses a threat to effective operations, and this inability reflects 2 challenges: accurately measuring and modeling social dynamics. Currently, we lack the capability to 1) objectively measure the transitions of large-scale collectives from one state (e.g., peaceful protest) to another state (e.g., violent mass protest); and 2) causally and predictively model these transitions across cultures. These limitations have not only hampered Army operations, but they have also stymied the development of generalized theories of social dynamics. Yet, new approaches to measuring and modeling social dynamics, including new sensing technologies, statistical strategies, and complex computational models, are emerging with the potential to augment and in some cases supplant current approaches. For instance, new capabilities such as geospatial sensors, biometrics, tracking of utility use, and multimedia data are enabling more objective ways to measure social action compared to traditional approaches such as observation and survey methods, which are often fraught with bias. Complex modeling approaches and new statistical techniques are being developed that more accurately represent the distributions of large-scale action and the often punctuated shifts from one collective state to another. This meeting investigated how these developments—in sensing and modeling dynamics—can be leveraged to improve the prediction of social trends and enable greater scientific understanding of the mechanisms and trajectory of how large-scale groups transition across social states.

3.2 Objective

The objective of this meeting was to improve the Army's capacity to understand, model, and predict large-scale collective dynamics by charting a path forward for research to improve measurement and modeling of social dynamics. It focused on identifying new opportunities and challenges for sensing and modeling social dynamics that are emerging through biophysiological sensors (i.e., microbiometrics) and Big Data available from a range of sensors (e.g., social media, "Internet of Things", utility usage, geospatial sources). In addition, the participants sought to identify new data management strategies to integrate different data sources that vary widely in temporal and spatial scales. That is, integrating different data sources is a challenge. Furthermore, participants discussed the range of emerging strategies to dynamically model complex social systems. To address this objective, participants tackled 3 challenge questions:

- **Challenge Question 1:** What established and new technologies exist for sensing social dynamics, including more micro biophysiological sensors and more macro sensors (e.g., geospatial sensing, utility use, multimedia sensors) and to what degree of granularity can these sensors assess social dynamics (i.e., collective action) as opposed to human behavior (i.e., individual actions)?

- **Challenge Question 2:** What are the limitations of existing sensing technologies for assessing social dynamics, including integrating different sources across cultures/regions?

- **Challenge Question 3:** What are the challenges for modeling social dynamics; what strategies currently exist; what are their limitations; and what improvements are needed to enable validated predictive, causal models? How can operators use models of social dynamics to shape the current/future operating environment?

Twenty-five attendees from academia, Army, Department of Defense, and the private sector participated in the meeting, through a series of keynote addresses, summary overviews of the 3 challenge questions, breakout groups to address them and report back to the group their findings, and generalized discussion regarding a way ahead for the Army in terms of remaining opportunities and challenges.

3.3 Opportunities

The meeting began with a "fire-starter" talk and a keynote address, followed by a discussion of the current state of the art in social science research and the failure to discriminate between individual human behavior and social dynamics. The former focuses on individual cognitions, attitudes, and actions, while the latter depends on the complex and dynamic nature of how relationships change—how information channels among actors shift over time. Yet, the latter also depends on the different propensities of the individuals comprising the collective and how those change over time, which is itself conditioned on the nature of the collective. How to sense and predictively model those changes, or dynamics, is part of the difficulty that meeting participants addressed through the challenge questions. For each challenge question, a set of recommendations emerged from the discussions.

3.3.1 Challenge Question 1 (Sensing Social Dynamics)

The group concluded that there is an ever-growing array of opportunities to sense shifts at micro (individual and small-group) levels and macro (large-scale collective and population) level. At the micro level, the group identified existing technologies commonly used to more objectively measure individual behaviors: Galvanic skin response, heart/respiration rate, brain sensing (e.g., functional magnetic resonance imaging [fMRI], electroencephalography [EEG]), thermography, pupilometry, eye gaze, voice patterns, and epigenetics. Wearables and social badging technologies are also gaining attention, as they identify ways to track groups through time, space, environments, and social contacts. Wearables can also serve as "alert" systems in an "Internet of Things" environment. Several participants noted that the use of these technologies is being expanded from individual-level data collection sources to tracking dynamics in small and large collectives—for example, to track how brain activity changes among members of a group who are experiencing an external threat, such as discrimination or injustice. Another example is the capacity to track the diffusion of emotions, such as anger, throughout a collective using facial thermography. These attempts to sense collective dynamics, however, are still nascent but represent an opportunity to substantially advance abilities to obtain objective measures of social dynamics (as distinct from individual behavior), insofar as many of the biophysiological shifts that these technologies detect are extremely difficult for individuals to control.

At the macro level, a wide range of emergent technologies capable of capturing large-scale collective dynamics was identified. Among them are geospatial data, vehicle traffic data, computer visioning, utility use (e.g., water, electricity, cellular phone), financial data, and media analytics (to include both social media and print media)—all of which can detect very large and even population-level dynamics. These data are increasingly available at varying levels of granularity, depending on the region of the world.

3.3.2 Recommendations for Challenge Question 1

- There is a dearth of research on how these different sensors at both micro and macro levels work cross culturally. With respect to biophysiological sensing, there has been some cross-cultural research on "emotion work" demonstrating that certain types and levels of emotional responses are learned and may vary cross culturally. Generally, however, research suggests that this learning occurs at very early developmental stages (especially during brain development). Consequently, there is an opportunity through cross-cultural research to begin to pinpoint these differences.

- Similar challenges with regard to macro-sensor availabilities and penetration of technologies such as "Internet of Things" exist. Research to baseline and then track these technologies within different cultures and regions is needed, alongside systematic research to understand their cultural significance. For instance, social media is used very differently in some cultures than others. Moreover, very recent research has discovered that adversarial groups often use social media to present and disseminate negative information as broadly as possible about a target individual or collective, whereas nonadversarial groups use social media to establish virtual communities and communicate with close friends and family with whom they interact. That is, social media is used by different groups vis-à-vis their position in a complex social system.

- A third challenge related to established and emerging sensing technologies involves issues of privacy and ethics. There has been little systematic research on whether using technologies such as those discussed in the meeting would have unintended adverse effects in terms of building alliances as a result of perceptions of privacy intrusions. Existing perceptions of the impact of using these sensing technologies to better understand social dynamics are largely anecdotal. As meeting participants noted, however, changes in the Common Rule (a process for determining whether a research study involving humans requires Institutional Review

Board [IRB] assessment to ensure ethical conduct of research) have relaxed requirements, such that many data sources that are publicly available no longer require IRB monitoring. Nonetheless, even with anonymization of data, researchers have demonstrated conditions under which integration of data from diverse sources can lead to exposure of personally identifiable information (PII). Many of the macro sensing technologies (especially social media, Internet of Things) represent ungoverned or semi-governed spaces and research on such spaces (e.g., if/how they evolve into formally governed spaces in a global world) is currently lacking.

3.3.3 Challenge Question 2 (Integrating Data from Social Sensors)

The discussion on cross-cultural challenges that arose with respect to challenge question 1 provided a good segue for challenge question 2. The group identified 2 critical dimensions that need to be considered when it comes to data integration: temporal and spatial. With regard to temporal challenges, some sensors discussed operate at the scale of milliseconds when it comes to detecting change (e.g., brain scanning technologies), while others may reflect changes occurring at hourly or daily rates (such as vehicular traffic patterns). Regarding the spatial dimensions, experts in the group noted that how widely over physical space a sensor can detect changes varies. Geospatial technologies, for instance, are becoming increasingly malleable to detect both highly localized changes and very broad changes in landscape and population activities. To date, there have been only relatively coarse attempts to integrate multiple social sensors of the type discussed for challenge question 1, owing to these temporal and spatial differences in their capabilities. They seem limited to staying within the micro/biophysiological domain or the macro domain. For instance, studies using galvanic skin response in concert with heart/respiration rate and brain imaging to track such phenomena as the diffusion of threat perceptions across a group or development of cohesiveness have been conducted, but these are largely controlled experiments (and invasive). At the macro level, financial and utility data are commonly used together to index emergence of state fragility that places a government at risk of civic uprising. Integrations of, and rationales for, integrating biophysiological data and macro data may be limited.

The group did identify a few possible empirical benefits of research on how to integrate different sensors across micro and macro domains if more effort was placed on data integration research, analytic strategies (see Section 3.4, Way(s) Forward), and visualization research. For example, such research programs might include research on how Warfighter wearables could be used to collect, analyze, and visualize data in their environments to detect risk environments (e.g., areas where collectives are mobilizing against our forces based on macro data sensor inputs conveyed to the Warfighter once the challenges for sensing and modeling social dynamics noted in Sections 3.3.1 and 3.3.5 are addressed). The conclusion was that this should be an area to monitor through an Army social dynamics research agenda.

3.3.4 Recommendations for Challenge Question 2

- Clearly, the use of sensors varies widely regionally and culturally. For example, although Twitter is an often-sought source to detect social networks, diffusion of information, shifts in sentiments, and emerging conflicts, its penetration is problematic. Some countries ban or heavily control and monitor its use by citizens; others have their own systems that are more popular than Twitter (e.g., Russia's Vk [VKontakte]). Consequently, basic research on the factors that drive users to different platforms is needed.

- When it comes to macro sensors, regions vary dramatically in the extent to which data from those sensors are available. Yet, this may be an important source of data on large-scale social dynamics and the interface between institutional structures (e.g., economics and governance institutions). For instance, in some regions, water and electricity are state-controlled and even made unavailable at times. Participants with research experience in these regions noted that communities seemed to adapt to that control, changing daily activities in accordance with these control structures, with little impact on sociopolitical dynamics. Thus, a shift in such utility use in such cases would not necessarily signal an emergent change in the population. There is, however, research showing that when unavailability of such resources is abrupt and unexpected, it can lead to mass uprisings. The tipping points and how they relate to institutional and cultural dynamics are not known, suggesting the need for ongoing basic research in this domain.

- While availability and penetration of different types of macro sensors and the data they produce varies considerably, the group felt that research on how different sensors are used and cultural interpretations of them was valuable because the availability and penetration would likely increase in the decades ahead.

3.3.5 Challenge Question 3 (Modeling Social Dynamics)

Challenge question 3 centered on overcoming the current state of the art related to how social dynamics are analyzed. Current tendencies are to use inferential statistics based on assumptions such as normal distributions of behavior (at both individual and collective levels), independence of observations, and continuous evolution of collective states. As participants noted, collective-level action is often not normally distributed; and it is rarely independent. Randomized experiments at large-scale levels (e.g., crowd-sourced experiments) are just beginning to come about, but even then, addressing assumptions of independence are problematic or opaque. Additionally, collective action is not generally continuous. Collective action often abruptly shifts from one state to another (as when a group suddenly turns violent). Another point that the group focused on was the fact that most social and behavioral science research tends to focus on dramatic events (e.g., deviance, government overthrow, violent confrontation between groups, war). Recently, researchers have argued that scientists are over-sampling on variables of interest without considering the distribution of events. This may be partly because data on events such as peaceful protest are more difficult to obtain (i.e., they are not reported as frequently as dramatic events).

On the other side of the coin, participants voiced the position that often it is the rare events/actions in a distribution of social dynamics that are of both the greatest interest and the greatest risk. Relevant to this is the phenomena sometimes referred to as “Normal Accidents”, which reflects complex interdependencies among social, natural, and physical systems that can lead to disastrous outcomes (e.g., airplane crashes, nuclear power plant failures, oil spills). More recently, emerging research has demonstrated that commodities markets (which tend to be very turbulent, with trading partners in a value chain frequently shifting even in the course of a day) can have extremely powerful effects on global sociopolitical dynamics. The point is that capacity to model long-chain, complex interdependencies may enable new predictive models of risk points in a complex social system that cascade and snowball, leading to unexpected dramatic outcomes.

Related to this, on the one hand, participants in the meeting cited research that moves away from traditional social science statistical approaches toward modeling approaches drawn from physics, natural sciences, and computer sciences that are emerging to capture complex crowd behavior (e.g., assessing how fluid dynamics models may capture evolving social dynamics; models on swarming, flocking, herding from nonhuman social species might fit human social patterns; and mortality models from epidemiological research that treat social states as punctuated equilibria that persist or die). On the other hand, some participants noted that while these models may fit a particular data set, validation seemed missing from the normal research cycle. The concurrence was that analytic strategies exist but have not been widely adopted in the social sciences. Likewise, however, existing social theory and research are often not adequately attended to by computational scientists.

Another concern that the group discussed was the need for information to be conveyed more clearly so that operators and analysts could not only rapidly ingest it, but also make sense of it. This may be a somewhat applied concern, but it does require research on how operators work with data and data analytics related to complex systems, the cognitive processes they use, and how collective operational decisions are made.

3.3.6 Recommendations for Challenge Question 3

- There is an opportunity to develop truly integrated computational research programs that equally engage social scientists and computational scientists, so as to advance analytics but also avoid reinventing the wheel when it comes to social science theory and research. However, this opportunity is not without challenges. There are incentive structures that often discourage such interactions (such as requirements of publishing in one's own disciplinary silo in tenure-track systems). Support for multidisciplinary *integrated* research teams may be a productive path forward. The Multidisciplinary University Research Initiative (MURI) program, and perhaps "mini-MURIs" that are MURI seedlings, could be effective.
- Another avenue to promote truly multidisciplinary research would be to leverage the Open Campus model to bring disciplinarily diverse scientists together around specified challenge problems at ARL. This would most likely require a funding mechanism to enable their time/effort.
- Developing stronger ties between data scientists and experts in cognition and decision making may enable better strategies to facilitate data/analytics ingestion by operators.

3.4 Way(s) Forward

The Army is the only agency that has developed an investment portfolio that balances micro-level technology development in biometrics (spanning cyber, engineering, physical, psychological, biological sciences) with macro-level social sciences and computational sciences. Moreover, the Army integrates basic and applied sciences through a combination of funding directed toward basic science research that draws on the intellectual capital vested in universities and funding that spurs commercial-academic-Army partnerships. This tripartite approach has positioned the Army to be the leader in the science of social dynamics. In the last session of the meeting, several suggestions were made to continue the synergy that developed in the meeting across a diverse group of scientists, agencies, and organizations. These included the following:

- Investigating the publication of an edited volume of a special journal edition that would collect expanded versions of overviews of topics/approaches/ challenges cited in this report.
- Developing a MURI topic based on one or more of the challenges on which the meeting focused.
- Participating in the Networking and Information Technology Research & Development (NITRD) Program of the National Science and Technology Council, which was established by the High-Performance Computing Act of 1991 to develop a framework for federal agency funding of IT research and development (R&D) as well as the Office of Science, Technology, and Policy under guidance and direction of the Executive Branch, including input on social computing (Military Services is a member agency of NITRD).
- Proposing a cofunded mini-MURI seedling program to pilot test multidisciplinary approaches to the challenges described in this report.
- Collaborating across academia, industry, and ARL to facilitate integration of intramural and extramural research project to address the challenges noted in this report.

3.5 Conclusion

National security risks are increasingly arising from social threats—that is, threats arising related to social action from group to large-scale cross-national collective adversarial efforts, which often take violent turns, put citizens' lives at risk, and compromise global sociopolitical order. These social dynamics may result in undermining or destroying fundamental institutions that stabilize societies and facilitate democracy, including our political systems, kin systems, financial systems, religious systems, public health systems, and educational systems. National security threats that place our Warfighters in danger—both domestically and abroad—arise when adversaries disrupt democracy, capitalism, humanitarian operations, and citizen access to fundamental services. Using a variety of tactics, adversaries are increasingly exploiting vulnerabilities in financial markets, cyber systems, social media, political processes, and even R&D activities by focusing on the social factors that make the foundations of any society susceptible to chaos. The capacity to detect and prevent multifarious attempts to undermine social systems is critical to the Army mission by providing situational awareness and more effective operational and resource allocation decision-making. For over 100 years, the state of the art in social science research has relied on observational and self-report (survey) methods to measure states and trends in large-scale collective attitudes and actions, as well as statistical models that defy social realities. Yet, these measurement methods and the models that use the measurements as data inputs are known to suffer from a high degree of bias, with relatively low predictive capacity. New technologies are emerging, however, to enable more objective and less intrusive capabilities to sense shifting social states (e.g., transitions from peaceful protest to violence; diffusion of social influence across a large collective). Moreover, new methods to model social dynamics, which treat social groups as complex systems as opposed to a simple aggregation of individual behaviors and actions, offer the promise of more accurate and predictive models. Still, challenges remain to be addressed that require new basic research approaches described in this report.

3.6 Acknowledgments

ARL organizers gratefully acknowledge the substantial contributions of many ARL, academic, and industry colleagues that participated in the ASPSM discussions. The organizers would like to specifically acknowledge and thank the 2 invited keynote speakers: Prof Alex (Sandy) Pentland from Massachusetts Institute of Technology and Prof Rene Weber from the University of California, Santa Barbara, as well as the Panel Chairs, Dr Morteza Dehghani from the University of Southern California, Mr David Fordyce from ARL, and Dr Tim Gulden from RAND Corporation.

4. Learning and Reasoning in Complex Data Environment

Authors: Tien Pham, Brian Henz, and Purush Iyer

4.1 Introduction

The increased prominence of AI approaches over the past 25 years has been boosted in large part by the adoption of statistical and probabilistic methods, the availability of large amounts of data, and increased computer processing power. Over the past decade, the AI subfield of ML, which enables computers to learn from experience or examples, has demonstrated increasingly accurate results, causing much excitement about the near-term prospects of AI. While recent attention has been paid to the importance of statistical approaches such as deep learning, impactful AI advances have also been made in a wide variety of other areas, such as perception, natural language processing, formal logics, knowledge representations, robotics, control theory, cognitive system architectures, search and optimization techniques, and many others.[1]

4.2 Objectives, Scope, and Goals

As numerous studies have made clear, the Department of Defense (DOD) must integrate AI and ML more effectively across operations to maintain advantages over increasingly capable adversaries and competitors.[2] It is envisioned that future military operations will involve teams of highly dispersed Warfighters and robotic agents operating in distributed, dynamic, complex, cluttered environments. Most current research (and success) in AI and ML is done with extremely large collections of relatively clean, well-curated training/operational data, with little background noise. Army domains, on the other hand, present rapidly changing situations; noisy, incomplete, and potentially erroneous data; small numbers of samples for many important cases; and strong impacts of deceptive adversaries. Research on learning and reasoning with such data is not yet well motivated for commercial applications. The goal of this ASPSM workshop on Learning and Reasoning in Complex Data Environment is to understand the near- and far-term implications of the AI and ML capabilities and developments within the challenging context of operating at the tactical edge.

ASPSM topics include but are not limited to the following:

- Adversarial distributed ML
- Robust inference and ML with conflicting sources
- Adaptive online learning in real time
- Adversarial reasoning integrating learned information
- Resource-constrained adaptive computing for AI and ML

4.3 Focus Areas and Discussions

The meeting focused on 2 specific questions:

1) What are the key AI and ML technical gaps that ARL needs to address with respect to the Army's complex data environment?

2) What are the key user requirements and use cases with available relevant data sets?

4.4 Focus Area 1: What Are the Key AI and ML Technical Gaps that ARL Needs to Address with Respect to the Army's Complex Data Environment?

Discussions during this breakout session identified multiple research gaps that ARL must address while maintaining reasonable expectations of outcomes, while drawing distinctions with the research interests of the computer science industry. The gaps identified are contained within 4 themes, namely, Trust in Human–Agent Teaming (HAT); Continuous Learning; Common-Sense Reasoning; and the Tradeoffs between Information-Poor Data and a Plethora of Sensors, and Impact on ML Approaches. Deemed important, the gaps provide topics for collaboration between academia and ARL through ARL's Open Campus initiative.

4.4.1 Trust in Human–Agent Teaming (HAT)

The vision of the future Army of 2035 as presented through the ARL AI/ML Essential Research Area (ERA) campaign leads portrays autonomous agents (robots) working in collaboration with Soldiers. Given this scenario, the panel discussed the main issue guiding interaction between a robot and a human handler as being one of trust, the ability to share each other's mental models, and management of limitations (physical and cognitive) given the limited bandwidth in exchanging information between the two. Of course, the problems are exacerbated when a group of human beings interacts with a group of robots/autonomous agents. Specific issues include brittleness in current ML-based approaches, as well as the effect of adversarial actions; indeed, insider threats and cybersecurity issues complicate interactions in human–agent teams especially under the fast tempo of actions in a theater of operation. The panel also considered the lessons that can be learned, generalized, and applied to HAT, including how canine units train and operate, how mental models are built, and how limitations are accounted for. Apart from these technical questions, the panel opined that ethical questions in a human–agent team need to be addressed, too; for example, the common question that is currently being posed in deploying autonomous vehicles is assignation of blame in the event of an accident. At the end of the day, all of the interactions assumed here between members of a HAT are predicated on the assumption that both humans and machines have access to their internal state, which they can communicate to other entities.

4.4.2 Continuous Learning

Explainable AI is the current buzzword—in which the goal is to explain the actions of an ML algorithm in terms of overall functionality of an implementation—and is definitely a holy grail at this point. While explainable AI is interesting in and of itself, it will likely be a critical component of continuous learning, where a robot (or an autonomous system) can learn from a human and its experiences in a continuous manner. Soldier after-action reports are written to drive the decision making, to learn about the adversary, and to devise tactics and strategies. If AI components will be part of a HAT, then it is obvious that continuous learning will have to play a part in developing robust AI systems of the future. Indeed, explanations can help with the learning process (in human–child or apprentice–master relationships) and help the learner develop a causal model/mental model to explain observations.

There are several questions that need to be answered. A primary question is how to formulate and address agent-centered teaching/learning. Lessons learned from the failure of massive open online courses and the prohibitively expensive small class sizes in human learning (in K–12 education) suggest that trying to teach an autonomous agent new skills (based on an analysis of its capabilities at any point in time) would be a challenging technical problem, and likely an expensive problem to solve. However, it will be necessary if continuous learning needs to be successful. The earlier reference to the relationship between a dog and its handler is germane here; will a robot need to be trained by a Soldier so that they can adjust to each other's idiosyncrasies? If so, what are the ramifications of "particularization" to interactions in a group setting? A second foundational question that was discussed by the panelists was the role that generalization from examples, or domain adaptation, could play in continuous learning. In particular, the panelists wondered if domain adaptation could be made unobtrusive and continuous. Current methods of domain adaptation are in their infancy, but there is hope that implementations of Bayesian learning, if done in a low-cost, incremental manner, could lay the foundation for true continuous domain adaptation. There are at least 2 examples of never-ending learning being tried right now: the Never-Ending Language Learning at Carnegie Mellon University and the RoboBrain project at Stanford. In some sense, these experiments can be considered as brute-force approaches, making invention of a science of continuous learning necessary. However, learning implies the communication of ideas, which implies the need for a language to express thoughts and ideas (or concepts in a robot's internals). Clearly, therefore, there is a need for continuous generation of ontology characterizing new concepts for continuous learning to succeed—a great challenge in and of itself.

4.4.3 Common-Sense Reasoning

Children do not solve partial differential equations to predict trajectories in a number of occasions: skip steps to climb, throw a ball, catch a ball, or tip over the cookie jar on a high shelf. Similarly, children bond with their grandparents without any instructions, or learn to say "thank you" or "please" with minimal instructions. There is, of course, a plethora of social science theories and principles that explain human behavior, but none of them have computational content. Principles such as path of least resistance or temporal continuity in parsing stories are used to explain a behavior post-facto. Similar invariants can be stated in explaining dynamics or motion (for an example, see the study of predatory behavior of brown bats using sonar beams by PS Krishnaprasad at the University of Maryland[*]). However, synthesizing such behavior from ground zero is a lot harder. But that is precisely what autonomous systems of the future need to be capable of doing, without explicating high-dimensional representations of the world around them and complicated algorithms acting on those representations. While there are attempts at formalizing common-sense reasoning, the area is still in its infancy. A potentially lucrative approach could be to generalize Kahneman and Tversky's work, for instance, to derive computational principles that could be the basis for quick reasoning by autonomous systems.

4.4.4 Tradeoffs between Information-Poor Data and Plethora of Sensors, and Impact on ML Approaches

While this topic was a hodgepodge of ideas, all discussions centered on the current popularization of deep learning. While it is clear that convolutional neural networks (CNNs) work like a charm in certain specific applications, the use of CNNs in situations where there is a lot of nonsensical data, or less meaningful data, as in military settings, is an open question. In particular, the panel discussed whether it is possible to build sensor models into CNNs. Furthermore, the panel opined that the flip question of whether CNNs can be built into a network of sensors is also worth pursuing. The latter is especially important if we wish to have reasoning at the edge, in the sensor fabric, so that the need for communicating huge amounts of data to a central server is reduced. An additional point made was that, in spite of the success of CNNs, inference mechanisms in current use are too slow—on the order of 30 frames per second on multicore hardware. Being able to process bursty data on a sensor fabric would definitely be challenging. In particular, processing rates of 100 frames per second would indeed revolutionize the use of CNNs at the edge.

[*] https://isr.umd.edu/~krishna/images/bats_in_pursuit_3348.pdf.

Assuming sensor fabrics are targeted for research on ML, which falls squarely into the domain of ARL's efforts in the Internet of Battlefield Things (IoBT), the central question would be how nodes of an IoBT network can start with a model and learn from data, in a self-aware manner. Importantly, it was argued that confidence-based learning, and correlation among sensors, could be used to build robustness against attacks from adversaries during the learning process. Finally, it was argued that learning needs to be calibrated to be both smooth and slow to avoid data poisoning adversarial attacks.

A final holy grail-style challenge was raised at the meeting: Can learning be parameterized so that it can be used in a "plug and play" manner? Indeed, this lofty goal would require that typical ML applications be robust, over-provisioned (such as the human immune system), self-aware, capable of carrying out self-repair, and so on.

4.4.5 What Must ARL Do to Address These Gaps?

Silicon Valley companies such as Twitter, Facebook, and Google have access to voluminous data that they guard zealously. However, when they do need help, they release versions of the data and set up challenge problems for the academic community to participate in. Could ARL, using its Open Campus initiative, create such a framework? Clearly, ARL is privy to data that has national security implications. But it does have the chance to bring people and algorithms in, and test their data on new ideas and new approaches. Making this happen in a consistent way that meets the constraints of the US Government and rules that academics play by could have a positive impact on DOD's capabilities in the essential areas of AI and ML.

4.5 Focus Area 2: What Are the Key User Requirements and Use Cases with Available Relevant Data Sets? Gaps and Recommendations?

4.5.1 Complexity of Data for the Army; Collection, Storage, Processing, Analysis and Understanding

One of the primary gaps identified by the attendees was the lack of relevant Army data sets for training AI and ML algorithms. Possible causes of this gap were identified as access control, knowing the location, true lack of data, and lack of useful data. One suggestion that was voiced multiple times was the relevancy of available data, so while the Army has a number of data sets, these are either not applicable or uncollated/labeled and oftentimes both. A possible solution to this is to develop a capability to collate these data sets and make them available to a wider user base. Some potential sources of existing data were identified:

- Night Vision and Intelligence Center of Excellence (ICoE) with availability on the Defense Systems Information Analysis Center (DSIAC) (formerly the Military Sensing Information Analysis Center [SENSIAC]).
- US Army Aviation and Missile Research, Development and Engineering Center (AMRDEC) data sets include data on missile and target acquisition, Mission Command Battle Laboratory, Leavenworth, Kansas:
    - Issues include variations in signatures collected given various scenarios.
    - IBM Watson attempts to learn doctrine from tactics, understanding critical information, and the utility of trained AI at the tactical level.
- US Army Training and Doctrine Command (TRADOC), G2 data includes multiple scenarios, and some data archives:
    - Issues include need for subject-matter expert input and feedback for training sensors at the edge, online learning to identify "zero day" attacks, incremental learning, and DOD-specific scenarios.

4.5.2 Goals for the AI and ML in Complex Environments

A number of goals were identified by participants in this panel that will provide the Army with useful deployable AI capabilities:

1) Focus on lower echelon.
2) Help manage forces; attempt to capture the wisdom of the commander.
3) Recognize, act as filter for commander.
4) Reduce operator workload.
5) Teach with humans for learning new environments.
6) Learn by instruction.
7) Intelligence, surveillance, and reconnaissance (ISR) synchronization.
8) Situational development—tracking classifying targets, etc.
9) Support targeting—human in loop, battle damage assessment.
10) Situational understanding—learn patterns, identify missing pieces, and so on.

11) Intelligence Preparation of Battlefield (IPB) update threat models, support wargaming.

12) Operational resource management—very important at the Tactical Operations Center level as fuel and water are the biggest problems in the Army (Communications-Electronics Research, Development and Engineering Center [CERDEC])

This list is not exhaustive but can be summarized as a need to support the lower echelon by providing useful information to reduce cognitive load, with systems that learn over time from the operator (gain wisdom). One of the difficulties, or advantages, of adaptive AI, is that over time this AI becomes subjective and more tailored to the individual user's preferences and risk tolerance.

4.5.3 Utility of Simulations and Synthetic Data

Current efforts have not been successful due to accuracy of simulations, unrealistic simulated battles, and limited fidelity of real clutter. Most successes of generated or synthetic data currently are through perturbation of real imagery. Use of simulations in a tactical environment to provide back-of-the-envelope calculations for analysis is viewed as a potential advantage for planning.

4.5.4 How Do We Validate That Systems Are Improving?

Much of the currently deployed ML does not adapt or learn on the fly. One open question is how do we objectively measure if a decision is correct? Some applications will require this type of adaptivity, including network security where a static model may not identify a zero-day attack. What is the best way of adapting (i.e., using a sliding window to only learn from recent observations, to continuously incorporate new observations, or something in between using some plasticity in learning)?

4.5.5 How Do You Learn at the Edge?

There is not a significant amount of active research in industry or academia focused on learning with limited storage space, connectivity, power, and other physical constraints. Industry is working under the assumption of full connectivity and is focused on a cloud-based platform. The Army must consider hardware, software, and algorithms that are resource aware. There should also be an emphasis on approximation and accuracy. For example, to compete against near-peer competitors, we need fast and accurate automatic target recognition.

4.5.6 Reasoning About Adversary's Intent

An existing gap is the ability to perform planning and reasoning in a simulation environment. We do not have models of emergent behaviors, enemy behaviors are different than ours, and every theater is different (e.g., Korea vs. European Union). The Army needs the ability to provide context and reasoning; for example, social media can be used to partially provide this picture, including movement of equipment and random cyberattacks.

4.5.7 Other Challenges

- Robustness to surprise
    - Enemy tactics, techniques, and procedures (TTPs) and perceived enemy intent
    - Emergent behaviors
    - National Geospatial-Intelligence Agency
    - Enemy does not follow traditional phases
- Gray zone scenarios
    - Info, cyber, kinetic/terrorist attack
    - Causal exploitation
- Multidomain battle
    - Limited fuel for autonomy
    - Reduced signature

4.6 Highlights and Identified Gaps

This meeting, through presentations and discussions with Army, industry, and academic subject matter experts, identified 3 technical gaps that ARL must address for the Army to achieve and maintain a tactical offset over adversaries through 2035. These gaps are refinements of the AI and ML ERA gaps but also provide the basis for developing concrete efforts to address these gaps. Following these highlights, recommendations are provided to address the 3 gaps identified in both focus area discussions as follows:

- The first technical gap identified in both focus areas is that of trust in the deployed algorithms. This question of trust in AI deals with finding answers to a number of open questions for research including V&V of algorithms (particularly adaptive or online learning), providing realistic expectations for AI, and learning individual preferences for cooperation and risk analysis. Particularly highlighted by Army experts is the fact that loss of trust in deployed AI will result in resistance to use in the field. Closely related to this question of trust is the continued validation of deployed adaptive AI algorithms on learned tasks and verification of expected behavior.

- During this meeting it was recognized that many Army data sets exist from field tests, laboratory tests, and deployed systems. Most all of the Army attendees provided lists of data sets that exist within their respective facilities. The challenge is that these data sets are not collated in central locations, often contain unlabeled data, are in many different formats, and are of unknown or limited quality and value. Whereas academia and industry rely heavily upon labeled, voluminous, clean data sets for learning, the Army faces a real gap in this area.

- Highlighted by most all Army attendees is the real gap in deploying AI due to the locality of AI, computing, and communications. This is viewed by many attendees as a developing gap as sensors are proliferating on the battlefield while the ability of the network to transport this data to the point of need is not improving commensurately. Prior to this meeting, this was a known gap, but the extent of the issue and the current growth in demand for pushing AI to as close to the point of collection as possible were a surprise.

4.7 Recommendation(s)

- ARL must develop an internally led program that provides a quantifiable V&V of adaptive AI algorithms and communicates the AI reasoned solutions to a Soldier with trust and understanding. This program will cross the AI and ML and the HAT ERAs. The close interaction of humans and AI agents in stressful environments is a unique defense application of AI and must be led by the Army.

- ARL, and the Army at large, should provide resources, and ARL must lead the effort to capture and maintain data from across the Army as it is required to develop future AI algorithms and software. These data must be in an accessible location for Army collaboration and development efforts. It is also clear that for the Army to be successful in developing deployable AI, we must address the question of appropriate data for training; industry and much of the academic AI development is dependent upon these labeled, voluminous, clean data sets for learning. Success in this effort will lead to pathways to addressing Army gaps through bringing in external collaborators, through data sharing; generating synthetic data with transfer learning to the real world; and providing a central location for Army researchers to share common data and environments.

- Finally, any developed programs must consider from the beginning how their outcomes are affected or viable in an adversarial environment.

4.8 Acknowledgments

Focus Area 1: What are the Key AI and ML Technical Gaps that ARL Needs to Address with Respect to the Army's Complex Data Environment?

- Dr Pedro Domingos (University of Washington)
- Dr Tom Dietterich (Oregon State)
- Dr Rama Chellappa (University of Maryland)
- Mr Joseph Cox (Army G2)
- Dr Liyi Dai (ARL/Army Research Office [ARO])
- Dr Augustus Fountain (Deputy Assistant Secretary for Research and Technology [DASA(RT)])
- Dr Carl Hart (US Army Corps of Engineers)
- Mr David Masters (Department of Homeland Security [DHS])
- Dr Cory Miller (ARL/Weapons and Materials Research Directorate [WMRD])
- Dr Tien Pham (ARL – Information Sciences)
- Dr Brett Piekarski (ARL – Sciences for Maneuver)
- Dr Brian Rivera (ARL/Computational and Information Sciences Directorate [CISD])
- Dr Mudhakar Srivatsa (IBM Watson Research Center)
- Mr Shane Thompson (AMRDEC)
- Dr Gene Whipps (ARL/Sensors and Electron Devices Directorate [SEDD])

Focus Area 2: What are the Major User Requirements and Use Cases with Available Relevant Data Sets?

- Arnold Boedihardjo (Engineer Research and Development Center [ERDC])
- Edward Colbert (ARL/CISD)
- James (Vic) Fink (US Army ICoE)
- David Gunning (Defense Advanced Research Projects Agency [DARPA])
- Nandi Leslie (Raytheon)
- Patrick McDaniel (Penn State)
- Alex Miller (Army G2)
- John Fossaceca (ARL/CISD)
- Andrew Ladas (ARL/SEDD)
- Kesny Parent (CERDEC/Intelligence and Information Warfare Directorate [I2WD])
- Dale Shires (ARL/CISD)
- Steven Vanstone (AMRDEC)
- Stephen Russell (ARL/CISD)
- Justin Wright (CERDEC/Night Vision and Electronic Sensors Directorate [NVESD])
- Shane Thompson (AMRDEC)

5. Uncertainty Quantification

Authors: Ernest SC Chin, MaryAnn Fields, Jaroslaw Knap, and Brian Jalaian

5.1 Introduction

With the increasingly dynamic, complex, and volatile contested operational environment, there exists a critical need to acquire and process operational information at the "speed of fight" to facilitate robust decisions for the course of action. Understanding and integrating UQ to account for variabilities within an ensemble of rapidly advancing and maturing edge computing, physical and social models, ML, and augmented autonomy will empower superior decision aids for Soldier adaptability at the changing speed of fight.

In the following sections, we discuss 3 topics focused on UQ future readiness for capabilities in 1) offensive–defensive adversarial disruptions, 2) robust design, and 3) achievability at the tactical edge. These topics culminate and converge to a set of research recommendations toward addressing risk and trust by means of UQ for robust decisions in unforeseen future battlefield scenario:

1) Offensive–defensive adversarial disruptions: Can and how shall UQ be employed and incorporated into active learning to deal with sparse, dynamic, incomplete data and information for adversarial detection, mitigation, and preemptive actions?

   This topic is rooted in the research strategy to achieve tactical battlefield superiority with Soldier-augmented decision capability, relevant to Army's needs in i) Asymmetric Vision and Decide Faster, ii) Training and Extrapolation, and iii) Cognitive Augmentation. This capability is envisioned as near-real-time abilities to a) collect, b) learn and analyze, c) reason, and d) infer from heterogeneous (multiple types of) data. UQ is crucial to these processes:

   a) To Collect: Heterogeneous data continuously collected from various types of sensing devices and human intelligence are often sparse, spatially and temporally distributed, and incomplete. Incorporation of UQ practices into near-real-time data collection will allow machine-tracking the quality of these data (e.g., fidelity, resolution, limitations) and updating the confidence/trust level for augmented decision on-demand.

   b) To Learn and Analyze: Data analytics and ML methodologies are the key to understand, model, and predict from collected data. UQ is necessary to address confidence in the machine learning from its output. It guides where additional data are necessary to increase the output certainty for artificial decision or prediction. It is an approach to detect anomalies in the data, possibly deceptive in nature, and also to track rare events that ML from transferred knowledge (prior trained data) would most likely neglect. This is critical to Soldier agility and adaptability as it allows combatants to anticipate the unknown and mitigate disruptive situations.

   c) To Reason: Reasoning often derives from attaining trust on the basis of cumulative simulations, mathematical causality, and statistical-probabilistic analysis that converge to a set of conclusions for optimal situational outcomes. This constitutes the heart of UQ such that cumulative uncertainty will converge toward trust.

d) To Infer: Inference is the final step to abstract and tailor outcomes from reasoning. Outcomes include robust options on courses of action in high-tempo operational environments. When communicated within the context of trust from UQ, this will augment and unburden combatant time-of-fight decisions to victory.

Though the goals and expectations of UQ in the described scenario appear systematic, the technical approach and methodology to computationally apply, implement, and track uncertainties are quite formidable. Among these UQ challenges include the following:

a. How to develop/extend/adapt fundamental mathematical/statistical theories to have UQ for machine learning and artificial reasoning models, which generate accurate UQ measures in addition to merely an output state, measure, or prediction.

b. How to best organize and track fidelity of data from distributed and heterogeneous sources (e.g., hierarchical or ad hoc).

c. How to evaluate the components of the system/situation and the overall system/situational independent and dependent variables, hypotheses, and objective function.

d. Which parts of these computational tasks and decisions are best accomplished by machines and which by humans?

e. How and when to best communicate UQ-derived outcomes, propagating up and down between elements of the distributed system (human, agents, sensors, system).

In conclusion, UQ must be incorporated at every level of data processing and machine learning to move from the physical data to actionable information to augment human decision. It is not clear that current UQ methodologies and approaches are suitable and can be employed through the data-to-decision process. The following section discuss insights from UQ advances in multiscale, multidisciplinary modeling that can address these challenges.

2) Robust design: Can and how UQ be utilized to assess risk and robustness (with respect to uncertainty, adversarial actions, incomplete information, and low-probability and high-cost events) associated with a course of action/decision?

Advances in UQ applied through statistics and probability theory have made a significant impact spanning from the finance market to the nuclear industry. The tools are used to evaluate specific designs, courses of action, and public danger. These include CVaR (also known as superquantile risk), distributionally robust (data-driven) optimization, risk quadrangle (risk-informed modeling), and capacity-based UQ. These methods are particularly well-suited to problems for which a single objective can be readily quantified, QoI. Adapting these methodologies for anomaly detection leading to adversarial mitigation schemes and predicting low-probability/high-costs events is not trivial but is very promising (relevant to the Army's need for resilient command, control, communications, computers, intelligence, surveillance, and reconnaissance [C4ISR]-risk-adaptive optimization/anomaly detection).

Recent advances through a strategy in abstraction and surrogate modeling of multiscale complex systems have demonstrated methodologies to overcome computational cost time from research to development. The emergence of stochastic modeling for risk-informed design is paving the way for a new generation of beyond novel materials to improve the survivability and decrease the weight associated with our combat vehicles.

These computational advances in multiscale modeling and UQ are opportunities for inclusion, adaptation, and optimization with risk and trust algorithms for a set of reasonable courses of actions or decisions (risk-adaptive optimization). This is crucial to unforeseen scenarios, where objectives are not readily quantifiable. Robust human-augmented decisions can be the difference between victory and tragedy. Trust and the ability to articulate and communicate uncertainty relative to an acceptable level of risk are the keys to these human endeavors.

In conclusion, UQ is pervasive and foundational from finance to engineering industry for decades. Recent advances in multiscale modeling and computational sciences stimulated a new paradigm for discovery and design of complex systems. Advancing UQ in concert with risk and trust will elevate and accelerate capabilities for robust decisions. Attaining the capability to articulate situational awareness, robust courses of action, trust, and other relevant information in a manner to unburden the Soldier's command and control will be UQ's biggest impact.

3) Achievability at the tactical edge: What new methodologies are needed to allow UQ-based risk analysis and efficient/effective communication of risk to humans for decisions executed on the battlefield with limited edge computing capabilities, and heterogeneous, dynamic, and sparse computational resources?

Within the context of tactical edge today or the far future, any capability will be resource limited. This includes size, weight, power, computation, and time. In the previous discussions, the opportunities to apply UQ to enhance and enable an unprecedented level of the Soldier's augmented ability to adapt and win the unknown battles are evident. This capability is envisioned at the tactical edge only through co-development of hardware and software for speed-to-action on the battlefield.

Recognizing the opportunities from physics-aware modeling-to-machine learning-to-decisions, a holistic framework for integrating uncertainty with risk that balances parameters such as information, knowledge, resources, and trust is necessary. This demands quantifiable metrics for uncertainty and risk representations that complement models of human–machine interaction, including trust. Continuous learning and evaluation of this framework's uncertainty is necessary to account for information degradation, malicious information insertion, dense versus sparse information, and processing tradeoffs (overload/underload).

Finally, given the recent trends in evolving architecture space heterogeneous :FPGA+GPU+CPU+…etc., a new class of constrained computational problems that include building models, optimization and learning, performance-accuracy tradeoffs, and human–agent messaging that converges toward a hierarchy of decision making is needed. Architectural design needs to align with such hierarchy for speed and power efficiency, learning from the training environment prior to the battlefield. Conversely, a systematic process of evaluating and optimizing architecture capabilities depending on resource constraints will provide a dynamic and adaptive capability for all scenarios (relevant to the Army's need for Close Combat Capabilities – Hierarchy of Decision Making to match Architectural Resources – AI Understanding Humans). A key challenge to enabling support for Warfighters interacting with these systems is organic tools and methods in this framework that assist humans in understanding and shaping the policies, values, and preferences assumed by automated systems. The technology gap awaiting for discovery is the mechanism to interface and communicate computational outcomes within the context of the operational environment for human decisions. This will support new kinds of human–unmanned autonomous system collaboration with mixed-initiative interaction and augmented human cognition capabilities.

In conclusion, having augmented decision tools at the tactical edge requires the co-development of hardware and software optimized to implement capability needs and adaptive to austere-resource constraint environments without compromising speed. A novel framework is necessary to complement machine and human differentiation of trust, risk, and data.

5.2 Recommendations

In summary, UQ is pervasive across numerous engineering and human sciences domains. Advances and implementation of UQ in a number of far future Army science and technology needs include the following:

- Asymmetric Vision and Decide Faster – Training and Extrapolation – Cognitive Augmentation
- Resilient C4ISR – Risk – Adaptive Optimization/Anomaly Detection
- Close Combat Capabilities – Hierarchy of Decision Making to match Architectural Resources – AI Understanding Humans
- Materials by design and on-demand to lower combatant risk with improved survivability and decreased weight burden associated with our combat vehicles – Risk Informed Design – Stochastic Modeling

The following are strategic recommendations to harness the strength of UQ for Army priorities:

1. Capture low-hanging fruit; adapt, modify, and integrate into data analytics and information research:
   a. Combine Bayesian frameworks and generative adversarial networks (GANs) for anomaly detection.
   b. Apply nonlinear autoregressive schemes on multifidelity data in combination with dynamical systems theory (tracking manifolds) for robust inference or classification, avoiding deceptive data.
2. Develop self-optimizing framework UQ compatible with data/model-to-decision:
   a. Combine risk measures and uncertainty sources by including uncertainty in the objective functions and constrains, both for physical and information systems.
   b. Compare stochastic outputs using objective measures (e.g., superquantiles).

3. Develop methodologies that are suitable for continual co-design of hardware and software, and are adaptive to the needs of implementing and communicating UQ-enhanced Soldier-augmented decision.

5.3 Topic Participants

- Topic 1: **Machine learning** enabling (agility and adaptability)
- Topic 2: Advances in **physical and social models** (fidelity and predictability)
- Topic 3: Rapidly evolving and maturing **edge computing** (battlefield speed relevant)

Topic 1: Future readiness for adversarial detection, mitigation, and preemptive actions	**Topic 2: Future readiness for robust decision and design**	**Topic 3: Future readiness at the tactical edge**
Dr Patrick Langley, Institute for the Study of Learning and Expertise	Dr George Karniadakis, Brown U	Dr Emre Neftci, U of California, Irvine
Dr Brian Jalaian (CISD)	Dr Jaroslaw Knap, ARL/CISD	Dr Mary Anne Fields, ARL/VTD
Dr Mark S Dennison-ARL West	Dr Brent Kraczek, ARL/CISD	Dr Rajgopal Kannan, ARL/CISD
Dr Huan Liu, Arizona State	Dr Michael Shields, Johns Hopkins U	Dr Viktor Prasanna, U of Southern California
Dr David Barajas-Solano, Pacific Northwest National Laboratory	Dr Lirong Xia, Rensselaer Polytechnic Institute	Dr Michael Kirby, U of Utah
Dr Amit Roy-Chowdhury, U of California, Riverside (UCR)	Dr Drew Kouri, Sandia National Laboratories	Mr Craig Fullman, nVidia
Dr Krishnamurthy Srikanth, UCR	Dr Douglas Allaire, Texas A&M U	Mr Michael Meneghini, Army Capabilities Integration Center
Dr John O’Donovan, U of California, Santa Barbara	Dr Fariba Fahroo, DARPA/Air Force Office of Scientific Research	Dr Ernest Chin, ARL/CISD
Dr Roman Garnett, Washington U in St Louis	Dr Joseph Myers, ARL/ARO	Dr Brian Henz, ARL/CISD
Dr Katie Rainey, Space and Naval Warfare Systems Command, Navy		
Dr Lance Kaplan, ARL/SEDD	…	…
Dr Tien Pham, ARL/CISD		

5.4 Problem, Barriers, and Approaches

The following table summarizes much of the workshop discussions by organizing them as problems of multidomain battle, with corresponding technical limitations and barriers, and the possible UQ-based approaches to overcoming those limitations and barriers.

Relevance to multidomain battle	Limitations and barriers	Possible UQ approaches to overcome limitations and barriers
Asymmetric vision and decide faster (agility and adaptability)	Speed from transferring expertise and data-to-decision	Training and extrapolation (limited resources) from different environment
• Dinky, dirty little data	Communicate and propagate uncertainty up and down syst.	Continuum of superquantilies; adapt risk preference base on Soldier state
Resilient C4ISR networks (fidelity and predictability)	Robustness; adversarial; risk-to-course of action	Risk-adaptive design/optimization; physical model -> anomaly detection
• Dinky, dirty little data	Predicting low-probability/high-cost events	Buffer probabilities to quantify tail weight and probability; set-based (multimodel)-estimate bounds on prob.; information-theoretic methods
Close combat capabilities (battle-speed relevant)	Spine for continuous information (knowledge, trust, risk, degradation, malicious info., etc.) analytics	Edge computing - hierarchy of decision making and need to match architectural resources to hierarchy
• Local (tactical) assessment and decision	Distributed system optimization	Physics-aware statistical models; Bayesian neural networks
Future modernization - design	Robustness	Risk-informed design: eval. components and system (ind. and dep vari., hypotheses, obj. func., simu.); stochastic modeling

5.5 Agenda

The workshop was organized along the following agenda.

Time	Event	Topic lead
0800–0900	Registration (and continental breakfast)	
0900–0910	Welcome/introduction and purpose of the meeting	Dr Kott (via Phone)
0910–0930	(Overview of the Army multi-domain battle/UQ challenges)	Mr Michael A Meneghini
0930–0945	UQ technical context	Dr Tien Pham
0945–1000	UQ Topic 1: Future readiness for adversarial detection, mitigation and preemptive actions	Dr Brian Jalaian Dr Patrick Langley
1000–1010	Break	
1010–1025	UQ Topic 2: Future readiness for robust decision & design	Dr J Knap Dr George Karniadakis
1025–1040	UQ Topic 3: Future readiness at the tactical edge	Dr Mary Anne Fields Dr Emre Neftci
1040–1100	Breakout session guidance	
1100–1130	Breakout sessions (intro, frame of reference, development of discussion points)	Session facilitators
1130–1230	Group lunch	
1230–1500	Concurrent breakout sessions (discussion, conclusions, recommendations, etc.)	Session facilitators
1500–1515	Break	
1515–1600	Brief out: Topic 1 – Outcomes in priority	Dr Patrick Langley
1600–1615	Brief out: Topic 2 – Outcomes in priority	Dr George Karniadakis
1615–1630	Brief out: Topic 3 – Outcomes in priority	Dr Emre Neftci
1630–1700	Discussion/summary	Dr Chin/Knap
1700	Meeting adjourned	…

6. References

1. National Science and Technology Council (NSTC), Networking and Information Technology Research and Development Subcommittee. The national artificial intelligence research and development strategic plan. Washington (DC): NSTC; 2016 Oct [accessed 2018 May 9]. https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf

2. Establishment of an algorithmic warfare cross-functional team (Project Maven) [memorandum]. 2017 Apr 26 [accessed 2018 May 9]. https://www.govexec.com/media/gbc/docs/pdfs_edit/ establishment_of_the_awcft_project_maven.pdf.

List of Symbols, Abbreviations, and Acronyms

AI	artificial intelligence
AMRDEC	US Army Aviation and Missile Research, Development and Engineering Center
ARL	US Army Research Laboratory
ARO	Army Research Office
ASA(ALT)	Assistant Secretary of the Army for Acquisition, Logistics, and Technology
ASPSM	Army Science Planning and Strategy Meeting
C4ISR	command, control, communications, computers, intelligence, surveillance, and reconnaissance
CERDEC	Communications-Electronics Research, Development and Engineering Center
CISD	Computational and Information Sciences Directorate
CNN	convolutional neural network
CPU	central processing unit
CVaR	conditional value-at-risk
DARPA	Defense Advanced Research Projects Agency
DASA(RT)	Deputy Assistant Secretary for Research and Technology
DHS	Department of Homeland Security
DOD	Department of Defense
DSIAC	Defense Systems Information Analysis Center
EEG	electroencephalography
ERA	Essential Research Area
ERDC	Engineer Research and Development Center
fMRI	functional magnetic resonance imaging
FPGA	field-programmable gate array
GAN	generative adversarial network

GPS	global positioning satellite
GPU	graphics processing unit
HAT	human–agent teaming
I2WD	Intelligence and Information Warfare Directorate
ICoE	Intelligence Center of Excellence
IoBT	the Internet of Battlefield Things
IPB	Intelligence Preparation of Battlefield
IRB	Institutional Review Board
ISR	Intelligence, surveillance, and reconnaissance
IT	information technology
ML	machine learning
MURI	Multidisciplinary University Research Initiative
NITRD	Networking and Information Technology Research and Development
NVESD	Night Vision and Electronic Sensors Directorate
PII	personally identifiable information
QoI	quantity of interest
R&D	research and development
SEDD	Sensors and Electron Devices Directorate
SENSIAC	Sensing Information Analysis Center
SLAM	simultaneous localization and mapping
TTPs	tactics, techniques, and procedures
U	university
UAV	unmanned aerial vehicle
UCR	University of California, Riverside
UQ	Uncertainty Quantification
VTD	Vehicle Technology Directorate

V&V	validation and verification
WMRD	Weapons and Materials Research Directorate

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA	
2 (PDF)	DIR ARL IMAL HRA RECORDS MGMT RDRL DCL TECH LIB	
1 (PDF)	GOVT PRINTG OFC A MALHOTRA	
13 (PDF)	ARL RDRL D A KOTT RDRL VT B PIEKARSKI B SADLER RDRL VTV B GLAZ RDRL VTA M FIELDS RDRL ROP L L TROYER RDRL ROI N P IYER RDRL HRF BA P KHOOSHABEH RDRL CI B J HENZ RDL CIH C E CHIN J KNAP B KRACZEK RDRL CII B B JALAIA	