

Reports of the 2018 AAAI Fall Symposium

Aaron Adler, Prithviraj Dasgupta, Nick DePalma, Mohammed Eslami, Richard G. Freedman, John E. Laird, Christian Lebiere, Katrin Lohan, Ross Mead, Mark Roberts, Paul S. Rosenbloom, Emmanuel Senft, Frank Stein, Tom Williams, Kyle Hollins Wray, Fusun Yaman, Shlomo Zilberstein

■ *The AAAI 2018 Fall Symposium Series was held Thursday through Saturday, October 18–20, at the Westin Arlington Gateway in Arlington, Virginia, adjacent to Washington, D.C. The titles of the eight symposia were Adversary-Aware Learning Techniques and Trends in Cybersecurity; Artificial Intelligence for Synthetic Biology; Artificial Intelligence in Government and Public Sector; A Common Model of Cognition; Gathering for Artificial Intelligence and Natural System; Integrating Planning, Diagnosis, and Causal Reasoning; Interactive Learning in Artificial Intelligence for Human-Robot Interaction; and Reasoning and Learning in Real-World Systems for Long-Term Autonomy. The highlights of each symposium (except the Gathering for Artificial Intelligence and Natural System symposium, whose organizers failed to submit a report) are presented in this report.*

Adversary-Aware Learning Techniques and Trends in Cybersecurity

Machine learning–based (ML-based) intelligent systems are becoming ubiquitous in the technology surrounding our daily lives. Intelligent home surveillance systems, e-mail filtering and computer virus detection software, online merchandise recommender systems, social media feeds — all rely on automated ML-based systems to make automated decisions and interact with humans. With our increased reliance on such technology, how can we make these ML-based systems robust to security attacks from cyber adversaries, so that humans can use them safely and reliably? Our symposium posed this question as the adversary-aware learning problem in cybersecurity to the AI community.

The symposium featured 3 invited talks and 10 peer-reviewed research papers from academic, industry, and federal research laboratories. The first invited talk by David Martinez of the MIT Lincoln Laboratory discussed the canonical architectures used in AI and their evolution toward robust behavior. The talk culminated by emphasizing that adversary-aware ML systems should be designed while keeping in mind key performance assessment characteristics, including robustness, resiliency, responsiveness, scalability, and explainability. William Treadway of US Navy OPNAV N2/N6 presented the next invited talk, “Artificial Intelligence and Machine Learning Requirements for Naval Decision Superiority at the Tactical Level.” His talk included a novel classification of information and AI techniques along known versus unknown dimensions and encouraged the research community to direct efforts toward investigating the unknown dimensions. A novel approach of integrating principles of control theory to solve a constrained optimization formulation representing the adversarial learning problem was presented in the final invited talk by Xiaojin (Jerry) Zhu of the University of Wisconsin. Participants were particularly intrigued by the data camouflage problem presented as a novel challenge of adversarial learning in the talk. Two additional invited talks on adversarial learning by Matthieu Sinn of IBM Ireland and Tien Pham of the US Army Research Laboratory were attended by participants as part of a joint session with the AAI Symposium on AI in Government and the Public Sector.

The research papers presented at the AAI Symposium on Adversary-Aware Learning Techniques and Trends in Cybersecurity were grouped into three theme-based sessions. The first session, Adversarial Data Generation and Adversarial Training, included four papers that described theoretical and empirical results on techniques that an adversary could use to generate adversarial binary, character, or text data and addressed the performance of supervised learning-based ML algorithms against such adversarial data.

Countering Adversarial Attacks in Cybersecurity was the theme of the second session. Papers presented in the session described techniques including intelligent agents used to generate software triggers in response to cybersecurity threats, lexical link analysis in conjunction with big data visualization tools to identify hacked or hacking computers within a computer network, and an overview of inefficiencies in the cybersecurity exercise life cycle.

The third and final session was on Novel Approaches in Adversarial AI. The first paper presented an interesting approach to generate adversarial data by projecting low-power laser light onto small three-dimensional objects like toy vehicles, demonstrating that ML-based classifiers could be tricked by the projected light into misclassifying the original object. The two remaining papers in this session were on a coevolutionary framework for adversarial AI and coordination-driven learning for multiagent problem spaces.

The finale of the symposium was an open forum group discussion in which the participants discussed open issues, challenges, and potential research directions for adversarial AI. Topics that came to the forefront of discussions as future research topics included adversarial learning in human-machine teams, adversarial trust and deception, game theory techniques for modeling asymmetric behavior in adversarial learning, adversarial and cybersecurity issues in the context of the Internet of things, and building a repository of data sets focused on adversarial learning, similar to existing ML data repositories.

Overall, the symposium was very interactive, with attendees participating enthusiastically in Q&A following every presentation and in the group discussions. The symposium concluded with identification of a road map for future research problems based on the research themes evolved from the presented papers and with a plan to convene at a similar venue next year. The symposium papers are available online as CEUR Workshop Proceedings, Volume 2269, at ceur-ws.org/Vol-2269.

The symposium was chaired by Joseph B. Collins, Prithviraj Dasgupta, and Ranjeev Mittu.

Artificial Intelligence for Synthetic Biology

Synthetic biology integrates biology and engineering—mixing theoretical and experimental biology; engineering principles; and chemistry, physics, and mathematics. As the fields grows it is apparent that there are many opportunities as well as a need to apply AI techniques to several complex problem areas in the field. The AAI symposium on Artificial Intelligence for Synthetic Biology was a way to bring the two communities together. The symposium consisted of a mix of 13 technical talks, 3 invited talks, 3 discussion sessions, and a government panel.

Participants in the symposium had different backgrounds in the two fields. Two invited talks lay the foundation for the technical talks and discussions. Eric Young (Worcester Polytechnic Institute) gave the introduction to synthetic biology, and Hector Munoz-Avila (Lehigh University) gave the introduction to AI. Ron Weiss (MIT) gave the keynote talk, which covered programmable organoids and a discussion of modularity, a key idea in computer science but not a natural property of living systems.

The symposium included a government panel that discussed funding opportunities, focused on health and defense, at the intersection of AI and synthetic biology. Panelists included representatives of the Defense Advanced Research Projects Agency, the National Science Foundation, the National Institutes of Health, and the Edgewood Chemical Biological Center. Discussions included opportunities to apply AI to create, manipulate, or optimize genetic circuits, to apply AI to high-fidelity experimental data sets, and

to apply AI or ML to transfer between biological models and show that predictions are accurate.

Talks addressed existing synergies between the fields from academia, industry, and government that described variants of the design-build-test-learn cycle. AI techniques that were used included neural nets, deep learning (for example, as applied to cell-free systems), active learning, Bayesian optimization, unsupervised learning techniques, planning, and semantic state models. These techniques were applied to a variety of data sets, such as instrument properties, simulated models, gene expression data, and sequence information. Challenges in applying such models to the domain were also presented. Some of the speakers, members of the synthetic biology industry, brought the unique perspective of focusing on the practical and cost-effective applications of AI in optimizing synthesis of products through synthetic biology.

The symposium included three discussion sessions. The first centered on identifying the big, hard problems in synthetic biology. Identified challenges include data processing at scale, lack of quality data and metadata, outlier detection, and the need to store negative results. Other challenges were in knowledge gaps in mapping DNA to its function, in transferring results between model systems, and in predicting biology. A lack of trust in ML, the need for explainability of computer suggestions, the need for knowledge in multiple fields, and the need for controlled and repeatable experiments were also discussed. The second session included highlights of AI expertise of attendees and how the problems from the first discussion might be addressed by AI techniques. The discussion centered on the need for high-quality data. Data collection, repositories, standards, and incentives were discussed, along with suggestions for various test cases. The final discussion addressed ethics issues in AI and synthetic biology. Discussion topics included dual-use concerns, gender balance of the fields, boundaries or parameters for research, and ensuring sufficient upstream public engagement.

The symposium concluded with a discussion of next steps, target publications, and future meeting venues. Aaron Adler (BBN Technologies), Mohammed Eslami (Netrias), Jesse Tordoff (MIT), and Fusun Yaman (BBN Technologies) served as cochairs of the symposium. Some papers and talks are available on the symposium website, www.synbiotools.com/ai-for-synbio-fss-2018.

AI in Government and Public Sector Applications

AI adoption in government and the public sector faces unique challenges and opportunities, including a higher standard for transparency, fairness, explainability, and operations without unintended consequences. Keynotes, panels, and formal presentations are summarized here; and innovative contributions

addressing these needs are published in the conference proceedings.

This was the fourth year that the AAAI symposium on AI in Government and Public Sector Applications has been held. Over time, we have seen a transition in the talks from conceptual and largely academic presentations to the practical reality of operating AI in government and public sector applications.

As we get beyond considering only the technology challenges, governance of AI operation has become an important issue. An explicit theme of the 2018 symposium was how government and public sector domains are unique. There is a wealth of publicly owned data, although access needs strict oversight; compared with industry, there is a lack of computing resources; and the consequences of failure are serious, and risk tolerance is conservative. Nevertheless, there is a deeply rooted culture of accountability and transparency — and a strong interdisciplinary culture — that arguably puts this community in a prime position to address areas such as bias mitigation, safety, and human+computer teaming challenges.

Lynne Parker (White House Office of Science and Technology) opened the symposium with a look at the government's AI research and development strategic plan. She discussed some of these strategic areas, including AI's impact on the workforce; how to design ethical, safe, and trustworthy AI; AI's role in cybersecurity; and building an AI workforce. Justin Herman (General Services Administration) talked about lowering the barriers to AI adoption in government by applying AI for information technology modernization, establishing a federal data strategy, establishing partnerships with industry and academia, and providing more ways to test and evaluate AI. Tien Pham (US Army Research Laboratory) discussed the challenges of developing AI for their unique environment including complex data types and the resource-constrained tactical edge.

The international keynote address was given by Gavin Pearson (UK Defence Science and Technology Laboratory). He identified issues that are blocking UK defense from fully benefiting from AI, setting these in the context of a systems reference model for the AI value train.

A joint session was held with the AAAI symposium on Adversary-Aware Learning Techniques and Trends in Cybersecurity. Mathieu Sinn (IBM) discussed practical defenses against adversarial threats to AI. He described how deep neural nets don't actually learn to recognize objects but learn instead to discriminate among objects in a training set in an optimized, but not robust, manner. Adversarial attacks work by pushing the data inputs across a nonrobust learned decision boundary. Jerry Zhu (University of Wisconsin), from the Adversary-Aware Learning Techniques and Trends in Cybersecurity symposium, joined Mathieu Sinn and Tien Pham for a lively panel discussion on vulnerabilities, trust, and computer security threats.

The symposium included panel discussions on growing and retaining AI talent for the US government,

on AI and policy, and toward a mission-based research road map to address unwanted bias.

The participants agreed that they would like to attend future symposia to share experiences and address some of the challenges posed.

Frank Stein (IBM) and Chuck Howell (MITRE) served as cochairs of this symposium. The organizing committee included Lashon Booker (MITRE), Alun Preece (Cardiff University), Michael Garris (NIST), Mihai Boicu (George Mason University), Shali Mohleji (IBM), and Jim Spohrer (IBM). Session papers were posted at arxiv.org/abs/1810.06018.

A Common Model of Cognition

The AAAI fall symposium on a Common Model of Cognition followed up on the results of the AAAI 2017 fall symposium on a Standard Model of the Mind, which was itself inspired by an analogy with the notion of a standard model in physics as a consensus on a scientific domain that is internally consistent yet may still have major gaps. In the standard model symposium, the aim was to reach a consensus on the data and models — in particular, the integrated models expressed as cognitive architectures — that inform us about the structures and processes underlying cognition. This effort has grown into a larger online community, with a change of name that was driven by community consensus, and online working groups have arisen covering the following topics: (1) procedural and working memories; (2) declarative memory; (3) metacognition and reflection; (4) language processing; (5) emotion, mood, affect, and motivation; (6) higher-level knowledge, rational and social constraints; (7) lower-level neural and physiologic constraints; and (8) perceptual and motor systems. The intent of these working groups is to develop a statement of the best consensus in each area given the community's current understanding of these components of cognition and how they fit together. The primary goal for this second symposium was to provide a forum focused on extending the model based on the progress made in the working groups, while engaging new participants in the process.

The symposium focused on two distinct types of sessions. The first involved presentations based on a subset of the papers accepted to the symposium, falling roughly into the following sets of topics: introduction, formalism, and validation; higher (that is, rational and social) bands and metacognition; attention, physiology, and emotion; and knowledge, memory, and language. The papers and presentations in large part focused on what is missing from or wrong about the current formulation of the common model, with most reflecting the particular perspectives of individuals or small groups of researchers. However, three were interim progress reports from working groups that reflected at least the beginnings of broader consensus on metacognition; emotion; and higher-level knowledge, rational, and social levels. There was also an invited keynote presentation that explored a

number of intriguing analyses of the range of cognitive architectures (more than 180) that have been developed over the past four decades.

The second type of session involved parallel breakout groups that were immediately followed by plenary discussions. These were true working sessions in which each breakout group was able to focus on progress toward identifying potential loci of consensus within a single topic, with immediate feedback then available during the subsequent plenary discussion. Each topic appeared in two different sessions of the symposium. In this manner, we hoped to initiate a discussion of each topic, while providing feedback from the full symposium along with time to think further about the topics, before holding a second discussion and feedback period. Planning for the symposium began with the assumption that a breakout group would be appropriate for each of the eight working group topics, but after a survey of the community this was compressed to six topics: language, metacognition, procedural memory and perceptual-motor behavior, declarative memory, higher bands, and attention, emotion, and neural physiology.

Although these sessions did not by themselves lead to specific extensions being made at the symposium to the current draft of the common model, they did highlight and structure a number of important possibilities for further consideration. The last session was then devoted to discussing appropriate next steps for the community toward both extending the common model and expanding the community involved with its development, with the primary focus on the nature and timing of the next opportunity for the community to meet physically rather than virtually. As organizers, we were particularly pleased with the continued level of enthusiasm expressed during this session for pursuing a community-driven consensus toward comprehensive models of the mind.

Paul S. Rosenbloom, John E. Laird and Christian Lebiere served as cochairs of this symposium. Most of the accepted papers were published in volume 145 of *Procedia Computer Science*.

Gathering for Artificial Intelligence and Natural System

The Gathering for Artificial Intelligence and Natural System at the AAAI fall symposium was organized by Ioana Baldini (IBM Research AI), Richard “Doug” Riecken (Air Force Office of Scientific Research), Prasanna Sattigeri (IBM Research AI), and Vikram Shyam (NASA Glenn Research Center). No report was submitted by the organizers.

Integrating Planning, Diagnosis, and Causal Reasoning

Planning, plan execution, diagnosis, and causal explanation have each been examined by various research

efforts, but little attention has been paid to how to integrate them within a single system; when the unexpected happens, how to adapt models for future success; or how best to respond to unexpected events. The AAAI fall symposium on Integrating Planning, Diagnosis, and Causal Reasoning brought together researchers to explore these questions, with four challenge talks about integrated systems and seven paper presentations on relevant themes.

Christophe Guettier (SAFRAN) presented Planning and Safety Challenges in Autonomous Driving Systems. J. Benton (NASA Ames) presented Autonomous Air Vehicles. Mark Micire (NASA Ames) presented Distributed Spacecraft Autonomy. Jeremy Frank (NASA Ames) presented The Europa Lander Mission: A Space Exploration Challenge for Autonomous Operations. We encouraged extended conversations about the themes with small group brainstorming and large group discussions. Next we highlight themes related to models, system integration, problem solving, evaluation, and human interaction.

Models define the representation used for problem solving. The biggest challenges for integrating planning and diagnosis result from their differing models. Model fidelity between such models may mismatch, exacerbating the adage “all models are wrong, some are useful” because model mismatch is not commonly researched. Integrating models with different input and output types and parameters is a substantial verification and validation challenge for integrating models. Early model integration reduces risk, where analyzing interfaces between models or components can identify not just differences in syntax (that is, Boolean true versus 0/1) but also differences in semantics. Ideally, a living document should describe what is designed, implemented, and maintained over time, reflecting how models may change and interrelate.

Mapping between models is not commonly considered with respect to execution monitoring. During execution, there is merit in maintaining the independent state for each model, but potential exists for better mechanisms and monitoring multiple models at varying fidelities. After execution, the mapping between models (and individual models) can be refined from observed traces, where errors result from mismatches from the observed and predicted traces.

System integration challenges differ from model integration; it is not clear how to determine what one component needs from another. What kind of feedback about a fault, or its consequences, does a planner require to replan? How much detail from a plan should be provided to the execution system? These questions can be resolved by defining and maintaining design principles that reduce the risk of future project phases with a clear distribution of responsibility across planner, executive, system health management, and all other planning/execution assets, such as lower-level software (for example, hardware controllers), support software (for example, additional domain-specific planners), and, if relevant, additional planners/executives on other systems.

Evaluation and testing of integrated systems poses another major hurdle. How do we verify that the developed system behaves like the designed system? While validation and verification can be accomplished through scenario testing and stress testing to ensure timely and well-behaved execution, it can be difficult to answer, “What is good enough?” Unlike human-operated deterministic systems, the threshold for success of autonomous planner and executive-operated systems has little precedent. Defining a priori thresholds will promote acceptance and deployment of autonomous systems.

Consideration of planning and diagnosis systems’ interactions with humans included topics such as explanations, detecting and accounting for the differences between the user’s model and that used by the system, and how to design systems accomplishing these objectives. Topics such as mixed-initiative systems and user interfaces were elaborated on by more nuanced and broader ethical considerations and questions. Can a diagnosis system detect or manage cognitive impairment on the part of users? If so, when should it intervene? Should they be rebel agents (for example, acting against human wishes in the interests of safety)? Even when humans are not impaired, human cognition is limited; how do we design interaction between computers and humans in light of these limitations?

The symposium was organized by Jeremy Frank (NASA), Matt Molineaux (Wright State), and Mark Roberts (Naval Research Laboratory). Summary contributions were provided by Christian Muise, Rashied Amini, Michael Rubin, and Shakil Khan. Further information can be found at the symposium website: makro.ink/sip/sip18.

Interactive Learning in Artificial Intelligence for Human-Robot Interaction

The fifth AAAI symposium on Artificial Intelligence for Human-Robot Interaction was held in October 2018 under the theme of interactive learning. This symposium provides a gathering place for researchers working at the intersection of the fields of AI and human-robot interaction (HRI) — an interdisciplinary area that historically has presented unique challenges. Accordingly, the previous iterations of the symposium respectively focused on (1) creating a venue for work at this intersection, (2) improving interactions between the AI and HRI communities, (3) critically analyzing the nature of work conducted at this intersection, and (4) presenting new challenges for the AI and HRI communities. The 2018 symposium focused on one specific research challenge at this intersection, with the intention of attracting new attendees to the symposium and holding more focused research-oriented discussions rather than community-oriented discussions.

The chosen focus topic was interactive learning: how robots can interact with humans to learn online.

Interactive learning differs from classic ML by using numerous small updates to the behaviors and actively involving both the learner (the robot) and the teacher (the end user) in the learning process. Interactive learning holds significant promise for the fields of AI and HRI as it aims to give access to ML tools to a wider range of the population and to adapt the outcome of the learning process more precisely to the users' needs.

Twelve papers on the topic were presented at the symposium: seven from US universities, two from European universities, one from an Australian university, and two from the US Army Research Laboratory. The presentations covered many relevant topics, including advances in human-in-the-loop ML methods, cycles that switch between learning approaches within an interactive system, applications and metrics for AI-HRI, sensor simulations for data set generation, and socially aware autonomous robot behavior. Attendees also participated in breakout discussions that highlighted not only central questions as to what can and should be learned but also higher-level concerns surrounding trust and privacy that arise when interactive learning techniques are used in interactive robots. There was also discussion about whether the current metrics used in AI and HRI research are sufficient when studying works at the intersection of the fields, which led to the realization that most existing data sets for learning fail to capture the diversity of environments and human behaviors that are vital to interaction.

One of the most consistent themes that emerged from the symposium was the importance of learning from many of the same communication modalities used by humans. This approach often results in multimodal interaction models that support identifying and generating speech, touch, and movement. For instance, three papers were written to support linguistic communication as well as other modalities; one was written to support touch-based interactions, which was further supported by one of our keynote presentations that highlighted this common form of human communication; one was written to support keyframe learning by demonstration; and two were written to recognize user intent to support interactive learning. A common challenge in these interactive learning settings is user fatigue, which can occur shortly after interactions begin. One hypothesis was to focus on cognitive load and user experience to alleviate and extend interactive demonstrations. One of the talks focused on the adaptability of the users with whom the robot interacted; it was shown that humans are learning and adapting to the robot as well. One study focused on the idea of passive demonstrations used by the robot to signal its intent to the user. It was found that the human partners adapted to these signals and increased fluency through repeated interactions with the robot.

Interspersed with these presentations and discussions were multiple invited talks. Invited talks were given by Sonia Chernova (Georgia Institute of Technology), Cynthia Matuszek (University of Maryland,

Baltimore County), Dylan Glas (FutureWei Technologies), and Greg Trafton (Naval Research Laboratory). Sonia Chernova discussed not only the benefits to robustness that come with interactive learning but also the challenges that a focus on learning may present to robots' reasoning capabilities and their overall usability. Cynthia Matuszek discussed the benefits of learning directly from language and how both supervised and unsupervised learning techniques can be leveraged for the purposes of robust language grounding. Dylan Glas reviewed a variety of approaches to developing intelligent interactive robots, basing them on the principles of designing versus learning behaviors as well as symbolic versus concrete representations. This review included a variety of examples for robots learning to be proactive store clerks by observing experts, identifying context through conversations, and presenting their own "personality" through intention-based planning. Greg Trafton explored a recent case study involving the development of robots that can work in teams alongside humans in firefighting scenarios; the story discussed the challenges of not only making the robots effectively responsive to their human peers but also implementing communication methods that are already intuitive to firefighters so that these experts do not need to learn anything new for practical interaction.

The participants were supportive and enthusiastic to discuss the current works and upcoming challenges for interactive learning in AI-HRI. In addition to the identified issues and questions for the AI-HRI community, an open challenge was presented at the plenary session to the AI community as a whole to develop simulations of human behavior that would allow more effective evaluations of HRI systems without the need to recruit many human subjects. Although live testing is important before deployment, there is a lot of work that goes into initial evaluations during development of intelligent interactive robotic systems; like robotics, performing virtual tests in simulation has many benefits before performing physical tests in the real world.

The organizers of this symposium were Kalesha Bullard, Nick DePalma, Richard G. Freedman, Bradley Hayes, Luca Iocchi, Katrin Lohan, Ross Mead, Emmanuel Senft, and Tom Williams. The proceedings were uploaded to arxiv.org/html/1809.06606.

Reasoning and Learning in Real-World Systems for Long-Term Autonomy

Over the past decade, decision-making agents have been increasingly deployed in industrial settings, consumer products, health care, education, and entertainment. The development of drone delivery services, virtual assistants, and autonomous vehicles has highlighted numerous challenges surrounding the operation of autonomous systems in unstructured

environments. These challenges include mechanisms to support autonomous operations over extended periods of time, techniques that facilitate the use of human assistance in learning and decision making, learning to reduce the reliance on humans over time, addressing the practical scalability of existing methods, relaxing unrealistic assumptions, and alleviating safety concerns about deploying these systems.

The AAAI fall symposium on Reasoning and Learning in Real-World Systems for Long-Term Autonomy consisted of 18 paper presentations and 3 invited talks, concluding in a lively and interactive panel discussion moderated by Joydeep Biswas. In total, there were 12 long papers and 6 short papers, ranging in topic from planning and learning to architectures and real-world systems. The three invited talks by Nick Hawes, Maarten Sierhuis, and Peter Wurman presented work with fully operational deployments of assistant service robots, semiautonomous vehicles, and large-scale multiagent warehouse robots, respectively. The techniques leveraged across the papers and talks included hierarchical and multiobjective (PO)MDP models; reinforcement learning with general value functions; deep learning for grasping, environment understanding, and risk-aware planning; multiagent models for system robustness; and robotic architectures with a focus on tight component integration. Applications included autonomous vehicles, delivery robots, activity recognition in smart homes, mobile warehouse robots, air traffic surveillance, dual-arm grasping robots, and mobile home-health-care robots.

Throughout the symposium, four key themes emerged as topics for long-term autonomy research: (1) integration of multiple AI components beyond traditional architectural, hierarchical, and multiobjective approaches; (2) methods to proactively leverage humans to overcome any exceptional issues encountered, diminishing this reliance over time; (3) standard metrics and verification methods to properly measure the effectiveness of long-term autonomous agents, such as by their exceptional issues encountered, number of human help requests, effect of system improvements made, and degree of learning performed; and (4) a focus on the robustness of the system to enable these long-term deployments.

Conclusions drawn during the panel discussion at the end of the symposium suggested long-term autonomous systems benefit greatly from a symbiotic collaboration with other connected agents, humans, and a cloud-based AI central support system. To be sufficiently robust also requires a much tighter development of the theoretical frameworks with the implementation itself. Finally, objectively evaluating the system continuously over time and across many metrics is crucial, both as it is developed and as it autonomously learns. Evaluation can be done using general metrics, such as how many tasks were completed, which tasks were completed and their completion times, how many failures occurred, and what kinds of failures occurred and their failure times.

In addition, specific domain-related metrics can improve this measurement, such as how far an autonomous vehicle has driven or how many objects per day a robot grasped. Such an array of metrics allow us to confirm that the holistic AI system is robust and capable of long-term autonomy.

The symposium was organized by Kyle Hollins Wray (chair), Julie A. Shah, Peter Stone, Stefan J. Witwicki, and Shlomo Zilberstein. The papers from the symposium were published by the University of Massachusetts Amherst as Technical Report UM-CS-2018-009.

Aaron Adler is a senior scientist at BBN Technologies in Columbia, Maryland.

Prithviraj Dasgupta is a professor in the Computer Science Department at the University of Nebraska, Omaha.

Nick DePalma is a research engineer at Samsung Research of America.

Mohammed Eslami is the chief data scientist at Netrias in Arlington, Virginia.

Richard G. Freedman is a researcher at Smart Information Flow Technologies (SIFT) and a PhD candidate at the University of Massachusetts Amherst.

John E. Laird is the John Tishman Professor of Engineering in the Division of Computer Science and Engineering at the University of Michigan.

Christian Lebiere is a research faculty member in the Psychology Department at Carnegie Mellon University.

Katrin Lohan is an associate professor of computer science at Heriot-Watt University.

Ross Mead is the founder and CEO of Semio AI.

Mark Roberts is a researcher at the US Naval Research Laboratory.

Paul S. Rosenbloom is a professor in the Department of Computer Science and director for cognitive architecture research at the Institute for Creative Technologies at the University of Southern California.

Emmanuel Senft is a research fellow at the University of Plymouth, United Kingdom.

Frank Stein is the director of the A3 Center at IBM.

Tom Williams is an assistant professor of computer science at the Colorado School of Mines.

Kyle Hollins Wray is a graduate student in the College of Information and Computer Sciences at the University of Massachusetts Amherst.

Fusun Yaman is a senior scientist at BBN Technologies in Cambridge, Massachusetts.

Shlomo Zilberstein is a professor and associate dean of research and engagement in the College of Information and Computer Sciences at the University of Massachusetts Amherst.