

# Reports of the AAAI 2017 Fall Symposium Series

*Arjuna Flenner, Marlena R. Fraune, Laura M. Hiatt, Tony Kendall,  
John E. Laird, Christian Lebiere, Paul S. Rosenbloom, Frank Stein,  
Elin A. Topp, Vaibhav V. Unhelkar, Ying Zhao*

■ *The AAAI 2017 Fall Symposium Series was held Thursday through Saturday, November 9–11, at the Westin Arlington Gateway in Arlington, Virginia, adjacent to Washington, DC. The titles of the six symposia were Artificial Intelligence for Human-Robot Interaction; Cognitive Assistance in Government and Public Sector Applications; Deep Models and Artificial Intelligence for Military Applications: Potentials, Theories, Practices, Tools, and Risks; Human-Agent Groups: Studies, Algorithms, and Challenges; Natural Communication for Human-Robot Collaboration; and A Standard Model of the Mind. The highlights of each symposium (except the Natural Communication for Human-Robot Collaboration symposium, whose organizers did not submit a report) are presented in this report.*

## Artificial Intelligence for Human-Robot Interaction

Over the last several years, a group of researchers have established the notion of artificial intelligence (AI) for human-robot interaction (HRI), and HRI as a research field lingering somewhere between the core topics typically discussed at AI or HRI venues. While most current, traditional HRI research involves investigating ways for robots to effectively interact with people, HRI's overarching goal is to develop robots that are intelligent, autonomous, and capable of interacting with, modeling, and learning from humans. These goals greatly overlap with some central goals of AI; however, much AI research takes place without consideration of humans, and so it does not consider the inherent uncertainty in dynamics, structure, and interaction that humans and human-populated environments bring with them. We therefore believe that HRI is an extremely interesting problem domain to challenge and refine with research in AI and robotics.

Despite the importance of research in AI for HRI, however, it has been traditionally difficult for work in this intersectional area to be welcomed by the underlying communities. For the last four years, the AI-HRI AAAI fall symposium has attempted to address this difficulty by bringing together researchers in the area to discuss technical results, to support research in this area, and to build up the AI-HRI community. The symposium had a number of invited talks that discussed connections between AI and HRI, including how to use AI methods as tools in HRI, how observations made in HRI research could inform models used in AI systems, and how interactions between people and robots can be modeled in an AI system. The talks spanned domains from interacting with robot swarms to improving robot autonomy and addressing social HRI. The talks also gave very clear statements of the open challenges in their respective areas, which provided younger researchers attending the symposium ideas on how to frame their contributions in the future.

In the presented papers, posters, and also in the intermittent discussion slots, two topics frequently occurred: deep learning and explainability. Deep learning was discussed in the context of whether it was helpful in the given context or not, and also how to use deep learning in one's work while also recognizing that deep learning is only a small subset of AI. The topic of explainability was discussed in the context of reflections on the cognitive aspects of HRI (how humans explain their actions, and what that would entail for a robot interacting with humans), as well as the technical aspects of explainability in AI.

Another central issue discussed was how AI-HRI researchers could develop HRI-related challenges for the AI community (and vice versa). The conclusion of this discussion was that there is a need for the acceptance of more contributions considering complete systems in real-world conditions in the AI community, as well as a need for more AI-related technical contributions in the HRI community. The symposium concluded with a discussion of further venues, formats, and activities for the now seemingly quite established AI-HRI community.

Laura M. Hiatt and Elin A. Topp served as cochairs of the symposium, which was also organized by Kalesha Bullard, Emmanuel Senft, Tian Zhou, Luca Iocchi, Marc Hanheide, Tom Williams, Frank Broz, Katrin Lohan, Ross Mead, and Dan Grollmann. The papers of the symposium were published as AAAI Press Technical Report FS-17-01.

## Cognitive Assistance in Government and Public Sector Applications

The goal of the Cognitive Assistance in Government and Public Sector Applications AAAI symposium was to present state-of-the-art cognitive assistant projects and to expand the dialog within the community on

building, accepting, and using these systems. During this symposium, innovative contributions were solicited on both the research and development and the application of cognitive assistance technology for use in government, education, and healthcare. The public sector is especially challenging because human experts must balance legal, social, and ethical requirements while delivering accurate and timely service.

Building on the success of the 2016 symposium, we worked to increase understanding of current use cases and what the future might look like. Last year's presentations largely dealt with getting systems deployed in specific professions and domains that included law and the courts, patent processing, teaching, health, cybersecurity, aviation, air traffic control, defense, and intelligence. This year, the emphasis shifted to include realized value and challenges.

Opening keynote speaker Mark Maybury (MITRE) observed that humans guiding AI systems — like a virtuoso playing a Stradivarius — can result in unprecedented performance. But there are also risks, such as AI-enabled fake news. Research topics included confidence versus control in trust environments, human-machine trust, security, autonomy, transparency, human-robot interfaces, and verification or validation.

William Regli's keynote reviewed DARPA investments in a new type of thinking that is a joint human-machine effort to solve problems and advance society. The advancement of humanity from simple tools to intelligent machines, with increasing output per person on the planet, was discussed, as well as how DARPA is advancing this trend.

Day two opening keynote speaker Kenneth Forbus (Northwestern) presented a talk on making computers that can keep up with humans. He has created software collaborators that act as aides-de-camp. Future research includes commonsense reasoning, rapid learning from few examples as people do using analogies, and systems that pursue their own learning (that is, using expectation violation to revise internal models).

Chieko Asakawa from IBM presented smartphone-based cognitive assistance applications to help blind users actively participate in the real world. A current impediment to navigation for the blind is the lack of open, standards-based maps and sensors for both indoor and outdoor areas. Jim Spohrer, also from IBM, discussed a look to the future, in which he sketched an optimistic scenario enabled by Moore's Law and AI that postulated the advent of narrow digital workers comparable in cost to human workers in 20 years and giving way to digital workers at much lower costs in about 40 years.

The session on pilot assistants included Randy Bailey (NASA), who introduced the automation paradox, where an autopilot in the cockpit masks

deskilling of the workforce until a failure occurs. Kevin Burns (MITRE) presented the results of a reliability model showing the potential safety impact of a cognitive copilot and what features would provide the most benefit.

The session on cybersecurity included examples by Gheorghe Tecuci (George Mason University) and Jim Whitmore (formerly of IBM), followed by a lively discussion concluding that the volume of cyber attacks could force the move toward automation even if cogs don't perform equal to human experts.

There was a session on natural language processing for US federal acquisition regulations for adjudicating administrative law and one on using NLP to monitor social media for situational understanding. We also had presentations on cognitive assistance for dealing with autism, helping with housing for homeless youth, and education.

The symposium audience asked if AI can be used to generate fake images, videos, and blogs, how will we know truth from fiction? If AI identifies a new drug, who gets the credit? If AI misidentifies a disease, who gets the blame? How do we adapt the workforce to working with cognitive assistants as collaborators and partners? The participants agreed that they would like to attend future symposia to share experiences and address some of the challenges posed this year.

Frank Stein (IBM) and Chuck Howell (MITRE) served as cochairs of this symposium. The organizing committee included Lashon Booker (MITRE), Chris Codella (IBM), Eduard Hovy (Carnegie Mellon University), Anupam Joshi (University of Maryland, Baltimore County), Andrew Lacher (MITRE), Jim Spohrer (IBM), and John Tyler (IBM). The session papers were published as AAAI Press Technical Report FS-17-02.

## Deep Models and Artificial Intelligence for Military Applications: Potentials, Theories, Practices, Tools, and Risks

Recent advancements in artificial intelligence enable new technologies to assist modern warfighters by automatically analyzing big data at timescales much faster than a human can achieve. Deep learning (DL) is the core of the new AI revolution, demonstrating that not only can machines classify quicker than humans, but they can also classify more accurately than humans. These technologies have revolutionized many commercial applications but are not designed to solve military problems.

Fundamentally, the field of machine learning (ML) seeks to learn the parameters of a function, given a data set. DL refers to a subfield of machine learning that consists of a large number of parameters (for example, deep models) for accurate classifications or

predictions (for example, convolutional neural networks [CNN]). DL was initially demonstrated in the breakthrough results for supervised learning of machine vision applications. Academic and industrial DL, ML, and AI are active in the applications of machine vision, speech recognition, chat, and autonomous driving.

Four of the main challenges in military applications are a lack of adequate samples for classification tasks (or none at all), short timescales for learning, fewer computational resources, and adversarial behavior. In our workshop context, deep models are broadly defined as all analytic models that can handle big data or no data at all and that can perform ML and AI. Our goal is to collaborate and form communities to constantly discuss the potentials, theories, practices, tools, and risks of deep models for military applications in order to remain competitive in technical leadership and innovation. We want to maximize the present and future impact of our approaches by reaching out to the strategic thinkers and leaders of our nation.

The capability to learn from data and experience is the most critical element of human intelligence. Data-driven deep models, including DL, ML, and AI, provide tremendous potentials for modernizing military applications. Data fusion and data correlation algorithms can discover associations among distributed and heterogeneous data. In the panel discussion, we discussed, since everything can be represented as a graph or network, the theories that can lead to innovative ways to store data more efficiently, speed up information retrieval, and reduce latency when applying deep models in real-time learning and adaptation, such as in reinforcement learning. Network analysis can be also used for the visualization of deep models such as the data-driven documents (D3) visualization presented in a paper given by Quinn Halpin et al. (Cornell University). This panel discussion was motivated by Raluca Gera (Naval Postgraduate School), who presented an invited talk on ranking nodes in a network.

ML and AI will lead to the discovery of previously unknown insights from multiple distributed and heterogeneous data sources, as demonstrated in the paper by Ying Zhao et al. (Naval Postgraduate School) on lexical link analysis, and with broadcast (ADS-B) data in the talk by Tony Kendall (Naval Postgraduate School).

Important for DOD applications are autonomous systems and classification. A paper by Richard Wu et al. (University of Massachusetts) demonstrated deep and reinforcement learning towards self-learning of moving autonomous objects in a virtual environment. In the area of autonomous systems, dealing with adversarial behavior is an important concept, and game theory provides a mechanism to address adversarial behavior that provides robustness to learned models. Prithviraj Dasgupta et al. (Universi-

ty of Nebraska) addressed this issue using Stackelberg game models. The papers of Arjuna Flenner (NAVAIR, China Lake) and Deanna Needle et al. (UCLA) discussed using graphs to improve classification and using compressed sensing techniques to obtain theoretical results on single-bit classification, respectively.

In a military context, ethics equates to an operational commander taking responsibility, which requires formal understanding of the micro- and macro-level aspects of autonomous vehicle tasking. A paper given by Don Brutzman et al. (Naval Postgraduate School) addressed rational behavior using human-robot ethical constraints and a paper given by Bonnie Johnson (Naval Postgraduate School) addressed a system of systems approach to battle management aids.

Traditional data sciences, including statistics, numerical analysis, machine learning, data mining, business intelligence, and artificial intelligence, have evolved into big data analytics and deep models. We discussed Amazon's AWS, the industry cloud computing leader's big data platform, which includes the ingestion tools Snowball, S3 storage, Glue Data Catalog, Kinesis analytics, and managed Hadoop/Spark. We also discussed industrial infrastructure tools such as GPU and IoT, and the analytic engines of Apache tools, Caffe, Theano, TensorFlow, Keras, and Torch.

The presentations by Wallace Bow et al. (Sandia National Laboratories), Philip Chan et al. (University of Maryland Baltimore County), and Krishnendu Ghosh et al. (Miami University) and the invited talk by Roshan Punnoose (Enlighten IT Consulting) explore these ideas in more detail.

Deep models, ML, and AI will become the lifeblood of military applications. But with opportunity can come risk. Can AI be trusted? AI can be weaponized and data can be poisoned. However, opportunities are plenty if we foster broader communities and collaboration. Inscrutability is inevitable, as the system of systems approach becomes more complex or in situations where human intelligence is not easily understood. Risks exist on autonomous systems against humans, the weapons of mass destruction. A paper given by Scott Humr (Marine Corps University) discussed autonomous outcomes on shaping the future data environment to build trust in artificial intelligence and learning applications.

Ying Zhao, Arjuna Flenner, and Tony Kendall served as cochairs of this symposium. The papers of the symposium were published as AAAI Press Technical Report FS17-03.

## Human-Agent Groups: Studies, Algorithms, and Challenges

As robots and artificial agents become more prominent in human lives, they are also increasingly

becoming parts of groups and teams. Group interaction of humans and agents are present in a diverse set of AI applications — for instance, a digital assistant for the home, a social robot operating in a mall, or a group of robots and virtual agents supporting first responders. Despite the growing avenues for human-agent group interaction, however, a majority of the research on interaction between humans and artificial agents still focuses on one human interacting with one agent.

Research on group interactions between humans and artificial agents (both virtual agents and physical robots) is important, but often more challenging than studying dyadic interactions. It requires gathering groups of humans and artificial agents, and addressing additional factors that contribute to successful group interaction (such as intragroup dynamics). Further, while several research domains do tackle the challenges associated with group interaction, the focus on human-agent groups has been limited. For instance, research on multiagent systems within AI has primarily focused on teams of artificial agents, while research in social psychology and human factors engineering has primarily focused on human teams. The goal of this symposium on human-agent groups was to bring together scholars from a variety of perspectives to discuss the state of the art and novel challenge in groups of humans and AI.

The symposium involved scholars from many research fields, including autonomous agents and multiagent systems, knowledge representation, conversational agents, decision support, human-in-the-loop planning, robotics, human-robot interaction, social networks, social psychology, design, and science policy. It featured a set of six invited talks and nine presentations from authors of contributed papers, including empirical studies, novel algorithmic challenges, and potential solutions for human-agent groups.

Talks and discussion on empirical studies of human teams and human-agent groups provided foundational insights for modeling and representing groups from an artificial intelligence perspective. For instance, humans associate with multiple groups (related to work, family, nationality), and group membership drives their behavior and cooperation within the group. Although an artificial agent engaging in group interactions can benefit from the ability to represent and identify such flexible and evolving group memberships, the ability to represent flexible group memberships is currently missing from classical AI models of groups. Brian Lickel (University of Massachusetts, Amherst) pointed out that humans have been successfully interacting with groups of animals, which are nonhuman agents with different physical and cognitive capabilities — and that insights from these group interactions might be helpful for informing the research on human-agent groups. Yuichiro Yoshikawa (Osaka University) pre-

sented an empirical study demonstrating the utility of multiple robots interacting among each other to improve the sense of successful human-robot conversation.

The contributed papers included a variety of novel AI solutions for human-agent groups, including an approach to generate explanations of AI decisions for a group of humans, a path planner for navigation in human crowds, and decision support systems for human-in-the-loop planning. Ana Paiva (Universidade Tecnica de Lisboa) discussed her recent and ongoing research on creating human-robot groups and shared her insights on factors that impact humans' choices of robot partners. David Sirkin (Stanford University) discussed designing robot behavior for group interaction and the way that human groups often negotiated among themselves before engaging in interaction with robots. Shiqi Zhang (Cleveland State University) discussed approaches that leverage the strengths of both symbolic planning and decision making under uncertainty and how these approaches can be applied to human-agent groups. Christopher Amato (North-eastern University) discussed his research on scalable models and algorithms for multiagent systems, highlighting both their utility and the challenges that remain to develop algorithms for coordinating human-robot teams.

Along with invited talks and author presentations, the symposium also included a set of breakout sessions to encourage in-depth discussion among participants and seed collaboration with researchers from diverse perspectives. Jeff Nickerson (Stevens Institute of Technology) led one of the breakout sessions to discuss a new NSF convergence research collaboration network. The discussions in the concluding session of the symposium emphasized the importance of further research on human-agent groups and future workshops to facilitate collaboration on this pressing issue in human-technology interaction.

Marlena R. Fraune (Indiana University), Vaibhav V. Unhelkar, Bradley Hayes, Julie Shah (all from the Massachusetts Institute of Technology), Selma Šabanovic (Indiana University), Friederike Eyssel (Bielefeld University), and Malte Jung (Cornell University) served as cochairs of this symposium. The papers of the symposium were published as AAAI Press Technical Report FS-17-04.

## Natural Communication for Human-Robot Collaboration

The organizers of this symposium, Jean Oh (Carnegie Mellon University), Matthew Walter (Toyota Technological Institute at Chicago), and Zhou Yu (University of California, Davis), did not submit a report of their symposium nor were any papers submitted or published in the technical report.

## A Standard Model of the Mind

The AAAI symposium on a standard model of the mind brought together researchers from artificial intelligence, cognitive science, neuroscience, and artificial general intelligence to explore the general notion of a standard model of the mind, to assess a particular proposed start towards that model, to make progress beyond this start where possible, to understand the extent of support from the community for pursuing this notion further, and to plan for the future should the community agree. It accomplished all of these goals to varying extents, and perhaps even more importantly, it yielded a range of discussions over the three days of critical topics concerning the development of comprehensive/integrated/unified models of the mind that too often fail to fit within other venues.

The first day of the symposium began with an introduction to the concept of a standard model of the mind and the proposed start towards it, based on a draft of the standard model in the winter 2017 issue of *AI Magazine*. This discussion was followed by sessions on the mapping of existing cognitive architectures onto this approach, the general foundations and requirements for a standard model, and alternative perspectives to it. The first of these sessions explored lessons that could be derived for the standard model from a group of existing cognitive architectures, with particular note being taken of mismatches and missing pieces. Most of the presentations in the second session focused either on understanding the space of capabilities that will ultimately need to be supported in a standard model, a number of which are not yet reflected in it, or pointed out limitations that result from the under-specification that results from only including aspects of the mind in the standard model if there is a consensus concerning them. Presentations in the third session focused on perspectives on the standard model from below-the-deliberate-act level (that is, from the brain) or above it (that is, from knowledge and higher-level capabilities), sometimes agreeing with the standard model in the process about what should be at the deliberate act level, and sometimes disagreeing.

The second day began with two sessions focused further on the levels/scales below and above the standard model, exploring their relationship to the standard model, what inspiration or constraint might be derived from them for the standard model, and the possibility of building more comprehensive models by explicitly including them. The day then continued with sessions exploring in more detail (1) two aspects already reflected in the standard model but that may need further work — long-term memory and procedural control — and (2) three bundles of interrelated aspects that are currently missing from the standard model: metalevels, reflection, and consciousness; visual, spatial, and imagery; and motivation, affect, mood, and physiology. With respect to



## HCOMP-2019

October 28–30, 2019

Skamania Lodge, Stevenson, WA USA  
(Columbia River Gorge National Scenic Area)

Edith Law (University of Waterloo) (Cochair)  
Jenn Wortman Vaughn (Microsoft Research) (Cochair)

Inquiries: [hcomp19@aaai.org](mailto:hcomp19@aaai.org)

the first point, there were discussions concerning how unified long-term memory should be, and the concept of a cognitive cycle, its parallelism versus sequentiality, and its discrete versus continuous nature. With respect to the second, there was general agreement that forms of these aspects are required, but the discussions did not go quite far enough yet to yield a consensus concerning what specifically should be added to the standard model.

The third (half) day focused on two discussion sessions, with the first assessing the overall status of the effort to develop a standard model, given what transpired at the symposium, and the second discussing what the future of this effort should be. In the first session, issues came up concerning how the standard model needs additional refinement and expansion, while avoiding over-constraint when not warranted; how it should not be proscriptive but continue instead to encourage alternative approaches; and how the material at the symposium itself was not well tied to either human data or real-world domains (although it was based on a variety of cognitive archi-

tectures, each of which has been directly tied to one or both of these). In the second session, the main question was whether to proceed, which received near-unanimous approval from those in attendance. As to how to proceed, given the need for this to be a community effort, the co-organizers of the symposium are to re-form themselves as an initial steering committee, with a mailing list to be developed, and discussions taking place via this list before major decisions are made. Working groups dedicated to specific topics were proposed as a tractable mechanism for evolving consensus. Venues for publication of standard model advances were also discussed. In addition, a variety of possibilities for future events at which further progress could be made as a community were discussed.

As organizers, we were pleasantly surprised by the level of enthusiasm expressed for pursuing a community-driven consensus for comprehensive models of the mind. It appears the time is ripe for groups to work together across many disciplines to develop, refine, and test such a consensus.

Paul S. Rosenbloom, John E. Laird, and Christian Lebiere served as cochairs of this symposium. The papers of the symposium were published as AAAI Press Technical Report FS-17-01.

**Arjuna Flenner** is a senior research physicist at NAVAIR, China Lake.

**Marlena R. Fraune** is a PhD candidate at Indiana University.

**Laura M. Hiatt** is a researcher at the Navy Center for Applied Research in Artificial Intelligence, part of the Naval Research Laboratory (NRL) in Washington, DC.

**Tony Kendall** is a research professor at the Naval Postgraduate School.

**John E. Laird** is a professor in the Computer Science and Engineering Division of the Electrical Engineering and Computer Science Department at the University of Michigan.

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

**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.

**Frank Stein** is the director of the Analytics Solution Center at IBM.

**Elin A. Topp** is a senior lecturer in computer science at Lund University, Lund, Sweden, and affiliated with the Lund University RobotLab shared between the Departments of Computer Science and Automatic Control.

**Vaibhav V. Unhelkar** is a PhD candidate at the Massachusetts Institute of Technology.

**Ying Zhao** is a research professor at the Naval Postgraduate School.