# Reports on the 2016 AAAI Fall Symposium Series

*Patrícia Alves-Oliveira, Richard G. Freedman, Dan Grollman, Laura Herlant, Laura Humphrey, Fei Liu, Ross Mead, Frank Stein, Tom Williams, Shomir Wilson*

■ *The AAAI 2016 Fall Symposium Series was held Thursday through Saturday, November 17–19, at the Westin Arlington Gateway in Arlington, Virginia, adjacent to Washington, DC. The titles of the six symposia were Accelerating Science: A Grand Challenge for AI; Artificial Intelligence for Human-Robot Interaction; Cognitive Assistance in Government and Public Sector Applications; Cross-Disciplinary Challenges for Autonomous Systems; Privacy and Language Technologies; and Shared Autonomy in Research and Practice. The highlights of each symposium (except the Accelerating Science symposium) are presented in this report.*

## Artificial Intelligence for Human-Robot Interaction

In 2016, AAAI hosted the third Artificial Intelligence for Human-Robot Interaction symposium. While previous incarnations of the symposium focused on finding a place for AI and human-robot interaction (HRI) work within the greater HRI community and improving the interaction between these groups, this year's symposium took a more introspective view. With the introduction of the paired talk and debate format, the symposium challenged attendees to examine the basic assumptions that underlie their work and to consider anew the various perspectives of design, cognitive models, and human-in-the-loop computation. Furthermore, organizers dedicated the final half-day to the discussion of tools, an oft overlooked, but vitally important, aspect of research.

The goal was that the symposium serve as the gathering point for the AI-focused community within HRI and allow for the sharing of the most exciting research in this area while cultivating a vibrant, interconnected research commu-

 Copyright © 2017, Association for the Advancement of Artificial Intelligence. All rights reserved. ISSN 0738-4602

nity. While in previous years the Artificial Intelligence for Human-Robot Interaction symposium has only explicitly solicited novel computational work, this year's symposium also solicited contributions from social scientists, philosophers, and industry professionals to promote discussion and cohesion within the multidisciplinary community interested in the intersection of AI and human-robot interaction.

The majority of papers presented at the symposium addressed novel computational mechanisms that enable robots to take better physical actions in the world, better perceive the structure of the world or interactions within the world, or better engage in linguistic interactions with human users. Moreover, a shared set of motivations emerged, as the majority of the contributions sought to either make interactions safer, make interactions smoother, or allow robots to learn better from humans. In addition, a number of papers were presented at the intersection of these areas. These papers presented novel computational mechanisms that enable robots to better perceive human actions to take better physical actions, better understand natural language to take better physical actions, or better generate natural language to communicate the robot's perceptions.

Finally, a number of submissions explored broader concerns in HRI, such as the concepts of trust, explainability, and reflection, or metaconcerns such as relations with other fields and practices within the field. These topics proved to spark the greatest amount of discussion, and all authors that presented full talks had the opportunity to comment on the relation between their own research and these larger concerns over the course of the author panels that followed each session of paper presentations. These author panels provided a space for authors and audience to engage in open conversation regarding the details and motivations of individual papers, the connections within each set of papers, and the relation between those papers and the evolving field of human-robot interaction.

There was further opportunity to discuss these issues through the invited speakers, who came from various aspects of the AI-HRI community and exposed the entire group to current issues and relevant works in their area. Many of these presentations were given in a debate format where two invited speakers presented a variety of views on their issues of choice, including playing devil's advocate to their own opinions and preferences. Topics ranged from questioning the relevance between AI and HRI to identifying the best practices and future for AI-HRI research with respect to design, cognition, human involvement and roles, and improving the safety and comfort of interactive experiences. After each presentation, breakout sessions and group discussions challenged the participants to think deeper about the issues and come up with their own insights to share afterwards. The speakers even joined the groups to engage and help spark the conversations that often continued during the coffee breaks.

This year's Artificial Intelligence for Human-Robot Interaction symposium builds on the successes of previous years, and further paves the path for research innovation and community growth for years to come.

Ross Mead and Dan Grollman served as general chairs, Tom Williams and Patrícia Alves-Oliveira served as program committee chairs, and Rick Freedman as panel chair of the symposium. The papers of the symposium were published by AAAI Press as Technical Report FS-16-01 within the compilation titled *The 2016 AAAI Fall Symposium Series: Technical Reports FS-16-01 – FS-16-05*.

## Cognitive Assistance in Government and Public Sector Applications

The concept of a cognitive assistant as a partner to help humans perform their work better dates to the early days of AI, including the 1960s writings of Douglas Engelbart and J. C. R. Licklider. For the sake of this symposium, cognitive assistance was defined as "a systematic approach to improving performance on complex tasks that require the processing of large amounts of information in which people and machines are treated as complementary cosystems working together." Building on the successful 2015 AAAI Fall Symposium on Cognitive Assistance, symposium organizers worked to expand the dialog between the user, academic, and commercial communities to increase understanding of the best practices and barriers to the adoption of such systems. Last year's presentations largely dealt with ideas. This year, attendees heard about systems in actual use and the adoption of such systems. Also, notable this year was increased focus on specific professions and domains, including law and the courts, patent processing, teaching, health, cybersecurity, aviation, air traffic control, defense, and intelligence.

The symposium also included three invited talks. Mark Maybury (vice president, MITRE) opened with historical context and then talked about goals (increased prosperity, safety, happiness) and challenges (transparency, trustworthiness, privacy and security, and control versus the triad of confidence, threat, and potential consequences). Ed Felten (deputy US chief technical officer) talked about the need for the AI community to inform the process in the areas of risks, employment, fairness, and justice. He encourated the community, as it moves forward with cognitive assistance, to consider such questions as: Does it work? Is it fair? and Is it acceptable to the public? Guru Banavar (vice president, IBM Research) discussed the meaning of expertise in the 21st century and how cognitive systems can help experts make better decisions. He suggested the need to integrate

many different training techniques to give a better result and to completely rethink software engineering for the cognitive age — "TrainOps rather than DevOps."

The panel discussion on the legal and judiciary domain discussed how cognitive assistants can help reduce the caseload on the courts and administrative judges. Brad Brown (MITRE) facilitated the discussion and reminded the panel that where adjudication is being done through automated means, it would be subject to dispute. They discussed the need to validate the system in each context and focus on assistance rather than automation of arbitration.

The second panel, facilitated by Jim Spohrer (IBM), was on creating a teaching assistant. Barriers that were noted included trying to build systems to handle all the students' questions reliably, authoring tools, getting the right data to build the systems, and matching the learners' aptitudes and styles.

The panel on cybersecurity assistants was chaired by Dan Tobin (NIST). Lessons learned included starting small, insuring interannotator agreement during the human annotation phase, and iterating to improve accuracy and recall. Panelists noted that the tools used to improve the process are getting better.

The DoD and Intel panel led by Scott Kordella (MITRE) raised the issues of user trust of the systems but also the possibility the machine-driven recommendations may induce "intellectual laziness" in decision makers. Also raised were issues around metrics for complex human-machine systems; user acceptance of recommendations; and efficient data management, architectures, and computational processing for large amounts of dynamic data. A provocative comment was made by the panel that while there are many technical challenges to be addressed by the research community, a larger issue is that of user acceptance and policy constraints for machine automation. Such constraints include the degree to which systems are operated "closed-loop" without human intervention, whether this is desirable or appropriate.

The last panel, focused on cognitive assistance in aviation, was led by Chris Codella (IBM). Several of the projects were designed to help pilots as a digital copilot. The Auto Ground Collision Avoidance System by AFRL is in active use and has saved lives since being fielded.

The symposium was useful in showcasing many different examples of cognitive assistant projects and approaches to implementing them. The participants agreed that they would like to attend future symposia to share experiences as the field develops.

Frank Stein (IBM) and Chuck Howell (MITRE) served as cochairs of this symposium. The papers of the symposium were published by AAAI Press as Technical Report FS-16-02 within the compilation titled *The 2016 AAAI Fall Symposium Series: Technical Reports FS-16-01 – FS-16-05*.

## Cross-Disciplinary Challenges for Autonomous Systems

Future technology concepts envisage highly autonomous systems that are fast, efficient, and adaptive, yet also human friendly and provably safe. However, no single field is capable of addressing all the challenges that face practical development and fielding of such systems. Organizers of this symposium therefore sought to bring together experts whose research spans two or more traditionally distinct disciplines in order to learn about recent successes in cross-disciplinary research for autonomous systems, discuss ways to encourage collaboration and education across disciplines, and spur ideas for future cross-disciplinary research.

The symposium consisted of nine paper presentations, three keynote speakers, a panel discussion, small group discussions, and large group discussions. Throughout the symposium, attendees were asked to self-identify their disciplines, which included computer science; cognitive science; cognitive psychology; aeronautics and astronautics; and mechanical, electrical, computer, chemical, and aerospace engineering. Attendees were also asked to self-identify their research areas, which included autonomics, machine learning, cognitive architectures, formal methods for design and verification, intelligent decision aids, agent frameworks, human-autonomy interaction, multiagent systems, and robotics, among others. Papers presented at the symposium shared some common themes, particularly the desire to better incorporate results from sociology, psychology, and cognitive modeling in design of autonomous systems and the desire to develop approaches to verify properties of autonomous or human-automation systems.

More specifically, paper topics included methods for incorporating models of human trust to improve human-automation system performance; verifying the behavior of human-automation systems; designing autonomous systems to guarantee certain safety or performance specifications or to classify and handle errors at run time; designing autonomics frameworks that incorporate machine learning for self-adaptation, formal methods for verification, and improved interfaces and methods of interaction for human users; and building flexible and extensible frameworks to study situated autonomy. Keynote speakers provided overviews of progress in cross-disciplinary research, particularly in robotics and in the interplay between planning and formal methods, and motivated the need for future cross-disciplinary research to produce capabilities such as explainable artificial intelligence. The panel discussion on ethics in encouraging trust in autonomous systems highlighted some of the potential pitfalls of designing autonomous systems to encourage trust, which was particularly germane given the emphasis on trust in the paper presentations.

During small group discussions, attendees divided into groups and considered questions related to what disciplines are required for the development of autonomous systems, where cross-disciplinary research has been successful and where it has stalled, and what would help facilitate more collaboration across disciplines. During large group discussions, the attendees collectively considered selected statements from recent government guidance for autonomous systems, including DoD Directive 3000.09: Autonomy in Weapon Systems, AF Autonomous Horizons Volume 1: Human-Autonomy Teaming, and the DoD Defense Science Board Task Force Report: The Role of Autonomy in DoD Systems. As with the paper presentations, some common themes were apparent. For instance, organizations by nature tend to group disciplines together, which hinders collaboration. Furthermore, strategic investments by funding agencies tend emphasize the technology *du joir,* leading to a focus on disciplines most closely associated with that technology and a failure to recognize the continued relevance of others. Attendees provided a number of recommendations to address these and related issues, including workshops, journals, academic courses, funding, and challenge problems specifically aimed at cross-disciplinary research. Attendees suggested drawing inspiration from successful cross-disciplinary programs, including the DARPA Grand Challenge and NASA's Mars Exploration Rover Mission, as well as learning lessons from less conventional sources, such as management strategies for large technical organizations.

This report was written by Laura Humphrey, who served with Chris Miller (Smart Information Flow Technologies), Satinder Singh (University of Michigan), Ufuk Topcu (University of Texas at Austin), and Moshe Vardi (Rice University) as cochairs of this symposium. The papers of the symposium were published by AAAI Press as Technical Report FS-16-03 within the compilation titled *The 2016 AAAI Fall Symposium Series: Technical Reports FS-16-01 – FS-16-05*.

## Privacy and Language Technologies

Computing applications continue to increase their reliance on natural language, with online social media, search engines, and spoken dialog systems enabling the collection of large amounts of language data for advertising, text mining, surveillance, and other purposes. Simultaneously, an opportunity exists for human language technologies to help people understand and protect their privacy by simplifying information they must comprehend (for example, privacy policies or other legal documents) or by helping them to reflect on the content they choose to share (for example, social media posts). The symposium invited paper submissions and participants who were interested in research in either of these two directions, with the goal of promoting interaction and synergy between them.

During the symposium, two themes emerged from the presentations and discussions: (1) the use of language technologies to inform users about their online privacy status and choices, and (2) the use of language technologies to protect users from the abuse of their personal information. Toward the first theme, online privacy policies were a major topic of discussion. Privacy policies are purported to inform users how their personal information is used and what choices they have, but these documents are infrequently read, leading to an opening for natural language processing efforts to extract salient details from them. Toward the second theme, participants discussed privacy preservation in online social networks, web search engines, and dialogue systems.

Talks by the two keynote speakers followed these two major themes. Norman Sadeh (Carnegie Mellon University) presented Using Artificial Intelligence and Machine Learning to Empower Users to Regain Control over Their Privacy, and Doug Oard (University of Maryland) spoke on The Other Side of the Coin: Proactive Language Technology for Privacy Protection. The keynote presentations spurred discussion regarding the effectiveness of the privacy notice and choice framework, the Internet of Things and privacy preferences, and the searchability and sharing of sensitive data, among other topics. It became clear that a variety of stakeholders, such as companies in private sector, legal scholars, media, privacy researchers, language researchers, and government regulators, could be involved to address the pressing issues of privacy protection. In addition to presentations and discussion forums, the symposium included a poster session and a joint session with the Cognitive Assistance in Government and Public Sector Applications symposium, to explore possible collaborations.

Participants agreed that the event produced particularly good discussions, in large part because of the diversity of disciplines represented by the attendees, such as social scientists, a lawyer, and a former police officer. They also expressed an interest in attending a future event with a similar focus, potentially as a workshop at a related conference.

The symposium was chaired by Shomir Wilson (lead organizer), Fei Liu, and Alessandro Oltramari. The papers of the symposium were published by AAAI Press as Technical Report FS-16-04 within the compilation titled *The 2016 AAAI Fall Symposium Series: Technical Reports FS-16-01 – FS-16-05*.

## Shared Autonomy in Research and Practice

Shared autonomy, where a robot and an operator collaborate to solve a complex physical task, provides an

exciting and encompassing theme for research in motion planning, control, machine learning, perception, and human-robot interaction. Furthermore, its applications extend human capabilities, to deep-sea exploration, disaster recovery, rehabilitation, and assistive care, enabling us to explore new possibilities safely and efficiently. This symposium provided a deep dive into the dual themes of fundamental theoretical algorithms for shared autonomy, as well as their real-world practical application through a unique partnership with the FDA. The goal of this symposium was to build bridges between research domains and also between researchers and practitioners, and to set the foundations for the future of shared autonomy.

Through keynote and contributed talks, a variety of different goals for shared autonomy were revealed. One use of shared autonomy was as a tool to teach human operators how to better perform a control task. In rehabilitation robotics, this could be seen through object avoidance assistance in powered wheelchairs that was reduced as the operators became more familiar with the wheelchair's capabilities. In another application, operators were taught to control a robot performing the classic inverted pendulum problem by rejecting operator inputs that were not consistent with the automated system's optimal policy. Another use of shared autonomy was to provide intelligent safety constraints in applications such as aircraft autopilots and surgical robotics. In these situations, there may not be time for the AI to communicate why a particular override is occurring, and it opened the floor for a discussion on creating "explainable" AI systems, and how transparent the AI should be with the operator about its intentions.

Another major theme was how to evaluate shared autonomy systems. Because shared autonomy often uses learned customization to each operator, who in turn has changing abilities over time, it is difficult to assess effectiveness of the overall system in a standardized way. The most common assessment tool being used is to lower bound performance by either the human or system operating individually, and use this baseline to compare improvements of the joint system. Ensuring safety with a constantly evolving and learning system was brought up during the panel discussion on commercialization of shared autonomy technologies. Evaluation is necessary for regulatory approval, as well as insurance coverage in the case of medical devices. Safety evaluation is an open challenge among all new AI technologies, from autonomous driving to assistive wheelchairs and surgical robots. Developing metrics to evaluate how closely operator intentions are being followed and bounding the probabilities on undesirable outcomes are important steps toward enabling shared autonomy systems to be broadly utilized.

The symposium participants discussed challenges inherent in all shared autonomy systems where a human and AI are interacting. These challenges included managing operator expectations, maintaining operator engagement at increasing levels of autonomy, identifying the right moments to have assertive AI, providing the AI with enough input to determine relevant context, and the impact of autonomy on the operator's trust of the system. This symposium brought together a wide range of participants from industry, regulatory bodies, and academe and sparked discussions and collaborations that may never have formed otherwise due to the diverse nature of shared autonomy applications.

Laura Herlant served as chair of this symposium. The papers of the symposium were published by AAAI Press as Technical Report FS-16-05 within the compilation titled *The 2016 AAAI Fall Symposium Series: Technical Reports FS-16-01 – FS-16-05*.

**Patrícia Alves-Oliveira** is a PhD student in psychology applied to human-robot interaction at Instituto Universitário de Lisboa and INESC-ID.

**Richard G. Freedman** is a PhD candidate in the College of Information and Computer Sciences at the University of Massachusetts Amherst.

**Dan Grollman** is the Robot Brain Architect of Sphero, Inc.

**Laura Herlant** is a PhD candidate at the Robotics Institute of Carnegie Mellon University.

**Laura Humphrey** is a research engineer at the Aerospace Systems Directorate of the Air Force Research Laboratory.

**Fei Liu** is an assistant professor in the Department of Computer Science at the University of Central Florida.

**Ross Mead** is the founder and CEO of Semio.

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

**Tom Williams** is a PhD candidate in the Human-Robot Interaction Lab at Tufts University.

**Shomir Wilson** is an assistant professor in the Electrical Engineering and Computer Science Department of the University of Cincinnati.