Editorial Introduction to the Special Articles in the Summer Issue

Innovative Applications of Artificial Intelligence 2015

David Gunning, Peter Z. Yeh

■ *This issue features expanded versions of articles selected from the 2015 AAAI Conference on Innovative Applications of Artificial Intelligence held in Austin, Texas. We present a selection of three articles describing deployed applications plus two more articles that discuss work on emerging applications.*

The AAAI Conference on Innovative Applications of Artificial Intelligence (IAAI) was founded in 1989 to showcase the successful application of artificial intelligence technology to real-world problems and its deployment into the hands of end users. Since then, we have seen examples of AI applied to domains as varied as medicine, education, manufacturing, transportation, user modeling, military operations, and citizen science. The 2015 conference continued the tradition with a selection of 6 deployed applications describing systems in use by their intended end users, 13 emerging applications describing works in progress, and three papers in a new category for challenge problems.

In the first article, Activity Planning for a Lunar Orbital Mission, John Bresina describes a deployed application of current planning technology in the context of a NASA mission called LADEE (Lunar Atmospheric and Dust Environment Explorer). Bresina presents an approach taken to reduce the complexity of the activity-planning task in order to perform it effectively under the time pressures imposed by the mission requirements. One key aspect of this approach is the design of the activity-planning process based on principles of problem decomposition and planning abstraction levels. The second key aspect is the mixed-initiative system developed for this task, the LADEE activity scheduling system (LASS). The primary challenge for LASS was representing and managing the science constraints that were tied to key points in the spacecraft's orbit, given their dynamic nature due to the continually updated orbit determination solution.

In our second article, Helping Novices Avoid the Hazards of Data: Leveraging Ontologies to Improve Model Generalization Automatically with Online Data Source, Sasin Janpuangtong and Dylan Shell describe an emerging application of an end-to-end learning framework for large-scale data analytics that allows a novice to create models from data easily by helping structure the model-building process. By treating the whole modeling process interactively and exploiting high-level

knowledge in the form of an ontology, the framework is able to aid the user in a number of ways, including in helping to avoid pitfalls such as data dredging. Their approach adopts the solution of using higher-level knowledge to assist the novice user in selecting relevant input attributes, and thereby constraining the hypothesis space. To validate the approach, models of four different problem domains were built using their implementation of the framework. Prediction error on unseen examples of these models show that their framework, making use of the ontology, helps to improve model generalization.

In our third article, Graph Analysis for Detecting Fraud, Waste, and Abuse in Health-Care Data, Juan Liu, Eric Bier, Aaron Wilson, John Alexis Guerra-Gomez, Tomonori Honda, Kumar Sricharan, Leilani Gilpin, and Daniel Davies describe a deployed application using graph-based analytics to detect suspicious activities in large health-care data sets. Each health-care data set is viewed as a heterogeneous network consisting of millions of patients, hundreds of thousands of doctors, tens of thousands of pharmacies, and other entities. Graph-analysis techniques are developed to find suspicious individuals, suspicious relationships between individuals, unusual changes over time, unusual geospatial dispersion, and anomalous network structure. A visualization interface, known as the Network Explorer, provides a good overview of data and enables users to filter, select, and zoom into network details on demand. The system has been deployed on multiple sites and data sets, both government and commercial, and identified many overpayments with a potential value of several million dollars per month.

In our fourth article, PSINET: Assisting HIV Prevention Among Homeless Youth by Planning Ahead, Amulya Yadav, Leandro Soriano Marcolino, Eric Rice, Robin Petering, Hailey Winetrobe, Harmony Rhoades, Milind Tambe, and Heather Carmichael describe an emerging application of partially observable Markov decision process (POMDP) techniques to model influence propagation to maximize education and training in human immunodeficiency virus (HIV) vulnerable communities. Homeless youth are prone to HIV due to their engagement in high-risk behavior such as unprotected sex, sex under influence of drugs, and other factors. Many nonprofit agencies conduct interventions to educate and train a select group of homeless youth about HIV prevention and treatment practices and rely on word-of-mouth spread of information through their social network. Previous work in strategic selection of intervention participants does not handle uncertainties in the social network's structure and evolving network state, potentially causing significant shortcomings in spread of information. They describe PSINET, a decision support system to aid the agencies in this task. PSINET includes the following key novelties: (1) it handles uncertainties in network structure and evolving network state; (2) it addresses these uncertainties by using POMDPs in influence maximization; and (3) it provides algorithmic advances to allow high quality approximate solutions for such POMDPs. Simulations show that PSINET achieves approximately 60 percent more information spread over the current state of the art. PSINET was developed in collaboration with My Friend's Place (a drop-in agency serving homeless youth in Los Angeles) and is currently being reviewed its their officials.

In our final article, Capturing Planned Protests from Open Source Indicators, Sathappan Muthiah, Bert Huang, Jaime Arredondo, David Mares, Lise Getoor, Graham Katz, and Naren Ramakrishnan describe a deployed application of phrase learning, probabilistic soft logic (PSL), and time normalization to detect planned protests from news and social media in several Latin American countries. In countries where civil unrest is lawful, qualitative analysis has revealed that more than 75 percent of the protests are planned, organized, or announced in advance; therefore detecting references to future planned events in relevant news and social media is a direct way to develop a protest forecasting system. The authors report on a system for doing that. It uses a combination of key-phrase learning to identify what to look for, probabilistic soft logic to reason about location occurrences in extracted results, and time normalization to resolve future time mentions. They illustrate the application of their system to 10 countries in Latin America: Argentina, Brazil, Chile, Colombia, Ecuador, El Salvador, Mexico, Paraguay, Uruguay, and Venezuela. Results demonstrate the successes of their system in capturing significant societal unrest in these countries with an average lead time of 4.08 days.

In the tradition of previous special issues on innovative applications of artificial intelligence, and consistent with the goals of the IAAI conference, the articles in this issue describe work that is strongly grounded in the needs of end users. We hope that you enjoy the articles, and that they both provide insight into the application development process and help to expand your view of what is possible with AI technology. We also invite you to submit a description of your next AI application to IAAI.

David Gunning directs PARC's program in data analytics, which includes projects in anomaly detection, behavioral analytics, privacy-preserving analytics, fraud detection, and analysis tools. He also manages PARC's research area in contextual intelligence, focused on developing context-aware computing technology to enable rich, predictive user models that capture and understand the user's situational context. Prior to PARC, Gunning was a senior research manager at Vulcan Inc., a program manager at DARPA (twice), senior vice president of SET Corporation, vice president of Cycorp, and a senior scientist in the Air Force Research Labs. Gunning holds an MS in computer science from Stanford University and an MS in cognitive psychology from the University of Dayton.

Peter Z. Yeh is a senior principal research scientist at Nuance Communications. His research interests lie at the intersection of semantic technologies, data and web mining, and natural language understanding. At Nuance, Yeh's research focuses on applying AI technologies to enhance interpretation intelligence within intelligent virtual assistants and to construct automatically large-scale knowledge repositories necessary to support such interpretations. Prior to joining Nuance, Yeh was a research lead at Accenture Technology Labs where he was responsible for investigating and applying AI technologies to various enterprise problems ranging from data management to advanced analytics. Yeh received his Ph.D. in computer science from the University of Texas at Austin.