5-2009

AAAI 2008 Workshop Reports

Mark Dredze
University of Pennsylvania, mdredze@seas.upenn.edu

AAAI 2008 Workshop Reports

Abstract
AAAI was pleased to present the AAAI-08 Workshop Program, held Sunday and Monday, July 13-14, in Chicago, Illinois, USA. The program included the following 15 workshops: Advancements in POMDP Solvers; AI Education Workshop Colloquium; Coordination, Organizations, Institutions, and Norms in Agent Systems, Enhanced Messaging; Human Implications of Human-Robot Interaction; Intelligent Techniques for Web Personalization and Recommender Systems; Metareasoning: Thinking about Thinking; Multidisciplinary Workshop on Advances in Preference Handling; Search in Artificial Intelligence and Robotics; Spatial and Temporal Reasoning; Trading Agent Design and Analysis; Transfer Learning for Complex Tasks; What Went Wrong and Why: Lessons from AI Research and Applications; and Wikipedia and Artificial Intelligence: An Evolving Synergy.

Comments
Sarabjot Singh Anand, Razvan Bunescu, Vitor Carvalho, Jan Chomicki, Vincent Conitzer, Michael T Cox, Virginia Dignum, Zachary Dodds, Mark Dredze, David Furcy, Evgeniy Gabrilovich, Mehmet H Göker, Hans Guesgen, Haym Hirsh, Dietmar Jannach, Ulrich Junker. (2009, April). AAAI 2008 Workshop Reports. AI Magazine, 30(1), 108-118.

Publisher URL: http://proquest.umi.com/pqdlink?did=1680350011&sid=1&Fmt=2&clientId=3748&RQT=309&VName=PQD

AAAI 2008 Workshop Reports

Sarabjot Singh Anand, Razvan Bunescu, Vitor Carvalho, Jan Chomicki, Vincent Conitzer, Michael T. Cox, Virginia Dignum, Zachary Dodds, Mark Dredze, David Furcy, Evgeniy Gabrilovich, Mehmet H. Göker, Hans Guesgen, Haym Hirsh, Dietmar Jannach, Ulrich Junker, Wolfgang Ketter, Alfred Kobsa, Sven Koenig, Tessa Lau, Lundy Lewis, Eric Matson, Ted Metzler, Rada Mihalcea, Bamshad Mobasher, Joelle Pineau, Pascal Poupart, Anita Raja, Wheeler Ruml, Norman Sadeh, Guy Shani, Daniel Shapiro, Trey Smith, Matthew E. Taylor, Kiri Wagstaff, William Walsh, and Rong Zhou

■ *AAAI was pleased to present the AAAI-08 Workshop Program, held Sunday and Monday, July 13–14, in Chicago, Illinois, USA. The program included the following 15 workshops: Advancements in POMDP Solvers; AI Education Workshop Colloquium; Coordination, Organizations, Institutions, and Norms in Agent Systems, Enhanced Messaging; Human Implications of Human-Robot Interaction; Intelligent Techniques for Web Personalization and Recommender Systems; Metareasoning: Thinking about Thinking; Multidisciplinary Workshop on Advances in Preference Handling; Search in Artificial Intelligence and Robotics; Spatial and Temporal Reasoning; Trading Agent Design and Analysis; Transfer Learning for Complex Tasks; What Went Wrong and Why: Lessons from AI Research and Applications; and Wikipedia and Artificial Intelligence: An Evolving Synergy.*

The Workshop on Advancements in POMDP Solvers brought together active researchers in the area of solving partially observable Markov decision processes (POMDPs). Participants discussed various approaches to solving POMDPs, and discussed as well as potential real real-world applications of the model. The AI Education Colloquium kicked off AAAI 2008's AI Forum, a series of events on the teaching and learning of AI. The colloquium convened AI practitioners passionate about improving both their students' and their own appreciation of our field's compelling ideas. The goal of the workshop was to examine and define the current state of the art research in agent systems research related to coordination, organizations, institutions, and norming. The Enhanced Messaging workshop brought together researchers from across the AI and computer science spectrum to discuss the state of research on e-mail and information overload. New connections between participants are driving forward work in this area and building a new research community. The Human Implications of Human-Robot Interaction (HRI) workshop concerned aspects of HRI that particularly call for multidisciplinary research and dialogue, representing AI and robotics as well as disciplines such as psychology, theology, sociology, and philosophy. The Intelligent Techniques for Web Personalization and Recommender Systems workshop was scheduled as a joint event, bringing together researchers and practitioners from the fields of web personalization and recommender sys-

 ISSN 0738-4602

tems. It focused on current and emerging topics related to web intelligence, particularly its application to recommender systems. The goal of the Metareasoning workshop was to explore the implications of a proposed model for metareasoning by examining its aspects, its use as a model of self, and its role in single-agent and multiagent applications. The Advances in Preference Handling workshop highlighted recent progress in eliciting and exploiting preferences for computational tasks from artificial intelligence, databases, and operations research. The Search in Artificial Intelligence and Robotics workshop brought together search researchers to share their ideas and disseminate their latest research results. It focused on finding common ground between search techniques used in artificial intelligence and robotics with great success. The Workshop on Spatial and Temporal Reasoning brought together related communities of researchers with an interest in the study of representing and reasoning about either space or time—or both. The Trading Agent Design and Analysis workshop focused on the design and evaluation of trading agents. The Transfer Learning for Complex Tasks workshop covered a wide range of topics, including regression, classification, reinforcement learning, planning, Markov logic networks, and neural networks. The What Went Wrong and Why workshop at AAAI-08 was dedicated to the propositions that insight often begins with unexpected results, and that clarity arrives in the ensuing response. The goals of the Wikipedia and Artificial Intelligence workshop were to investigate the mutually beneficial interaction between Wikipedia and AI and to foster a discussion on new applications and research directions that could benefit from this increasingly important relationship.

Advancements in POMDP Solvers

Over the past decade, much advancement was achieved in the field of POMDP solvers. The size of POMDPs that solvers can handle has increased by orders of magnitude. Solvers developed 10 years ago were hardly able to handle more than 10 states, while modern solvers can handle domains with millions of states. New techniques focus on computing approximate policies of manageable complexity, thus allowing us to handle these larger and more complicated POMDPs. This advancement was achieved by a few orthogonal approaches—the use of point-based techniques, finite-state controllers, efficient model representations, model compression techniques, hierarchical decompositions, inference-based techniques, and improved algorithms for online search.

The first part of the workshop provided an overview of a number of these approaches. These tutorials attracted many researchers from nearby areas, such as planning, who were interested in learning about the new developments in the field. We began by a tutorial on point-based value iteration methods. These methods, which contributed much to the scaling up of POMDP solvers, compute a solution over a subset of the belief space using the point-based backup operator. Next, we presented a tutorial on solving POMDPs through an online search over the belief space, starting at every time step from the current belief state. The key challenges for online methods include using efficient techniques for pruning the search space. We then discussed policy iteration using finite state controllers (FSCs), another important method that has shown the ability to scale up significantly. This tutorial provided an overview of the practical concerns of using FSCs, focusing on bounded memory controllers as a possible method for limiting the exponential growth of the FSC. The final tutorial reviewed methods for exploiting domain structure, focusing on factored POMDPs and algebraic decision diagrams (ADDs) to efficiently represent and compute policies.

The second part of the workshop included presentations on several examples of real-world applications of POMDPs, including assistance to elderly people using robots and cognitive reminders and a spoken dialog system aimed at fixing Internet connectivity problems. The demonstrations focused on unexpected concerns that are not traditionally handled when computing a POMDP policy. Participants agreed that it is important to better understand such concerns in order to apply POMDPs to real-world domains.

In the third part of the workshop, researchers presented new technical contributions in POMDP solvers. While point-based methods and finite state controllers still offer many opportunities for scaling up and continue to present many interesting open questions, researchers also presented work in other directions. We heard interesting ideas pertaining to multiagent scenarios, POMDPs with continuous parameters, and the integration of expert knowledge into solutions. The discussions throughout the meeting indicated that POMDP researchers are interested in strengthening the community and its impact on the development of autonomous systems. We will hence investigate several methods for supporting research in this area, such as offering a community web page, maintaining lists of active researchers, a bibliography of relevant papers, tutorials and presentations, and links to relevant software. Finally, as many researchers showed interest in additional meetings, we decided to hold a second workshop next year. This workshop will focus on bringing together researchers that develop new algorithmic contributions for POMDPs and researchers that apply POMDPs for real-world problems.

Joelle Pineau, Pascal Poupart, Guy Shani, and Trey Smith organized the workshop. The papers are available as AAAI Press Technical Report WS-08-02.

AI Education Colloquium

The AI Education Colloquium arose from the premise that teaching and learning AI is where we, from any discipline within our broad community, have the most to offer one another. The 11 talks and nine poster presentations spanned the full range of education's scale space: pedagogical strategies, curricular innovations, and outreach to new audiences that leverage AI in innovative ways. Taken together, this diverse set of contributions suggests that AI education and AI research are two phrases describing a single pursuit.

In any room of educators, teaching tips will out. Suggestions included online resources that enable students to contrast old and new AI systems, for example, Eliza and current chatbots. For material less easily "demoed," emulating peer-reviewing techniques can deepen student engagement with primary-source papers. "AI in the real world" discussions are energizing—both in showing emerging applications and when debunking overreported achievements. The group shared several nifty AI assignments that motivated students through a balance of peer cooperation and competition. Eric Eaton of University of Maryland Baltimore County presented one memorable example of a next-generation Wumpusworld, in which an image of Freeman Hrabowski (customizable, of course) executed students' plans and responded to changing conditions in order to save inhabitants of a search-and-rescue simulation. A mature set of online applets, AIspace, offers students accessible interaction with many basic AI algorithms. Two overarching pedagogical goals emerged amid the many ideas: first, that students not only learn about AI but engage actively in it; second, that students perceive the connection between their AI work and current research.

Strengthening ties with current research and applications of AI also motivated much of the colloquium's curricular innovation. Successful curricula based on web crawling highlighted AI's central role in modern life. A wide variety of game-based assignments reinforced AI's ubiquity and added playful and creative opportunities for students to express themselves while designing, implementing, and experimenting with agent behaviors. Reflecting on such approaches, the group felt it would serve the AI community to articulate a set of fundamental skills—as opposed to topics—that students would build through an AI survey course. Such skill-based scaffolding would offer students a concrete and transferable grounding in "doing AI" while allowing the leeway in approach, breadth, or depth balance exemplified by these innovative curricula.

That flexibility—and AI's applicability—was pushed even further as the group asked, "How can AI education serve more than just the AI community?" In answer, participants presented non-AI courses in which AI served as a central theme. Several used robots to motivate students to think computationally in early computer science or at the K–12 level. Another asked students to write a MMORPG as a capstone CS 1 project. Perhaps the broadest-reaching example came from Jim Marshall of Sarah Lawrence College: a year-long seminar in which AI served as a touchstone for critical thinking, reading, and writing about the ideas in Ray Kurzweil's *The Singularity Is Near*. The audience was entirely first-year liberal-arts students. To disseminate such resources—both for AI education and for education with AI—participants expressed interest in better leveraging the AAAI's AI Topics library; more generally, the group welcomed opportunities to consider AI education at future AAAI venues. The colloquium's discussion concluded in agreement that, while energy invested in AI education does serve future generations of AI practitioners, it perhaps serves the current generation even more.

Zachary Dodds, Kiri Wagstaff, and Haym Hirsh served as cochairs of this symposium. The papers are available as AAAI Press Technical Report WS-08-02. The event's participant list, many posters and slides, discussion summaries, and supplementary material have been archived at www.cs.hmc.edu/aieducation.

Coordination, Organizations, Institutions, and Norms

Multiagent systems (MASs) are often understood as complex entities where a multitude of agents interact, usually with some intended individual or collective purpose. Such a view usually assumes some form of structure or set of norms or conventions that articulate or restrain interactions in order to make them more effective in attaining those goals, more certain for participants, or more predictable. The engineering of effective coordination or regulatory mechanisms is a key problem for the design of open complex multiagent systems.

In recent years, social and organizational aspects of agency have become a major issue in MAS research. Recent applications of MAS on web services, grid computing, and ubiquitous computing enforce the need for using these aspects in order to ensure social order within these environments. Openness, heterogeneity, and scalability of MAS pose new demands on traditional MAS interaction models. Therefore, the view of coordination and control has to be expanded to consider not only an agent-centered perspective but societal and organization-centered views as well.

The overall problem of analyzing the social, legal, economic, and technological dimensions of agent organizations, and the coevolution of agent interactions, provide theoretically demanding and interdisciplinary research questions at different levels of abstraction. The MAS research community has addressed these issues from different perspectives that have gradually become more cohesive around the four notions that give title to the workshop: coordination, organization, institutions, and norms.

Virginia Dignum and Eric Matson were the cochairs of this workshop. The papers are published as an AAAI Technical Report WS-08-03.

Bringing Intelligence to E-mail

It has been more than a decade since the research community first became interested in e-mail overload. In that time, the HCI community has explored the effect of e-mail overload and tools for combating the problem. Five years ago, DARPA launched a major research initiative under the personal assistant that learns (PAL) program, which encompasses the cognitive assistant that learns and organizes (CALO) and reflective agents with distributed adaptive reasoning (RADAR) programs. These projects created intelligent systems that assisted users in managing large amounts of information. At the same time, efforts have emerged both at startups and in in-

dustry research labs to devise solutions to the e-mail overload problem. The result is an emerging community focused on intelligent tools for e-mail analysis and assistance, spanning diverse research areas including AI and HCI. Recognizing that new forms of communication, such as instant messaging and blogs, are becoming more prevalent, e-mail overload is now more generally known as information overload. The Enhanced Messaging workshop brought together a diverse group of researchers from academia and industry to discuss recent trends in messaging research and how we can address the increasing problem of information overload.

Gabor Cselle and Greg Duffy of Xobni Inc, an e-mail startup, spoke about their application, which improves e-mail organization, search, and navigation. They presented direct feedback from end users about how intelligent technologies improve the e-mail experience.

The 15 papers and posters at the workshop covered several broad themes. Building intelligent interfaces for e-mail enables new tools for e-mail management was one of these themes. Intelligent e-mail tools have been the focus of many researchers in recent years, represented by papers on recipient recommendation and leak detection and suggesting reusable replies.

Understanding e-mail as the modern information center yields an understanding of the modern information worker. Such analyses are useful for understanding the workplace and for mining useful information about an organization. For example, Christopher Diehl, Galileo Namata, and Lise Getoor presented techniques for identifying organizational relationships based on e-mail activity, a key part to understanding workplace roles. Andrew Lampert, Robert Dale, and Cécile Paris analyzed e-mails to learn how users make and commit to requests, which helps in understanding the nature of work and how users communicate about it. These techniques are also applicable to the emerging interest in e-discovery, where legal teams are responsible for analyzing millions of documents for relevance to litigation.

The PAL program focuses on learning procedures to complete user tasks, creating an automated assistant. Melinda Gervasio and Thomas J. Lee demonstrated the ability of the CALO system to learn these procedures from user actions, automating common tasks stemming from e-mail management.

Finally, a difficult aspect of e-mail research is a lack of resources due to the private nature of the data. Workshop participants discussed ideas for new data sources and annotations, useful for applications such as summarization (Jan Ulrich, Gabriel Murray, and Giuseppe Carenini).

In addition to strong participation from the research community, several industry representatives in attendance indicated an increased interest in deploying such technologies to combat overwhelmed users, a desire that has led to the recent founding of the Information Overload Research Group. The formation of this workshop indicates a continued interest in AI research on information overload in e-mail as well as the desire for a transition from research into workable technologies that can be deployed for information workers.

While the growing interest in e-mail has left a fractured community spread through many subareas, the workshop helped bridge this gap. Participants from many subfields of AI as well as the broader research community came together to discuss the state and future of the field. The workshop was an important first step toward building a community structure that will open channels of communication and collaboration as we move forward.

Mark Dredze (University of Pennsylvania), Vitor Carvalho (Microsoft Live Labs), and Tessa Lau (IBM Almaden Research Center) organized the workshop. The papers were published as an AAAI Technical Report WS-08-03.

Human Implications of Human-Robot Interaction

This workshop claims a history of its own, having consecutively been an element of the AAAI-06, AAAI-07, and AAAI-08 conferences. Although study of HRI certainly was in progress prior to celebration of AI's 50th birthday in 2006, this particular series of recent workshops has emphasized a distinctive focus of interest within that domain. Our interest was summarized concisely in a graphic that we used for this year's workshop call for papers and onsite conference poster. Created by Linda Pope, a United Methodist pastor, the graphic depicts a human looking into a hand mirror—but seeing the face of a humanoid robot in the mirror; the same robot, in an adjacent frame, sees the human's face in the mirror that it is holding. Immediately, of course, the graphic reminds us that an especially intimate and potent dynamic of interaction emerges as AI achieves greater capability to equip robotic artifacts with lifelike behavior. Indeed, the phenomenon of HRI may progressively become a process of coevolution.

Some familiar concepts of morality notably have illustrated this coevolutionary possibility in each of our three workshops. At least two of the presentations in 2006, for example, effectively suggested that sustained HRI with certain kinds of social robots could encourage humans to replace a historical concept of categorically free moral action with a much different notion of purely deterministic moral behavior. This distinction was revisited in one of the 2008 presentations, which observed that accepting a philosophical thesis of epiphenomenalism sets aside intensionality worries that the distinction presumes. In any case, philosophical positions that we adopt apparently do strongly condition the ways in which humans will interpret the moral status of robots—and themselves.

Again, it has been noted in both the AAAI-07 and AAAI-08 workshops that the enterprise of equipping robots to exhibit what humans regard as acceptable moral behavior reflexively can produce better understanding of our own moral ideas. In fact, a machine ethics presentation during the 2008 workshop demonstrated the ability of an AI system to infer new ethical principles that are meaningful.

Another multidisciplinary topic that has appeared repeatedly in our workshops is a concern with human acceptance of robots in specific HRI scenarios. More than half of our AAAI-06 workshop presentations engaged

this topic, reporting some results from empirical studies. This theme of interest continued in 2007, with presentations that explored acceptance in the context of privacy issues raised by robotic butlers, as well as the case of diminished human-to-human contact within healthcare applications. The AAAI-08 workshop also reflected interest in acceptance. Presentation of a paper reviewing Masahiro Mori's Buddhist perspectives on robotics was followed by some discussion of his uncanny valley hypothesis, according to which human acceptance drops sharply just when a robotic artifact becomes too lifelike. Results from cluster analysis of survey data also were presented, examining patterns relating responses in the areas of ethics, religion, and acceptance. Again, the 2008 workshop engaged human acceptance in the context of human trust of autonomous systems.

Perhaps it is not surprising that themes of ethics and acceptance have been conspicuous in this series of workshops. As our robotic creations become more lifelike, the dynamics of mutual adaptation between human and machine might naturally be expected to mirror those that take place among humans. After all, we commonly care very much whether a stranger is morally good and socially acceptable.

Lundy Lewis and Ted Metzler (chair) organized this workshop. The papers of the workshop were published as AAAI Technical Report WS-08-05 and are available from AAAI Press.

Intelligent Techniques for Web Personalization and Recommender Systems

Web personalization can be generally defined as the process of tailoring the content or visual presentation of a website, or the behavior of a web application, to the particular needs and preferences of an individual user or a group of users. Automated personalization systems typically accomplish this task through individualized information filtering and ranking, as well as through intelligent navigation support. They thereby rely on different sources of knowledge such as clickstream data, web usage logs, or explicit personalization rules.

Recommender systems represent one special and prominent class of personalized web applications. They focus particularly on the user-dependent filtering and selection of relevant information and aim to support online users in the decision-making and buying process. Recommender systems have already been a subject of extensive research in AI over the past decade. With today's soaring number of e-commerce environments on the web, the demand for new approaches to intelligent product recommendation has become more pressing than ever. The web now has more online users, more online channels, more vendors, more products, and, most of all, more complex products and services. These recent developments in the area of recommender systems generated new research demands, in particular with respect to interactivity, adaptivity, and user preference elicitation. These research challenges, however, are not confined to the area of recommender systems only, but are also in the focus of general web personalization research.

In the face of this increasing overlap of the two areas, the aim of this workshop was to bring together researchers and practitioners of both fields, to foster an exchange of information and ideas, and to facilitate a discussion of current and emerging topics related to web intelligence, particularly regarding its application to recommender systems. The workshop united topics from web personalization and recommender systems after a successful prior event at AAAI-07.

This year's workshop attracted a number of high-quality contributions from 12 different countries. Of these, seven papers (fewer than 40 percent) were accepted for full presentation, with an additional three accepted for short presentations. The accepted papers dealt with a variety of issues and techniques for creating more intelligent personalization systems, but generally fell into three broad categories. In the area of personalization in search and navigation, new proposals were made for the ranking of personalized search results based on machine learning, for client-side user monitoring to infer the user's search intent, as well as for community-based query expansion and personalized document filtering. In the personalization on the social web track, presenters addressed aspects of search and navigation improvement in social media platforms through tag clustering and tag-based user profiling, as well as the issue of possible attacks against such open tagging systems. Finally, in the technical track on recommender systems technology, the focus was on combining rating information with additional knowledge about item characteristics or about user relationships in a social network, to improve prediction accuracy.

In keeping with the increasing usage of social web and personalization features on media-sharing sites, this year's workshop also included an exciting invited address by Paul Lamere of Sun Microsystems Labs, titled, "Music Recommendation Is Broken, and Only You Can Fix It!"

The workshop ended with an open discussion and a reflection on the state of the art in the respective areas. The exploitation of additional knowledge sources to improve personalization and recommendation was identified as one of the central means for advancing research in web personalization and recommender systems. Such new sources include content-related information, for example, in the form of lexicons, semantic web ontologies, or web 2.0 community knowledge. They also include context-related information, such as the user's geospatial position, which, for example, the latest generation of location-aware handheld devices can provide to the personalization system.

The Workshop on Intelligent Techniques for Web Personalization and Recommender Systems was cochaired by Sarabjot Singh Anand, Bamshad Mobasher, Alfred Kobsa, and Dietmar Jannach. The papers of the workshop were published as AAAI Press Technical Report WS-08-06.

Metareasoning: Thinking about Thinking

The two-day Workshop on Metareasoning: Thinking about Thinking was

a sequel to the successful Workshop on Metareasoning in Agent-Based Systems at AAMAS-07 last year. This year the cochairs wrote a brief manifesto outlining a simple model of metareasoning and invited participants to compare and contrast their research to the conceptualization. The contention is that, like traditional reasoning that is composed of an action and perception cycle, so too metareasoning has distinct control and monitoring components. The resulting paper submissions were both exciting and novel, falling into one of four categories. Some papers examined the metalevel control of reasoning, while others discussed introspective monitoring of reasoning. Another set of papers reported research on distributed models of metareasoning and multiagent metareasoning that considered the issues of coordination at the metalevel and when and how agents could gain a common metalevel context for problem solving. Finally, a large group of papers put everything together to talk about computational models of self. Day one organized paper sessions from the first two groups, whereas day two contained the presentations for the latter two groups.

In addition to the standard 25-minute presentation and question periods, all sessions were followed by an interactive panel where each author presented views and comments regarding topics put forth by a select moderator. These sessions also gave the audience an opportunity to ask questions or make comments on issues that spanned individual paper presentations.

A special highlight of the workshop was the invited talks. On day one Don Perlis of the University of Maryland, College Park, spoke of two conceptions of reflective reasoning. The first one he called "Meta-Strata," referring to hierarchical architectures such as the one proposed in the workshop manifesto. Here one discrete layer lies above another and reasons about its workings. In another arrangement, called "Meta-Loopy" (in deference to Doug Hofstadter), an architecture examines itself as a snake biting its own tail. Much lively discussion erupted over the differences and similarities between these formulations.

On day two Aaron Sloman of the University of Birmingham in the UK discussed varieties of metacognition in natural and artificial systems. His talk was based on the requirements analysis in the CoSy robotics project and discussed various types of metacognition in intelligent biological individuals. He discussed the role of the environment in the development of natural intelligence where the environment includes the ways in which things outside that organism, for example, types of interactions, relate to design features of the organism. Several interesting videos of metacognition in natural systems were screened, and the implications of these metacognitive issues for architectures and representations were discussed.

The workshop participants discussed the commonalities and need for differences among the various metareasoning architectures presented at the workshop. The discussions also included whether a common model for metareasoning in the manifesto was sufficient and the implications of enforcing such a model. There was also discussion on the complexity of solving various metareasoning issues. The group reached a general consensus that metareasoning is an exciting area requiring more investigation. There was also interesting discussion on methods and metrics to evaluate the effect of metareasoning on overall system performance.

AAAI Press published the papers from this workshop as AAAI Technical Report WS-08-07. The collection is available in hard copy from the publisher and electronically from AAAI's digital library. The cochairs of this workshop were Michael T. Cox (BBN Technologies) and Anita Raja (University of North Carolina at Charlotte).

Advances in Preference Handling

Preferences are a central concept of decision making. As preferences are fundamental for the analysis of human choice behavior, they are becoming of increasing importance for computational fields such as artificial intelligence, databases, and human-computer interaction. Preference models are needed in decision-support systems such as web-based recommender systems, in automated problem solvers such as configurators, and in autonomous systems such as Mars rovers. Nearly all areas of artificial intelligence deal with choice situations and can thus benefit from computational methods for handling preferences. Preferences are also at the heart of social choice methods and are thus of importance for consensus methods, which are, for example, used in bioinformatics, and for multiagent systems.

These new perspectives on preferences led to the creation of a growing community of researchers from artificial intelligence, databases, operations research, and other computational fields who are interested in computational models of preferences and their applications to computational tasks. This workshop had an exciting program consisting of an invited talk, 23 technical presentations, a poster session, and a panel discussion.

The topic of preference elicitation found particular interest at the workshop. Craig Boutilier summarized major challenges and experiences about this topic in a fascinating invited talk. A particular challenge consists in keeping the number of questions about user preferences small while giving a guarantee about the degree of optimality of the recommendation produced by an interactive system. As shown by Boutilier and other participants, it is necessary to interleave elicitation and problem solving to achieve good results. A couple of other tasks discussed preference elicitation in concrete systems such as a smart home system that adapts to changes in a resident's preferences and multiagent systems for surveillance tasks. Preferences can also be acquired by observing an agent's behavior, in particular when the agent's environment is changed.

Another highlight of the workshop was the usage of preferences in game theory. Milind Tambe's group presented their work on Bayesian Stackelberg games, which is used for security scheduling at the Los Angeles International Airport. These games distinguish different players, namely a leader and followers, and pose particular challenges for preference modeling. Game theory was also the topic of

several other talks. Furthermore, Vincent Conitzer organized an excellent panel about whether game theory is necessary or beneficial for research on multiagent preferences. The panelists were Craig Boutilier, David Parkes, Yoav Shoham, and Milind Tambe. One conclusion was that game theory is important but should be used with care, as it is essentially based on the notion of an equilibrium.

The workshop also covered topics such as preference representations in form of graphical models or soft constraints, preference queries in databases and ontologies, and preference representations in planning and in combinatorial auctions. It was interesting to learn that preferences can also be represented in the form of rules such as those for change management in the service industry or for product configuration. Consensus methods in bioinformatics were illustrated for the problem of identifying sibling relationships from genetic data. A poster session allowed students to get feedback about their thesis projects.

The workshop was cochaired by Jan Chomicki, Vincent Conitzer, Ulrich Junker, and Patrice Perny. The papers of the workshop were published as AAAI Press Technical Report WS-08-09.

Search in Artificial Intelligence and Robotics

Search is one of the few areas of artificial intelligence (and beyond) that lack their own conference. This workshop therefore brought together researchers interested in this topic to share their ideas and disseminate their latest research results. It focused on finding common ground between search techniques used in artificial intelligence and robotics.

Heuristic search and related algorithms are currently very active areas of research. For example, researchers investigate how to search in real time, how to search with limited (possibly external) memory, how to search in parallel on several processors, how to solve sequences of similar search problems faster than with isolated searches, how to improve the run time of the searches through randomization or learning techniques, how to discretize continuous state spaces, how to trade off between the run time and memory consumption of the search and the resulting solution quality, how to select between different search strategies, and how to focus the searches with sophisticated heuristics such as pattern databases. Their results are published in different conferences such as IJCAI, AAAI, ICAPS, NIPS, ICRA, and IROS. This workhop brought these researchers together to exchange their ideas, crossfertilize the field, and combine various search techniques that originated in different research communities.

The two-day workshop had more than 35 attendees, in part thanks to generous support from NSF for student participation. It featured an overview that highlighted the similarities and differences of search in artificial intelligence and robotics and three invited talks (by Oliver Brock, Malte Helmert, and Maxim Likachev) on "Solving Hard Planning Problems in Robotics with Simple A*-like Searches," "Automatically Deriving Abstraction Heuristics," and "Search in Embodied Artificial Intelligence and Computational Biology."

The 15 oral presentations and more than 12 posters in a lively poster session displayed the diversity of research on search and its applications, covering topics such as abstraction, inconsistent heuristics, bounded suboptimality, performance prediction, learning, symmetry, real-time search, moving-target search, connections to probabilistic reasoning and applications to robotics, machine learning, and diagnosis. Among the highlights of the workshop were the presentations on the use of heuristic search in the first- and second-place vehicles participating in the DARPA Urban Challenge.

The organizing committee of the symposium consisted of David Furcy, Sven Koenig, Wheeler Ruml, and Rong Zhou. The papers of the workshop were published as AAAI Press Technical Report WS-08-10.

Spatial and Temporal Reasoning

The Workshop on Spatial and Temporal Reasoning was held on Sunday, June 13, the first of the two workshop days of the 23rd AAAI Conference on Artificial Intelligence (AAAI-08). It continued a tradition that started with a workshop 15 years ago at IJCAI-03 in Chambery, France, which led to a series of workshops with the common goal of bringing together related communities of researchers with an interest in the study of representing and reasoning about either space or time—or both.

As is the case with many other reasoning techniques, spatial and temporal reasoning is at home in many areas of artificial intelligence (and computer science in general), such as planning, robot control and guidance, natural language understanding, assembly plant sequencing and scheduling, ambient intelligence and smart homes, temporal databases, concurrent and distributed programming. The field of spatial and temporal reasoning has progressed significantly over the recent years, and some of the long-standing problems in the field have at least partially been solved, in particular those related to tractability for spatial calculi, explicit construction of models, characterization of important subclasses of relations, multidimensionality of spatiotemporal calculi, and handling of incomplete and imprecise information. Despite all these successes, there is still a lack in a deeper understanding of the foundations of the field, which might be the reason that it has not found as much enthusiasm among the practitioners in artificial intelligence, computer science, and information technology as it should have had. The aim of this workshop was to work towards overcoming this shortcoming.

A total of seven presentations spread over the whole day laid the basis for the workshop. The presentations were slightly biased towards spatial reasoning (as opposed to temporal reasoning), and one of them addressed both. However, due to the similarity of spatial and temporal reasoning, most of the presentations had an impact on both spatial and temporal reasoning. They provided a solid basis for the discussion sessions, which were interleaved with the presentation sessions. The discussion sessions tied together the individual presentations into larg-

er themes and provided a forum for clarification of ideas, exchange of points of view, assessment of results and methods, and suggestions for future work.

Although progress was made in various areas of spatial and temporal reasoning, it became obvious that there is still a large number of open problems. In particular, there is still a gap between the theories of spatiotemporal reasoning and their applications to real-world scenarios, which requires a significant amount of future work in the field.

Hans Guesgen (Massey University, New Zealand), Gérard Ligozat (LIMSI, Université Paris-Sud, France), and Rita V. Rodriguez (National Science Foundation, USA) served as cochairs of this workshop. This report was written by Hans Guesgen. The papers of the workshop were published as AAAI Press Technical Report WS-08-11.

Trading Agent Design and Analysis

Research in trading agent technologies has gained increasing prominence over the past decade, in part due to a drive to partially automate trading decisions in a number of different domains. This workshop focused on the design and evaluation of trading agents. Papers were invited on topics in trading agent architectures, decision-making algorithms, theoretical analysis, and empirical evaluations of agent strategies in different negotiation scenarios.

The workshop was held in conjunction with the finals of the 2008 Trading Agent Competition. Two game scenarios and two challenge events attracted 42 entries.

The supply-chain management (TAC-SCM) scenario places six agents in the role of a personal-computer manufacturer. Each agent has to procure PC components and sell finished goods in competitive markets while managing inventory and production facilities. The baseline TAC-SCM competition was also complemented by two additional challenge events: (1) A procurement challenge that requires agents to manage supply chain risk through a combination of long-term and one-off procurement contracts, and (2) a prediction challenge designed to test price-prediction capabilities of competing agents in both procurement and sales markets.

In contrast to the supply-chain scenario, which casts the competing agents as traders, the CAT scenario places agents in the role of competing exchanges. The CAT competition is motivated by the rise of independent for-profit stock and commodity exchanges that compete for the attention of traders. CAT agents compete by defining rules for matching buyers and sellers and by setting commission fees for their services. Profitability is the ultimate measure of performance in both the supply chain and CAT scenarios. Both scenarios involve multiple rounds, each of which features a large number of encounters aimed at capturing a broad range of market conditions.

The workshop featured seven (six extended and one short) paper presentations, in-depth discussions of the performance of competing techniques and agents, and two panel discussions. Paper presentations included in-depth analyses of the 2007 TAC-SCM procurement and prediction challenges. Designers of the prediction challenge emphasized how their challenge makes it possible to isolate the prediction performance of different agents in different areas, whereas the TAC-SCM scenario effectively allows agents to compensate for inaccuracies in their predictions by adjusting their procurement and sales activities. The procurement challenge presentation discussed lessons learned from the 2007 edition and provided a comparison of the top three entries and a discussion of changes introduced in the 2008 edition of the challenge. Another paper presented a survey of agent designs in TAC SCM. The survey showed that, in some areas such as modularity, there are common themes emerging in how to design a successful trading agent, while in other areas, such as coordination, there are strong differences in the designs. Another presentation revolved around an experimental study of bidding heuristics designed for the TAC Travel game (a game that was not part of the 2008 competition), showing that using as much distributional information as possible is an effective approach for an agent in one-shot simulated auctions settings. Another paper discussed how to classify bidding strategies in CAT and reported that using a hidden Markov model yields the best results. A second paper on CAT discussed how to design an effective e-market and provided an overview of the authors' CAT agent. One paper discussed the problem of multiunit multiattribute allocation through call auctions. The paper demonstrated that an iterative bidding protocol can often overcome the limitations of bidding language restrictions made to achieve clearing tractability.

The first panel discussed the benefits and challenges of developing mixed initiative variations of existing scenarios, where human decision makers would compete with the support of semiautonomous trading agents. The panel gave rise to lively discussions and generally suggested that the trading community would likely welcome such a game, most probably as an extension of TAC-SCM procurement challenge. Panel participants generally viewed the introduction of such a game as a possible way of engaging other subdisciplines both within and outside of the AI community. A game like this would also likely be an excellent teaching tool in business schools. Simply taking the procurement challenge and requiring a human decision maker to replace an agent would not work, given the amount of information the human would have to make sense of and the number of decisions she would have to make. Instead a more promising approach would be to find ways to put human decision makers in control of important decisions while allowing them to delegate more routine day-to-day pricing, procurement, and resource allocation decisions to semiautonomous trading agents operating under their overall strategic guidance.

The day ended with a panel and group discussion on the future of trading agent research, and to what extent results from TAC are influencing current practice. Discussions suggested that TAC has started to make an impact. Specifically, it appears that TAC-like techniques are being implemented

in several practical settings including energy markets, flower auctions, and procurement of electronic components. Mixed-initiative interfaces are an important issue in these deployments.

Wolfgang Ketter, Norman Sadeh, and William Walsh were organizers of this workshop. The papers were published as AAAI Technical Report WS-08-12. Additional information on the 2008 TAC competition can be found at www.sics.se/tac.

Transfer Learning for Complex Tasks

All machine-learning algorithms require data to learn, and often the amount of data available is a limiting factor. For instance, classification and regression require labeled data, which may be expensive to obtain. Reinforcement learning requires samples that must be collected through repeated interaction with an agent's environment. Typically, a learning system or agent treats every problem as distinct and must begin learning tabula rasa. The insight behind transfer learning is that past experience may assist learning a novel task, even if the tasks are very different. While the idea of transfer has long been explored in the psychological literature, it has only recently been gaining popularity as a general machine-learning technique.

In the transfer learning paradigm, one typically uses a set of source tasks to help learn one or more target tasks. Successful transfer allows faster, or better, learning, when compared to learning without previous knowledge, even if the source and target task data originate from different distributions.

The primary goal of this workshop was to bring together researchers working on different aspects of the transfer problem so that we could discuss current approaches and common problems. Roughly 30 researchers attended to discuss the 10 accepted papers. Papers covered a wide range of topics, including regression, classification, reinforcement learning, planning, Markov logic networks, and neural networks. Despite the large number of contexts, many of the same transfer-related questions were discussed by the different presenters.

One such open question is how to best determine task similarity autonomously. In one of the three reinforcement learning papers, Tom Croonenborghs, Kurt Driessens, and Maurice Bruynooghe presented a novel approach to learn how state variables are related in different reinforcement learning tasks. Research in transfer for reinforcement learning typically limits transfer to tasks that have the same state variables or relies on a human to specify how the tasks are related. In contrast, this approach may allow an agent to transfer between very different tasks without requiring a human in the loop.

Another long-term goal of transfer has been to enable successful knowledge reuse between very different tasks. In their work, Lilyana Mihalkova and Raymond Mooney show that Markov logic networks can successfully transfer between different domains. For instance, their method can exploit similarities between the social organization in academia and that in the movie industry to achieve transfer when target domain data is severely limited. Jesse Davis and Pedro Domingos also use Markov logic networks, but explicitly focus on deep transfer, where the domains transferred between are even more different, such as using molecular biology data to learn better in an academic domain. Experiments show that learned network templates, representing concepts like symmetry, transitivity, and homophily, enable significant improvements when learning in novel domains.

Will Bridewell and Ljupco Todorovski's paper had a similar deep transfer goal, but in a very different setting. The goal of inductive process modeling is to produce a model that explains the behavior of a dynamic system and predicts unseen data. After successfully learning on data from one ecosystem, the authors use transfer to learn a model of a second ecosystem, which had different characteristics and different organisms. Experiments showed that transferred constraints could reduce search time by an order of magnitude with little loss to model accuracy.

After the last presentation, attendees participated in a general debate on the relative merits of transfer. Topics included differentiating transfer from generalization and what goals are appropriate for deep transfer. While no firm conclusions were reached, the consensus was that transfer learning is a relatively young field with many open questions. Current results suggest that transfer can lead to substantial performance improvements in many different machine-learning contexts. Going forward, we expect that transfer will continue to generate many questions, as well as opportunities, for AI researchers.

Matthew E. Taylor wrote this report and was the primary contact for the workshop. The papers were published as AAAI Technical Report WS-08-13.

What Went Wrong and Why

Unfortunately, bugs, glitches, and failures are rarely mentioned in academic discourse, so their role in informing design and development is essentially lost. The first What Went Wrong and Why workshop addressed this gap at the 2006 AAAI spring symposium by inviting AI researchers and system developers to discuss their most revealing bugs and relate problems to lessons learned. Several of the articles and invited talks were published as a special issue of the summer 2008 *AI Magazine*. The second What Went Wrong and Why workshop continued this theme through a one-day program at AAAI-08 that emphasized methodological insights. It included invited talks by Kevin Ashley, Bruce Buchanan, Steve Chien, and Haym Hirsch, plus four papers.

Kevin Ashley focused on evaluating research in computational argumentation. He incorporated computational models of argumentation into his curriculum for first- and second-year law students (whose business is [arguably] to argue), and asked if existing models helped them learn argumentation skills. He discovered that learning in this realm was as hard to measure as evaluating argumentations. In the process, however, he developed a suite of diagnostic tools that led to better-targeted pedagogical advice. This story is a reminder that the methodology can be as important as the intended result.

Bruce Buchanan's talk examined what we learned from the expert system's boom. After noting the power and prevalence of such systems, he focused on cases where the reality of commercialization forced developers to address hard problems in AI and where pragmatic shortcuts let them avoid other problems. For example, the expectation that nonexperts could write rules proved largely false, but led to the designation of *knowledge engineers,* and clarified the target end user for expert system shells. The lesson that some knowledge requires intricate and subtle representational structures clarified the need for new technology without diminishing the role of flat rule bases or their utility.

Steve Chien discussed lessons learned for AI applications from autonomous science craft, drawing on his experiences at JPL/NASA. He noted that work on such large-scale integrated systems presents both organizational and technical challenges. From an organizational perspective, the key issue is to balance autonomy, seen as risk, against benefits in the form of cost reduction and scientific returns. This trade-off is evaluated by hard-nosed engineering calculations but has a social aspect as well. AI systems in space applications need to build a track record of trust before they can be deployed. Steve showed the potential for enormous returns: software onboard NASA's Earth Observing One (EO-1) mission documented more than a 10-fold increase in science return and more than $1M/yr in cost reductions. He concluded by noting that AI software has flown on five missions, and its successes are changing the acceptability of spacecraft AI.

Haym Hirsch examined what's going wrong in data mining from his perspective as a researcher and as the director of the Information and Intelligent Systems Division at NSF. He observed that the examples motivating large bodies of academic work have become dangerously out of date, and shape research in inappropriate ways. For example, the Irvine machine-learning repository supports incremental improvements in classification algorithms, while application challenges concern terabyte information extraction tasks over distributed sources that are riddled with incorrect and incomplete data. Haym argued that these elements transform the nature of the relevant research, and that we should review our examples in this, and other areas lest they blind us to a changing world.

Among the papers, Carl Hewitt discussed the history of logic programming in terms of the issues and responses characterizing its development. He concluded that paraconsistent logics are the new horizon, as they can plausibly infer properties of software written for practical domains, which is chock full of inconsistencies. Nestor Rychtyckyj and Alan Turski provided case studies in the development of commercial expert systems and concluded that organizational versus technical issues constituted the key barriers to acceptance (for example, the presence/absence of software life-cycle support mechanisms). Soumi Ray and Tim Oates presented an unusual what went right and why story, which asked the audience to help explain unreasonably fast convergence from a reinforcement learning algorithm that randomly scaled Q-values. The resulting discussion is still in process. Finally, Cindy Marling and David Chelberg discussed an unsuccessful team's entry into an international competition. The paper, titled "RoboCup for the Mechanically, Athletically, and Culturally Challenged," noted that stuff blew up, communication systems failed, the professor broke her ankle developing domain expertise, and the principles narrowly escaped arrest for shipping compressed gas. Despite these setbacks, the team concluded that failure is a catalyst for future progress.

In summary, this workshop provided researchers and developers with an informal, valuable, and enjoyable opportunity to share their experiences about What Went Wrong and Why. It illustrated, once again, that WWWW experiences offer novel insights and communicate interesting research lessons in concise ways. Mehmet H. Göker and Daniel Shapiro cochaired the workshop. The papers were published as AAAI Technical Report WS-08-14.

Wikipedia and Artificial Intelligence: An Evolving Synergy

As a large-scale repository of structured knowledge, Wikipedia has become a valuable resource for a diverse set of AI applications. Major conferences in natural language processing and machine learning have recently witnessed a significant number of new approaches that use Wikipedia for tasks ranging from text categorization and clustering to word-sense disambiguation, information retrieval, information extraction, and question answering. On the other hand, Wikipedia greatly benefits from numerous algorithms and representation models developed during decades of AI research, as illustrated recently in tasks such as estimating the reliability of authors' contributions, automatic linking of articles, and intelligent matching of Wikipedia tasks with potential contributors.

Consistent with the aims of the workshop, the paper presentations addressed a highly diverse set of problems, in which Wikipedia was seen either as a useful resource or as a target for algorithms seeking to make it even better. As a rich knowledge source, Wikipedia was shown to benefit applications in information extraction, machine translation, summarization, ontology mining and mapping, and information retrieval. We also learned of interesting applications, most of them using machine learning, that could further enhance the breadth and the quality of Wikipedia, such as predicting the quality of edits, vandalism detection, infobox extraction, creation of crosslingual links, and semantic annotation.

The workshop featured an invited talk by Michael Witbrock (Cycorp) on human-computer collaboration. Based on the premise that it is not always possible to employ humans to verify the correctness of an ever increasing number of Wikipedia submissions, Witbrock presented arguments for a large scale use of artificial agents that verify facts by repeatedly observing the behavior of others, assuming repeated behaviors and facts to be correct.

The workshop concluded with an

exciting panel discussion. Jamie Taylor (Metaweb) gave a short presentation on Freebase, a collaboratively edited database of world knowledge that derives its content from Wikipedia and other knowledge bases. He then argued for a better understanding of semantic issues within the Wikipedia community, as a necessary step towards turning Wikipedia into a semantic network. He also pointed to the challenge of transferring the current techniques for mining Wikipedia to the significantly noisier World Wide Web. Barney Pell (Powerset) drew attention to the rate of increase of Wikipedia submissions, which lately has been going down. As a possible solution, he proposed automatically creating initial versions of stub articles in order to motivate users to edit them. He also emphasized the need to improve the techniques for extracting knowledge from Wikipedia and crosslinking it with other knowledge sources. Michael Strube (EML Research) proposed going beyond the extraction of factual knowledge from Wikipedia in order to distill richer information such as opinions, procedural knowledge, and even scientific insights. Daniel Weld (University of Washington) presented an approach to populating infoboxes with information extracted from Wikipedia articles and the World Wide Web. He emphasized the importance of unobtrusively motivating users to contribute expertise. In this context, he described the approach employed by his group where online ads were placed on Google, Yahoo, and MSN in order to invite volunteers to verify automatically extracted facts. Other issues raised during the panel discussion included integration of information from multiple pages when showing search results, studying what users actually want from Wikipedia, allowing contributors to define and annotate relations in Wikipedia, and more generally assisting contributors to provide deeper semantic annotations.

Razvan Bunescu, Evgeniy Gabrilovich, and Rada Mihalcea served as cochairs of the workshop. The papers of the workshop were published as AAAI Press Technical Report WS-08-15.

Sarabjot Singh Anand is a senior lecturer in the Department of Computer Science, University of Warwick.

Razvan Bunescu is an assistant professor at the School of Electrical Engineering and Computer Science, Ohio University.

Vitor R. Carvalho is a scientist at Microsoft Live Labs.

Jan Chomicki is an associate professor of computer science and engineering at the University at Buffalo.

Vincent Conitzer is an assistant professor of computer science and economics at Duke University.

Michael T. Cox is a senior research scientist in the Intelligent Computing group of BBN Technologies.

Virginia Dignum is an assistant professor at Utrecht University.

Zachary Dodds is an associate professor of computer science at Harvey Mudd College.

Mark Dredze is a fifth year Ph.D. student in computer science at the University of Pennsylvania.

David Furcy is an assistant professor at the University of Wisconsin.

Evgeniy Gabrilovich is a senior research scientist at Yahoo! Research.

Mehmet H. Göker is the research director at PricewaterhouseCoopers Center for Advanced Research.

Hans Guesgen is a professor of computer science in the School of Engineering and Advanced Technology at Massey University, New Zealand.

Haym Hirsh is a professor of computer science at Rutgers University.

Dietmar Jannach is a professor at the Department of Computer Science, Dortmund University of Technology.

Ulrich Junker is a distinguished scientist at ILOG.

Wolfgang Ketter is an assistant professor at the Rotterdam School of Management, Erasmus University.

Alfred Kobsa is a professor in the Donald Bren School of Information and Computer Sciences of the University of California, Irvine.

Sven Koenig is an associate professor of computer science at the University of Southern California.

Tessa Lau is a research staff member at IBM Almaden Research Center.

Lundy Lewis is a professor and chair of the Department of Computer Information Technology, at Southern New Hampshire University.

Eric Matson is an assistant professor in the Computer and Information Technology Program, College of Technology, Purdue University.

Ted Metzler is director of the Darrell W. Hughes Program for Religion and Science Dialogue at Oklahoma City University.

Rada Mihalcea is an associate professor at the Department of Computer Science and Engineering, University of North Texas.

Bamshad Mobasher is a professor at the School of Computing, DePaul University.

Joelle Pineau is an assistant professor of computer science at McGill University.

Pascal Poupart is an assistant professor at the School of Computer Science University of Waterloo.

Anita Raja is an assistant professor in the Department of Software and Information Systems at the University of North Carolina at Charlotte.

Wheeler Ruml is an assistant professor at the University of New Hampshire.

Norman Sadeh is a professor at Carnegie Mellon University.

Guy Shani is a researcher with Microsoft Research.

Daniel Shapiro is the executive director of the Institute for the Study of Learning and Expertise and president of Applied Reactivity, Inc.

Trey Smith is a systems scientist at Carnegie Mellon University West/NASA Ames Research Center.

Matthew E. Taylor is a postdoctoral research scientist at the University of Southern California

Kiri Wagstaff is a researcher in the Machine Learning Systems Group at the Jet Propulsion Laboratory.

William Walsh is a senior research scientist at CombineNet.

Rong Zhou is a researcher at the Palo Alto Research Center.