AI Magazine Volume 27 Number 4 (2006) (© AAAI)

Reports on the Twenty-First National Conference on Artificial Intelligence (AAAI-06) Workshop Program

Wolfgang Achtner, Esma Aimeur, Sarabjot Singh Anand, Doug Appelt, Naveen Ashish, Tiffany Barnes, Joseph E. Beck, M. Bernardine Dias, Prashant Doshi, Chris Drummond, William Elazmeh, Ariel Felner, Dayne Freitag, Hector Geffner, Christopher W. Geib, Richard Goodwin, Robert C. Holte, Frank Hutter, Fair Isaac, Nathalie Japkowicz, Gal A. Kaminka, Sven Koenig, Michail G. Lagoudakis, David Leake, Lundy Lewis, Hugo Liu, Ted Metzler, Rada Mihalcea, Bamshad Mobasher, Pascal Poupart, David V. Pynadath, Thomas Roth-Berghofer, Wheeler Ruml, Stefan Schulz, Sven Schwarz, Stephanie Seneff, Amit Sheth, Ron Sun, Michael Thielscher, Afzal Upal, Jason Williams, Steve Young, and Dmitry Zelenko

■ The Workshop program of the Twenty-First National Conference on Artificial Intelligence was held July 16–17, 2006 in Boston, Massachusetts. The program was chaired by Joyce Chai and Keith Decker. The titles of the 17 workshops were AI-Driven Technologies for Service-Oriented Computing; Auction Mechanisms for Robot Coordination; Cognitive Modeling and Agent-Based Social Simulations, Cognitive Robotics; Computational Aesthetics: Artificial Intelligence Approaches to Beauty and Happiness; Educational Data Mining; Evaluation Methods for Machine Learning; Event Extraction and Synthesis; Heuristic Search, Memory-Based Heuristics, and Their Applications; Human Implications of Human-Robot Interaction; Intelligent Techniques in Web Personalization; Learning for Search; Modeling and Retrieval of Context; Modeling Others from Observations; and Statistical and Empirical Approaches for Spoken Dialogue Systems.

AAAI was pleased to present the AAAI-06 workshop program. Workshops were held Sunday and Monday, July 16–17, 2006, at the Seaport Hotel and World Trade Center in Boston, Massachusetts. The AAAI-06 workshop program included 14 workshops covering a wide range of topics in artificial intelligence. All but 2 of the workshops were held in a one-day timeslot. Each workshop was limited to approximately 25 to 75 participants. The 2006 workshop program was constructed to encourage dialogue and build bridges between researchers in different subfields. Several of the workshops were follow-ons from workshops held at more specialized AI conferences. Attendees were encouraged to attend and contribute to workshops that they were interested in even if they had no recent work directly in that area.

The titles of the 17 workshops were AI-Driven Technologies for Service-Oriented Computing; Auction Mechanisms for Robot Coordination; Cognitive Modeling and Agent-Based Social Simulations; Cognitive Robotics; Computational Aesthetics: Artificial Intelligence Approaches to Beauty and Happiness; Educational Data Mining; Evaluation Methods for Machine Learning; Event Extraction and Synthesis; Heuristic Search, Memory-Based Heuristics, and Their Applications; Human Implications of Human-Robot Interaction; Intelligent Techniques in Web Personalization; Learning for Search; Modeling and Retrieval of Context; Modeling Others from Observations; and Statistical and Empirical Approaches for Spoken Dialogue Systems.

AI Driven Technologies for Service-Oriented Computing

Service-oriented computing is an emerging computing paradigm for distributed systems that advocates web-based interfaces for the distributed business processes of any enterprise. The interfaces, called web services, hold the promise for diluting the traditional challenges of interoperability, inflexibility, and performance that have long plagued traditional distributed systems. Web services research represents an emerging application test bed with its own distinct challenges and presents an opportunity for AI techniques to enter and affect this emerging area.

While somewhat similar workshops

in the past have focused on the application of specific AI techniques in services-oriented computing, these did not bring together the broader AI community. The goal of this workshop was to investigate the application of a broad spectrum of AI techniques to services-oriented computing.

This full-day workshop featured 10 oral presentations and a 45-minute group discussion. All the selected papers were peer reviewed by members of the program committee. The speakers and participants represented several countries; diverse backgrounds that included academic, industrial, and defense agencies; and varying areas of expertise. The papers presented in the workshop addressed topics that included logics for reasoning about web services, symbolic and planning systems for composing web services, mixed initiative approaches for discovering and composing web services, and probabilistic techniques for adapting service compositions to dynamic environments. The group discussion at the conclusion of the workshop centered on the general and domain-specific challenges of developing realistic services-oriented architectures that form the technology infrastructure for enterprises. There was a general consensus among the participants about the need for such workshops that would facilitate a better awareness of the issues facing services-oriented computing among the AI community.

The papers presented at this workshop were published as an AAAI technical report (WS-06-01), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Prashant Doshi (LSDIS Lab, University of Georgia), Richard Goodwin (IBM T. J. Watson Research Center), and Amit Sheth (LSDIS Lab, University of Georgia).

Auction Mechanisms for Robot Coordination

Robot teams are increasingly becoming a popular alternative to single robots for a variety of difficult robotic tasks, such as planetary exploration or planetary base assembly. A key factor for the success of a robot team is the ability to coordinate the team members in an effective way. Coordination involves the allocation and execution of individual tasks through an efficient (preferably decentralized) mechanism. It is desirable to enable good decision making while communicating as little information as possible.

Recently, there has been significant interest in using auction-based methods for robot coordination, a trend that gave rise to our AAAI-06 workshop. In these methods, the communicated information consists of bids robots place on various tasks, and coordination is achieved by a process similar to winner determination in auctions. Auction-based methods balance the trade-off between purely centralized methods (full communication) and purely decentralized methods (no communication) in both efficiency and quality.

This emerging field of research has demonstrated significant progress in its few years of existence; however, heretofore there had been no official forum for involved researchers to share experience, establish foundations, and explore future directions. Our AAAI-06 workshop served precisely this purpose. It drew the leading researchers in this active area of research and gave them an opportunity to discuss and analyze both the practical aspects of the subject (distributed implementation, limited communication, target applications), as well as the theoretical ones (theoretical guarantees, computational complexity, communication complexity).

A total of 15 papers were presented either as talks (8 papers) or posters (7 papers). The compiled workshop report consists of 110 pages covering the current state of the art in the field and representing all aspects of auction-based coordination from sophisticated auction schemes and learned bidding strategies to applications in automated warehouse management and adaptive distributed sensor networks.

The workshop attracted about 30 participants from both academe and industry—a significant turnout for such a specialized topic. The workshop program was structured in a way that facilitated communication and discussion among workshop participants. Talks were followed by plenty of discussion, while posters were given the opportunity for short oral presentations before the poster session.

The highlight of the workshop was the invited panel discussion that concluded the workshop program. Our invited panelists were Tuomas Sandholm (Carnegie Mellon University), Alvin Roth (Harvard University), Brad Clement (NASA JPL), and Jeff Kephart (IBM Research). During the panel, panelists and participants had the opportunity to share their vision of auction-based coordination in the future, address practical design problems, and identify potential real-world target applications.

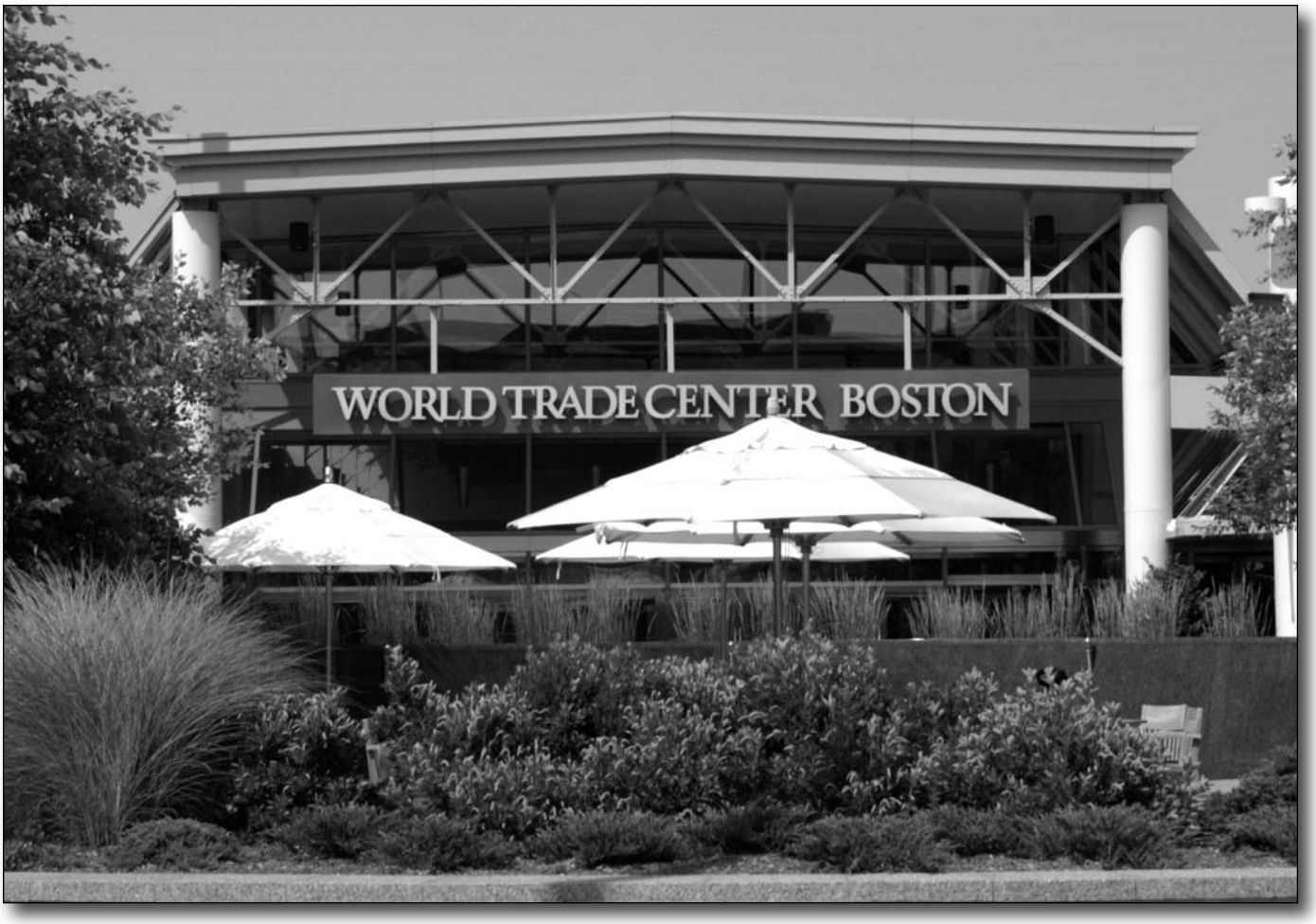
Jointly with other colleagues we offered an AAAI-06 tutorial on the same subject the day before the workshop. The tutorial and the workshop complemented each other and their combination was a perfect opportunity for AAAI attendees to familiarize themselves with this exciting field of research. It is our belief that this workshop helped to establish a new research community that has a significant potential to advance robot coordination in the near future.

The papers presented at this workshop were not published.

The cochairs of this workshop were M. Bernardine Dias (Carnegie Mellon University), Sven Koenig (University of Southern California), and Michail G. Lagoudakis (Technical University of Crete).

Cognitive Modeling and Agent-Based Social Simulation

Traditionally, artificial intelligence and cognitive science have focused on individual cognition at the expense of the sociocultural processes and their relationship to cognition. Over the last few years we have seen growing interest in the use of multiagent systems to address issues of social interaction (for example coordination and cooperation) among cognitive agents. Most multiagent system, however, assume unrealistic (for example, extremely cooperative) environments and do not model the cognitive ten-



The Workshops Were Held at the World Trade Center in Boston, Massachusetts.

dencies that people are known to have or model the complex dynamics of the real-world environments.

Social scientists who are interested in building models of real-world social environments have traditionally limited themselves to verbal or static game-theoretic equilibrium models, which force them to make unrealistic assumptions such as homogeneity among agents and limit their analysis to small populations consisting of very few agents. However, recently, scientists have turned to building agent-based social simulation (ABSS) because it allows them to build more accurate dynamic models from the bottom up. By showing how complex social patterns can emerge through simple interactions of agents with little or no internal structure, ABSS models allow exploration, validation, and refinement of social theories and untangling of cause and effect relationships between individual characteristics and social phenomena—something that has been a hard to achieve, though much desired, objective in much of theoretical social science.

Existing ABSS models are limited, however, in their ability to model human social patterns because of their inability to represent complex social knowledge and reasoning processes. This is unfortunate because it severely limits the relevance of these models to the real-world social phenomena of interest. There are two natural ways of extending the collaboration between the ABSS and AI and cognitive modeling communities: first, by incrementally enhancing the knowledge representation and reasoning capabilities of existing ABSS systems, and second by adding social interaction capabilities to cognitive modeling and AI systems to build multiagent versions of these systems. The six papers presented at the half-day-long workshop suggested advances along both lines.

Gal Kaminka and Natalie Fridman from Bar Ilan University presented a knowledge-rich agent-based social simulation model based on Festinger's social comparison theory to model the behavior of pedestrians walking in groups. Afzal Upal from the University of Toledo presented a multiagent society of comprehension-based agents to model the emergence of misinformation in human societies. Gal Kaminka and his students also presented a teamwork architecture built on top of the cognitive modeling system Soar to model agent interaction during task performance that is unac-

counted for by traditional team architectures. Alexie Samsonovich from George Mason University presented a new hybrid symbolic connectionist architecture that allows agents with rich knowledge representation capabilities and allows agents to interact with one another to model human behavior. Teresa Ko and Justin Basilico from Sandia National Labs proposed an integration of Sandia's ABSS tool called Seldon with Sandia's Cognitive Framework to model the idea propagation in human societies. Hong Jiang and Jose Vidal from the University of South Carolina proposed an extension of the BDI architecture to account for emotions.

This very first workshop had some success in bringing together AI, cognitive modeling, and social simulation researchers, but more events of this kind may be needed to promote a sustained dialogue and interaction between such diverse research communities.

The papers presented at this workshop were published as an AAAI technical report (WS-06-02), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Afzal Upal (University of Toledo) and Ron Sun (Rensselaer Polytechnic Institute).

Cognitive Robotics

This workshop was the fifth in a series that started in 1998 in conjunction with the AAAI Fall Symposium in Orlando, Florida. While research in robotics has traditionally emphasized low-level sensing and control tasks, the workshop on cognitive robotics has been concerned with the problem of endowing robots and software agents with higher-level cognitive functions that enable them to reason, act, and perceive in changing, incompletely known, and unpredictable environments. Such robots must, for example, be able to reason about goals, actions, when to perceive and what to look for, the cognitive states of other agents, time, collaborative task execution, and so on. In short, the topic of the workshop was the integration of reasoning, perception, and action with a uniform theoretical and implementation framework.

The use of both software robots (softbots) and robotic artifacts in everyday life is on the upswing, and we are seeing increasingly more examples of their use in society with commercial products around the corner and some already on the market. As interaction with humans increases, so does the demand for sophisticated robotic capabilities associated with deliberation and high-level cognitive functions. The research results that were presented and discussed at the workshop show that combining insights from the traditional robotics discipline with those from AI and cognitive science has been and will continue to be central to research in cognitive robotics. The workshop brought together researchers involved in all aspects of the theory and implementation of cognitive robots. For 1.5 days participants discussed recent results, current work, and future trends and directions. A total of 23 technical papers were presented on a variety of topics related to cognitive robotics, including RoboCup, manipulation, service robots, action / plan representation, action recognition, architectures, and applications. A highlight of the workshop was the invited talk by Rachid Alami, who introduced the participants to COGNIRON (the cognitive robot companion), a collaborative research funded by the European Commission in the framework of the "Beyond Robotics" work program. A further highlight was the panel "Why My Cognitive Architecture Is on the Road to Nirvana," where different paradigms to cognitive robotics were presented, followed by lively and controversial discussion.

The next workshop in this series is planned for 2008 in conjunction with the European Conference on AI.

The papers presented at this workshop were published as an AAAI technical report (WS-06-03), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Michael Beetz (Universität München), Kanna Rajan (Monterey Bay Aquarium Research Institute), Michael Thielscher (Dresden University of Technology) and Radu Bogdan Rusu (Technische Universität München). This report was written by Michael Thielscher.

Computational Aesthetics: Artificial Intelligence Approaches to Beauty and Happiness

Our aesthetic agency for beauty and emotion is one of the most celebrated bastions of humanity. If machines could understand and affect our perceptions of beauty and happiness, they could touch people's lives in fantastic new ways.

Shooting for the moon, this workshop brought together researchers who—despite being "officially" registered under diverse fields such as cognitive science, art theory, psychology, philosophy, or natural language processing—are all working quite directly on the modeling and manipulation of people's perceptions of beauty and happiness. The collective attitude of the convened researchers struck us as being extremely ambitious yet highly pragmatic. Committed to "drinking the kool-aid," the postworkshop dinner venue was chosen by a computer program that collected mood keywords from participants and psychoanalyzed the latency of everyone's cravings.

Computational aesthetics had up to now been a suitcase term, used in the generative art community to mean mathematical aesthetics and grammars but used in the semantics community to refer to poetics and taste-based research. Building on previous sympathetic events—such as the FLAIRS special track on AI in music and arts, the 2005 Eurographics computational aesthetics workshop, and the 2004 AAAI spring symposium on style and meaning—this year's workshop on computational aesthetics hoped to integrate these different traditions and to articulate a clear mission statement for this emerging field of research by reiterating everyone's shared top-level goal—to understand and affect people's sense of beauty and happiness.

The workshop struck an equal balance between theoretical and empirically driven frameworks for understanding the aesthetics of domains such as chess, gambling, algorithmic music, interactive drama, generative art, poetry, interactive fiction, storytelling, humor, and typography. Poster presentations showcased applications of aesthetic technology for affect-based story rewriting, mood-based music and cocktail recommenders, laughter detection, and automatic dream analysis.

Perhaps most pleasing was the unplanned emergence of methodologies central to modeling and generating aesthetic perceptions—machine learning for corpus-based feature discovery, linguistic and commonsense resources to bridge a person's affective context with the environment, and analogy-based reasoning for generation of narrative variations. There was not a shortage of plainspoken take-away messages either. Aesthetic reactions such as the sense of novelty and the genre of suspense are often the result of expectation violation and may be governed by underlying "sweet-spot" equilibriums. Much of the textual aesthetics lies not in the story but in the order and manner of telling—from the selection of story details, to the metaphors invoked, to the sensibilities represented in word choice.

We hope to continue this most successful and satisfying workshop by making this the first in what is to be a series of such events to support the emerging field of computational aesthetics.

The papers presented at this workshop were published as an AAAI technical report (WS-06-04), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Hugo Liu (Massachusetts Institute of Technology), and Rada Mihalcea (University of North Texas).

Educational Data Mining

This year's workshop on educational data mining is the second such workshop held at AAAI. The field of educational data mining is a new one and focuses on applying and developing data mining and data analysis techniques to large data sets, with a focus on answering questions about how people learn and represent the domain being taught and how they respond to different types of instruction. Typically these data are available through computerized educational software, although some researchers work with data from standardized or classroom-administered tests.

The workshop covered a wide variety of topics. Several papers were presented on predicting high-stakes state achievement tests for mathematics. In addition to providing an interesting metric for evaluating success (does the proposed work improve our ability to predict student performance), this work introduces a new consumer of the research. In the past, researchers analyzing data from educational software were typically trying to obtain results that would improve the system itself by changing how it taught students. This work highlights a new type of consumer for our research efforts: the educational practitioner.

Several other pieces of work presented at the workshop focused on these new consumers. One paper presented an approach to automatically analyze the content of a student's online discussion posts to determine whether the student was an active participant. Other work discussed using data from prior students to help advise students on which courses to register for. One area for future work is determining how a program can act on its knowledge. In the past, new knowledge could be directly used to improve the computer tutor from which the data were gathered. How confident should a program be before it alerts an instructor that a student is not participating? How many interruptions will an instructor tolerate? Can a system do anything if a student ignores its course recommendations?

In addition to work on emerging areas, there was also work on unifying two separate approaches from the prior workshop on determining the domain skills that underlie a set of test questions. Two competing approaches are to use student performance data to automatically derive the topics and to use experts to determine what aspects of the domain each question tests. These approaches yield rather different models of the domain, with the automatically generated model typically being more compact. The paper presenting this work contrasted the approaches to better understand the pros and cons of each.

One other theme at the workshop was the usefulness of having members from outside of the AI community participating in the workshop. AI researchers understandably focus on selecting the right modeling approach for the task and determining how to represent the task in a known framework. For solving practical problems in education it is also crucial to consider what is actually causing the variance in outcomes. For example, thinking about issues such as the impact of a student's teacher and school district may not come naturally to an AI researcher but is something an educator or psychologist would naturally think of. Accounting for such characteristics is likely to be more important for better predictive accuracy than the exact model form chosen. Similarly, statisticians and psychometricians have spent much effort thinking about issues relevant to our community such as modeling latent student traits and handling nonindependence of observations. Maintaining and extending this cross-disciplinary nature of the community is essential for its continued success, and we look forward to working with additional members of neighboring research communities.

The papers presented at this workshop were published as an AAAI technical report (WS-06-05), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Joseph E. Beck (Carnegie Mellon University), Esma Aimeur (University of Montreal), and Tiffany Barnes (University of North Carolina at Charlotte).

Evaluation Methods for Machine Learning

The call for papers for this workshop stated that one of its goals would be

"to encourage debate within the machine-learning community into how we experimentally evaluate new algorithms." We began that process through a lively and interesting debate involving all attendees. The discussion was stimulated by eight interesting papers and three very good invited talks. The workshop organizers began by pointing out the many problems with the current evaluation process. A common theme in subsequent presentations was the large variety of measures that could, and should, be used to evaluate algorithms. Some from the medical community not only measured important properties of classifiers but also were more acceptable to users in that community. Others were used to assess calibration, for model selection and to measure the explanatory power of Bayesian networks. The range of possible measures was exemplified in Richard Carauna's (Cornell University) invited talk. He discussed a very large-scale experimental study using many different measures on many different algorithms. He also used bootstrapping to analyze the interaction between measures based on entries to the KDD cup. He concluded there was a larger variability in the best algorithm between problems rather than between measures.

Robert C. Holte (University of Alberta), in another invited talk, criticized scalar measures, saying that the total order they imposed on classifiers was misleading. He argued that it was critical to determine not whether one algorithm is better than another but when it is better. He presented cost curves as a way of easily visualizing this. An alternative form of visualization was also presented. In another invited talk, Dragos Margineantu (Boeing) argued for the importance of good confidence intervals for measures and discussed using bootstrapping for bounding expected cost. He showed that very costly and infrequent faults have a large impact on confidence intervals. Understanding the complete distribution for different algorithms and different measures would lead to tighter intervals. Holte also stressed the advantages of confidence intervals but showed that they gave different conclusions dependent on external factors such as misclassification costs and class priors.

The workshop did produce a general consensus that the current means of evaluating learning algorithms has some serious drawbacks. We agreed that there are many important properties of algorithms that should be measured, some application specific. Current practice, which reports the best single classifier based on a single measure, is of questionable merit. To evaluate an algorithm, it is necessary to know how it will perform according to many different measures and under many different conditions. We also agreed that current practice is overreliant on the UCI data sets as a way to estimate general performance. These data sets do not reliably represent the diverse problems we encounter in practice. Simulated data would give better assurance of generalization, by exploring the larger space of possible variations seen in practice. Building the simulation around some existing data sets, points in the much larger data space, would prevent some possible criticisms of using simulated data.

The papers presented at this workshop were published as an AAAI technical report (WS-06-06), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Chris Drummond (NRS Institute for Information Technology), William Elazmeh (University of Ottawa), and Nathalie Japkowicz (University of Ottawa).

Event Extraction and Synthesis

The AAAI 2006 Workshop on Event Extraction and Synthesis was the first workshop to focus specifically on the problem of event extraction, primarily from text, rather than the more general problem of information extraction. Given its groundbreaking status, the workshop was intended to provide an opportunity for discussion of a wide variety of pertinent topics, ranging from fundamental notions, such as the definition of event extraction, to shared concerns, such as the complexity of the event extraction problem, to the presentation of specific extraction problems and techniques.

Event extraction is multifaceted, and different techniques appear best applicable for different aspects of the problem. Thus, the workshop drew researchers from many areas, such as information retrieval and extraction, natural language processing, semantics and knowledge representation, information modeling, and data management.

At the beginning of the workshop, the organizers posed some questions to the participants, to serve as "food for thought" during the day: What is the scope of event extraction? Is event extraction one large (possibly unwieldy) problem, or can it be decomposed into more tractable subproblems? Should we worry about the speed of event extraction? Are there applications of event extraction where speed is important? If so, what separates acceptable from unacceptable speeds? Is event extraction a problem that can be cleanly divorced from the particular domain or application? Or do domain-specific considerations tend to dominate in typical event extraction problems?

The intention was not to solicit quick answers to any of these imposing questions, but rather to initiate discussion, in the hope that the resulting appreciation of such issues could serve to lend clarity to our understanding of the challenge of event extraction.

Nine papers were presented at the workshop. A few papers focused on the problem of identifying, in a large document, which sections are relevant to a particular event, as a precursor to the extraction of event slot values. Much of this work sought to find the features that would enable a classifier to make the relevance determination.

Syntactic features (for example, indicative noun and verb phrases) and features based on document structure (such as sentence position within a document) were among those explored by presenters (Hilda Hardy, Vika Kanchakouskaya, and Tomek Strzalkowski and M. Naughton, N. Kushmerick, and J. Carthy).

There was also a presentation on a bootstrapping approach where identi-

fying patterns can be learned starting with a few seed examples.

Of course, the filling of event slots was the focus of several papers. Here, the interesting question appears to be where to fall on the spectrum between blank-slate machine learning and full natural language understanding. For example, Nancy McCracken, Necati Ercan Ozgencil, and Svetlana Symonenko presented a statistical approach augmented with grammar rules. On the other hand, Charles J. Fillmore, Srini Narayanan, and Collin F. Baker maintained that linguistic analysis using existing lexical resources, particularly FrameNet, is sufficient by itself to solve the typical problem of event extraction from text.

The workshop discussions extended to issues beyond slot filling, as well, including a paper presentation on the challenge of information consistency. Earl J. Wagner, Jiahui Liu, Larry Birnbaum, Kenneth D. Forbus, and James Baker highlighted the problem of multiple, possibly inconsistent values of an event slot being extracted from multiple sources (for example, multiple articles in a corpus of news stories), and presented a voting-based approach to choose a most likely single value from among the alternatives. This presentation also motivated discussion of issues related to the quality of the extracted information, its reliability and the reliability of its source, confidence measures, and so on.

One of the fundamental issues, that of event modeling, was discussed in the context of a paper on (multimedia) event clustering where the authors presented an event model as a data model capturing for event data as well as event relationships. Finally, there was a presentation on an end-to-end event extraction and analysis system with use cases in the business intelligence domain.

Ralph Weischedel of BBN gave the invited talk, in which he traced the history of and summarized experiences with the DARPA ACE program. A key challenge to the designers of the ACE event extraction initiative was defining the problem with sufficient clarity that adequate interannotator agreement could be achieved. Ultimately, after several iterations, the committee came to the realization that the problem of event extraction was too subtle to allow the same level of interannotator agreement (and, by extension, machine performance) as that of, say, named entity extraction. Weischedel shared several of the ambiguous sentences from this experience, leaving workshop participants with a heightened appreciation of the difficulty of the task. Weischedel also discussed recent efforts to bring more semantics to the extraction task, such as the OntoNotes project, arguing that the construction of such knowledge resources is critical. One of the things that the talk stressed was the need to contextualize any given extraction problem within its end goal, such as database or knowledge-base creation.

In summary, the workshop provided a fruitful day of high-quality research presentations and discussions among researchers in diverse areas related to extraction.

The papers presented at this workshop were published as an AAAI technical report (WS-06-07), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Naveen Ashish (University of California, Irvine), Doug Appelt (SRI International), Dayne Freitag (SRA), Fair Isaac (SRA), and Dmitry Zelenko (SRA).

Heuristic Search, Memory-Based Heuristics, and Their Applications

The AAAI-06 Workshop on Heuristic Search, Memory-Based Heuristics, and Their Applications took place on July 17, 2006, during the events of AAAI-06 in the World Trade Center in Boston. Heuristic search is one of the earliest, great ideas in AI, beginning with J. Doran and D. Michie's work on Graph Traverser and P. E. Hart, N. J. Nilsson, and B. Raphael's work on A*. Despite extensive study for four decades and attainment of a significant level of maturity, there is still much to be learned. Indeed, the past decade has seen a surge of research activity on heuristic search, with significant advances such as J. C. Culberson and J. Schaeffer's work on pattern databases and methods devised by R. E. Korf, R. Zhou, and E. Hansen for using disk space effectively during search.

In addition, there has recently been a large expansion of the application of heuristic search in other fields of artificial intelligence and computer science such as planning, model checking, dynamic programming, and weighted logical inference.

The workshop had two aims. The first was to bring together the members of this rapidly growing research community so they could meet one another, exchange ideas, discuss methods for evaluating heuristic search systems, and share a vision of the future. The second aim was to present, in a single venue, all the recent advances in heuristic search and their applications. The workshop was successful in both regards. It was widely attended, by senior researchers in the field as well as by new students who are only taking their first steps.

The workshop began with a comprehensive survey of recent advances by Rich Korf. Presentations were then made on a wide variety of topics, ranging from high-performance parallel implementations to the latest advances in AND/OR search.

An important contribution to the workshop was Chris Raphael's tutorial on coarse-to-fine dynamic programming, which provided evidence of the wide applicability of heuristic search. The workshop ended with a general discussion about various aspects of the field, including implementation techniques, domain independence, and methods for fairly comparing different heuristic search systems.

Finally, it was decided that the workshop should be held again, and David Furcy and Rong Zhou volunteered to organize it.

The papers presented at this workshop were published as an AAAI technical report (WS-06-08), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Ariel Felner (Ben-Gurion University of the Negev), Robert C. Holte (University of Alberta), and Hector Geffner (UPF).

Human Implications of Human-Robot Interaction

Autonomous humanoid robots have begun to display levels of humanlike behavior and appearance that call for attention to important psychological, sociological, philosophical, and spiritual implications of human-robot interaction (HRI). As ongoing commercial development of these robots makes human-robot interaction increasingly common, elements of human culture outside the technical communities of AI and robotics need to engage HRI phenomena with an aim to seek clearer understanding of their potentially significant long-term effects upon human individuals and human society. Moreover, artificial intelligence itself, now generally considered to be celebrating its 50th year, is achieving capabilities for mimicking human behavior that oblige its own technical community to examine reflexively the effects of HRI upon other components of human culture.

Accordingly, this HRI workshop deliberately invited papers from an especially multidisciplinary and international academic population, with an intent to cultivate dialogue of a scope and quality that would improve awareness and understanding of specifically human implications of human-robot interaction. Its call for papers suggested particular topics of interest that included potentially HRI-related alterations in traditional concepts of human identity, uniqueness, consciousness, freedom, moral status, and moral responsibility.

A distinctive extra feature of this full-day workshop, also motivated by its international scope, involved use of videoconferencing technology. Initial presentation of papers was followed by open international discussion during an afternoon session. Both sessions were linked in real time by means of multipoint videoconferencing to additional gatherings of invited scholars at the Universität Augsburg, in Augsburg, Germany, and at Oklahoma City University, in Oklahoma City, Oklahoma. Highlights of the dialogue generated within this workshop already indicate that it successfully accomplished its principal objective of improving awareness and understanding of specifically human implications of HRI. In broadest terms, its presentations and its discussions helped to clarify important distinctions (and interrelations) between the following types of questions: (1) How are people actually responding to HRI experience? (2) What kinds of responses arguably are warranted? (3) What kinds of responses should we accept?

Several notable observations generated discussion regarding question 1. First, the actual kinds of responses that people observably display in HRI appear to be influenced by a variety of factors such as types of robots involved, gender of the human involved, and various cultural features (such as religious beliefs) that might be represented. Second, a case was made for the hypothesis that ontological categories to which people assign the robots can evolve over time and may even overlap. Third, humans abundantly have demonstrated dispositions to bond with sufficiently humanlike robotic artifacts—a phenomenon highlighted especially in one of our presentations that included video of human interaction with MIT's social robot, Kismet.

Regarding question 2, several of our papers and presentations reminded us of two important claims concerning human consciousness: (1) the authentic presence of subjective conscious awareness in other humans, and in robotic systems, apparently remains beyond the reach of scientific demonstration, but (2) behavioral cues that we normally associate with such conscious awareness can potently influence our practical attributions of consciousness to other entities. Not surprisingly, the workshop did not manage to determine by consensus whether—and in what sense—such attributions may be warranted.

Question 3—reflecting a separate concern with how we ought to respond in HRI—was addressed from a number of perspectives. First, interaction with question (2) invited some uncomfortable questions involving the notion of being deceived. For example, should we accept the introduction of so-called "care giving" mechanical artifacts? Again, if one happens not to accept "compatibilist" freedom as the sort of freedom required for moral responsibility, should one accept mechanical artifacts as moral peers? Finally, one of our presenters—legitimately representing the technical robotics community itself—drew attention to inappropriate and misleading anthropomorphic language that is used (arguably, too often) in public descriptions of existing robotic systems.

Overall, perhaps the most significant point that surfaced repeatedly in this workshop was the observation that autonomous humanoid robots functionally can constitute "mirrors" in which humans may discern new images of themselves.

The papers presented at this workshop were published as an AAAI technical report (WS-06-09), which is now available for sale in hardcopy and by download from the AAAI digital library.

The chair of this workshop was Ted Metzler (Wimberly School of Religion). The cochairs were Lundy Lewis (Southern New Hampshire University) and Wolfgang Achtner (Justus-Liebig-Universität Giessen).

Intelligent Techniques in Web Personalization

The workshop on intelligent techniques in web personalization, the fourth in the series, continued it focus on the application of AI to personalizing user interactions with the web. To achieve effective personalization, a variety of types of data must be harnessed, including user profiles, web content and structure, and domain knowledge. Efficient and intelligent techniques are needed to mine this data for actionable knowledge and to effectively use the discovered knowledge to create user models. These techniques must address important challenges emanating from the size and the heterogeneous nature of the data itself, as well as the dynamic nature of user interactions with the web.

The workshop has traditionally attracted researchers from the fields of web mining, user modeling, adaptive hypermedia, semantic web, intelligent agents, and distributed AI working to-

wards the common goal of providing the users of the web with an interface that provides content and services that are relevant to their current needs without their necessarily having to request it explicitly.

A number of themes emerged within this year's workshop. The first of these was the need to develop techniques for identifying malicious attacks on recommender systems. Recommender systems, being essentially open systems that take input from all users of the system and make recommendations based on the data input, have been shown to be sensitive to attacks that either promote or nuke products. Previous work has proposed attack models that require varying degrees of insight into the distribution of user ratings stored within the recommender system. Runa Bhaumik, Chad Williams, Bamshad Mobasher, and Robin Burke (DePaul University) presented a paper at the workshop on using anomaly detection to detect attacks on recommender systems.

Another theme was the discovery of latent factors implicit within the user rating matrix. Papers presented by Panagiotis Symeonidis, Alexandros Nanopoulos, Apostolos Papadopoulos, and Yannis Manolopoulos (Aristotle University) and Bhaskar Mehta (Fraunhofer IPSI), Thomas Hofmann (Darmstadt University), and Peter Fankhauser (Fraunhofer IPSI) focused on the use of latent factors for improving scalability of the recommender systems and cross-domain recommendation, respectively.

Another recurring theme at the workshop has been that of using item knowledge bases within the recommendation process. There has been an increasing interest in developing hybrid models for recommendation generation and more recently into the use of deeper domain knowledge in the form of ontologies to improve recommendation quality. Santtu Toivonen (VTT Technical Research Centre of Finland) and Oriana Riva (Helsinki Institute for Information Technology) presented a paper that revisited the role of ontologies within content filtering applications.

Seung-Taek Park, David M. Pennock, and Dennis DeCoste (Yahoo! Research) and Rohini Uppuluri and Vamshi Ambati (International Institute of Information Technology) presented papers on reranking search results from Yahoo's Movie Search database and the Internet at large, respectively.

Finally, the plenary talk focused on the role of context within recommendation generation and user profiling. Sarabjot Singh Anand noted that the role of context has been researched extensively in cognitive science and context-aware computing and more recently in ubiquitous computing and information retrieval; however, it has been largely ignored in personalization. He suggested the use of context as a cue for retrieving those user ratings stored in long-term memory that are relevant to the current context and basing recommendation generation solely on those ratings. Results presented during the talk suggested that this strategy can improve recommendation quality.

The workshop closed with a discussion among the participants regarding the future directions in intelligent techniques for web personalization. The participants saw evaluation of personalization systems as a continuing challenge with most systems to date focusing on predictive accuracy rather than on aspects such as recommendation diversity, novelty, and serendipity. The need for more publicly available data sets and user studies was also expressed. As far as future directions are concerned, the following areas were noted as being key foci for the community: modeling and predicting context, extending recommender systems to be useful across multiple domains, developing techniques for attack prevention, and distributed recommender systems.

The papers of the workshop were published as AAAI Technical Report WS-06-10, and are available from AAAI Press.

The cochairs of this workshop were Bamshad Mobasher (DePaul University) and Sarabjot Singh Anand (University of Warwick).

Learning for Search

Heuristic search is among the most widely used techniques in AI. In its different varieties, tree-based search and local search, it provides the core engine for applications as diverse as planning, parsing, and protein folding. One of the most promising avenues for developing improved search techniques is to integrate learning components that can adaptively guide the search. Research in this field is of wide interest in the AI community, not only because of the variety of subcommunities directly involved (problem solving, learning, constraint programming, operations research) but also because of the wide range of applications areas in which search algorithms play an important role.

Many disparate techniques have arisen in recent years that exploit learning to improve search and problem solving. These techniques can be offline or online, based on hard constraints or probabilistic biases, and applied to tree-structured or local search. This workshop brought together researchers and practitioners from the various subcommunities where such methods have arisen in order to learn from each other, develop common understandings, and inspire new algorithms and approaches.

Note that we are not discussing how search can be used to improve a learning algorithm, but rather the other way around. Accepted papers discussed how the learning component can be employed to increase problem-solving performance, not just aid in scientific understanding.

This topic area is of particular interest at this time for two reasons. First, search itself is increasingly important. Most modern AI planning systems since the late 90s rely heavily on heuristic search, for instance, and search is also crucial in bioinformatics. The recent publication of Hoos and Stutzle's textbook and the continued success of the SAT conferences are further evidence. Second, machine learning has matured to the point where efficient methods are widely available for handling the large amounts of data generated by search processes. Knowledge of how to apply these methods is also more widespread. Many isolated attempts to apply learning to improving search performance

have been tried, dating back to the original founding of AI (GPS and SOAR, for example). However, to our knowledge there has been no recent venue where those interested in the topic could gather together to share ideas and experience.

Recent successful meetings that have addressed search include workshops at IJCAI, AAAI, and ICAPS. While many of the meetings have included elements that relate to learning for search from a particular perspective, none of them has focused exclusively on the topic, bringing together a wide range of perspectives at once.

We were pleasantly surprised by the number of submissions received (24). We have contributions from the constraints community, from planning, on real-time search, and on learning heuristics, on local search, and on search trees. Each one was carefully reviewed by at least two members of the program committee. We accepted 9 papers for oral presentations, 3 for each of the sessions clause weighting/constraint learning, run-time prediction/learning of heuristics, and real-time search. We also accepted an additional 12 papers for the afternoon poster session. Both categories of presentations are represented by full papers in the workshop technical report, which has been posted in AAAI's digital library. These papers truly represent a broad cross-section of the state of the art on learning for search. Due to the success of the workshop, plans are afoot to propose a similar session for AAAI-07, in coordination with participants of the 2006 workshop.

The papers of the workshop were published as AAAI Technical Report WS-06-11 and are available from AAAI Press.

The cochairs of this workshop were Wheeler Ruml (Palo Alto Research Center) and Frank Hutter (University of British Columbia).

Modeling and Retrieval of Context

Context and context-awareness are crucial not only for mobile and ubiquitous computing but for numerous application areas such as information sharing, workflow, health care, personal digital assistants, and e-learning. Contextual knowledge, including both application-specific and environmental knowledge, can serve as a major source for reasoning, decision making, and adaptation. Exploiting such knowledge depends on models, methods, and tools to enable structured storage of contextual information, to provide effective ways to retrieve it, and to enable integration of context and application knowledge.

The two-day AAAI 2006 Workshop on Modeling and Retrieval of Context (MRC 2006) was the third in a series of international MRC workshops. Seven papers were presented, accompanied by four posters and six application demonstrations, but the primary workshop emphasis was on discussion rather than presentations.

A highlight of the first day was Anind Dey's invited talk on "Usability in Context-Aware Applications," a broad overview of the implications of usability issues of control, feedback, privacy, and information overload for context modeling, based on his work at the Human Computer Interaction Institute at Carnegie Mellon University. Following a tradition in the MRC series, on the first day the participants selected two topics to be explored individually by discussion groups: "Evaluation of context elicitation and context models" (lead by Sven Schwarz and David Vallet) and "Context representation and context schema requirements" (lead by Philipp Mohr). Following another tradition, after the first day's sessions the participants gathered for a group dinner and more discussion, this time at the picnic tables of the Barking Crab.

The second day featured the demo session, which started with brief presentations to provide broad overviews of systems and continued with individual hands-on trials, developing contacts for collaborations beyond the workshop. The systems ranged from fully functional applications (such as the "Onlife" system for desktop context inference, by Edison Thomaz) to prototypes and work in progress; their application domains ranged from proactive knowledge management and context-aware information retrieval to mobile systems. The closing afternoon included intense group discussions of the topics selected the first day, followed by each group's presentation of its preliminary results. The topic "Context representation and context schema requirements" proved sufficiently rich that the group planned to continue its work and report later results. Participants expressed strong interest in continuing the workshop's dialogue in a future meeting, and options are being explored for a followup to the workshop in 2007.

The papers presented at this workshop were published as an AAAI technical report (WS-06-12), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Thomas Roth-Berghofer (TU Kaiserslautern), David Leake (Indiana University), Stefan Schulz (The e-Spirit Company GmbH), and Sven Schwarz (DFKI GmbH).

Modeling Others from Observations

The Modeling Others from Observations (MOO) workshop was the latest of a series of successful workshops designed to bring together researchers working in a number of different research communities all related to modeling the actions, behavior, and goals of human and synthetic agents. Since the work in this area is done under a number of different research headings, the sharing of research results has suffered from compartmentalization. We felt that a wider sharing of results from varying disciplines would help all researchers working in this area and therefore have organized a number of different MOO workshops over the last few years at different conferences including AAMAS, IJCAI, and this year at AAAI.

To achieve the goal of sharing research results more broadly we have attempted to cast the net as wide as possible to bring together researchers with common underlying objectives but very differing domains and techniques. As a testament to our success, MOO had papers covering a wide spectrum of research areas from more tra-

ditional plan recognition to activity recognition based on an RFID tagged environment to social network analysis from annotated photo sharing to policy learning.

Even given the diversity of the papers, a number of themes emerged during the workshop, the first being the application of these recognition and modeling techniques as just one component of larger systems for applied contexts. For example, there has been a recent surge in interest in the AI community as a whole on assistive systems for the disabled, and this interest was certainly evident at MOO with a number of papers reporting on applications to various portions of this domain. However, applications to other "hot topics" including social network analysis and even RoboCup were also reported on.

It was noted that this use of these technologies within larger applied systems may be contributing to diluting the impact of research results, since in some cases the recognition and modeling techniques themselves are not the focus of these applications. There was also significant discussion of the possibilities for the application of these technologies to domains where they are less common including computer network security and critical infrastructure protection.

The continuing resurgence of probabilistic methods in AI was also evident. A majority of the papers had at least a hybrid probabilistic approach if not being fully probabilistic. This stands in sharp contrast to much of the early rule-based work in plan recognition and has gone a log way to eliminating some of the weaknesses of very early work in this area.

A number of significant research questions were also identified that led to productive conversations after the workshop, extending well into the rest of AAAI and beyond. Further, since one of the reasons for this workshop is to attempt to share results from a wider range of researchers and research areas, we have begun discussions of how we could encourage still wider participation in future workshops. This discussion has resulted in a number of good suggestions about ways to raise the visibility of this work and even other pools of researchers that should be made aware of the work in this area. As a result we are very optimistic about the significant advances reported at MOO, the dissemination of these results, and even the application of these new results to new exciting domains.

The papers presented at this workshop were published as an AAAI technical report (WS-06-13), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Gal A. Kaminka (Bar Ilan University), David V. Pynadath (ISI/University of Southern California), and Christopher W. Geib (Honeywell Labs).

Statistical and Empirical Approaches for Spoken Dialogue Systems

Spoken dialogue systems help users accomplish a task using spoken language—for example, booking a flight, reserving a conference room, or obtaining restaurant information. Traditionally these systems have been designed by hand, but as the limitations of handcrafting have become apparent, researchers have begun approaching aspects of dialogue design as an optimization problem. This workshop focused on new work in this area. Each paper was reviewed by 3 members of the 31-member technical committee, and the workshop accepted eight papers.

Workshop researchers presented advances addressing some of the key challenges faced by spoken dialogue systems. First, speech-recognition errors are common, and as a result a computer can never know the true state of the conversation with certainty. To address this, workshop researchers presented new techniques for maintaining multiple hypotheses of the conversation state through partially observable models. Several papers also explored the related problem of how to exploit this state to perform planning and guide system behavior. Second, running trials between users and spoken dialogue systems is expensive, and as a result the field suffers from a chronic shortage of data. To address this, workshop researchers presented methods for faithfully modeling users, with the aim of assessing and improving dialog systems with little or no human contact. Third, there is a common problem of "new situations" unseen in training data, and workshop researchers presented new methods to cope with unseen dialog situations, to track shifting user behavior and needs, and to generate novel output language.

The organizers see spoken dialogue systems as a catalyst for collaboration between the language technology community and the AI community. Whereas the language technology community benefits from a deep understanding of the structure of conversation, language phenomena, and the workings of language technologies such as speech recognition, the AI community enjoys a long history and deep "toolkit" of core algorithms for planning, control, and utility maximization. Moreover, the AI community increasingly seeks real-world applications to test algorithmic advances. Toward this end, Pascal Poupart's keynote address reviewed recent techniques from the AI literature and interpreted their potential for the spoken dialogue systems domain. Both communities were represented in the workshop, leading to a lively discussion session, which identified a joint initiative capable of engaging both communities more broadly.

The workshop organizers would like to extend their thanks to the technical committee for their thoughtful paper reviews and to all of the participants for joining.

The papers presented at this workshop were published as an AAAI technical report (WS-06-13), which is now available for sale in hardcopy and by download from the AAAI digital library.

The cochairs of this workshop were Pascal Poupart (University of Waterloo), Stephanie Seneff (Massachusetts Institute of Technology), Jason Williams (AT&T Labs-Research), and Steve Young (Cambridge University).

.