2002 AAAI Spring Symposium Series

Jussi Karlgren, Pentti Kanerva, Björn Gambäck, Kenneth D. Forbus, Kagan Tumer, Peter Stone, Kai Goebel, Gaurav S. Sukhatme, Tucker Balch, Bernd Fischer, Doug Smith, Sanda Harabagiu, Vinay Chaudri, Mike Barley, Hans Guesgen, Thomas Stahovich, Randall Davis, and James Landay

The American Association for Artificial Intelligence, in cooperation with Stanford University's Department of Computer Science, presented the 2002 Spring Symposium Series, held Monday through Wednesday, 25 to 27 March 2002, at Stanford University. The nine symposia were:

- Acquiring (and Using) Linguistic (and World) Knowledge for Information Access
- Artificial Intelligence and Interactive Entertainment
- Collaborative Learning Agents
- Information Refinement and Revision for Decision Making: Modeling for Diagnostics, Prognostics, and Prediction
- Intelligent Distributed and Embedded Systems
- Logic-Based Program Synthesis: State of the Art and Future Trends
- Mining Answers from Texts and Knowledge Bases
- Safe Learning Agents
- Sketch Understanding

■ The American Association for Artificial Intelligence, in cooperation with Stanford University's Department of Computer Science, presented the 2002 Spring Symposium Series, held Monday through Wednesday, 25 to 27 March 2002, at Stanford University. The nine symposia were entitled (1) Acquiring (and Using) Linguistic (and World) Knowledge for Information Access; (2) Artificial Intelligence and Interactive Entertainment; (3) Collaborative Learning Agents; (4) Information Refinement and Revision for Decision Making: Modeling for Diagnostics, Prognostics, and Prediction; (5) Intelligent Distributed and Embedded Systems; (6) Logic-Based Program Synthesis: State of the Art and Future Trends; (7) Mining Answers from Texts and Knowledge Bases; (8) Safe Learning Agents; and (9) Sketch Understanding.

Acquiring (and Using) Linguistic (and World) Knowledge for Information Access

The symposium entitled Acquiring (and Using) Linguistic (and World) Knowledge for Information Access gathered 20 some researchers and practitioners from corporations, research institutes, and academic institutions from far corners of the world.

The subtitle was "Theory for systems; application for theories," and the aim was to find common ground between those who work on defining algorithms and methods and building systems for the acquisition of knowledge from text and those who use knowledge for information-access–related tasks, primarily for understanding text and user queries. The presentations ranged from new theories for computational modeling of brain information processing and applications of various types of multivariate statistical models to text categorization to name identification across documents and ontology-based search systems.

The discussion centered on representational issues. Representation is central to projects aiming at modeling the human brain, important to projects aiming at information reuse, less important to those aiming for behavioral verisimilitude or computational efficiency, and ad hoc but preferably handy and inspectable for those aiming at solving some concrete task. The question of how to compare and unify representations was identified as the focus point for future work.

The hopeful aim of the symposium was to find a common task to test methods and representations systematically, in the spirit of the Text Retrieval Conferences organized annually by the U.S. National Institute of Standards and Technology; the consensus at the workshop was that this was too tall an order at this time and that further separate experimentation would be necessary for systems and frameworks to mature to the point where bridging the gap between knowledge acquisition and application would be possible and where common tasks would provide a fruitful research motivation. A tentative aim is for the participants to meet again in some two years, after performing new sets of experiments, to discuss the same or similar questions.

– Jussi Karlgren and Björn Gambäck,
Swedish Institute of Computer Science
– Pentti Kanerva
Stanford University

Artificial Intelligence and Interactive Entertainment

The 2002 AAAI Spring Symposium on AI and Interactive Entertainment is the fourth in a series of meetings that were started to build bridges between AI researchers and game designers. This year, there were around 45 attendees, evenly mixed

between AI researchers and game designers and developers. Such a diverse collection of viewpoints, goals, and approaches provides a stimulating atmosphere, where the consequences of technical ideas can be explored hand in hand with how to make more compelling artistic works.

Several threads emerged in the panels and papers that were presented. First, a number of technologies were explored for controlling characters, ranging from behavior-based techniques from robotics (for action-oriented games) to hierarchical task network planning (for interactive stories). A second theme was how to achieve better conversations. It is generally agreed that advances in conversation by nonplayer characters (NPCs) will lead to radically new kinds of interactive entertainment. Today's computer games are mostly aimed at young males and already produce more gross revenue than the movie industry; what if they appealed to both genders and to the same age range as those who watch movies and read novels? A number of efforts are thus focused on improving conversational skills, including the cleverly engineered ELIZAS (for example, the trash-talking 14-year-old moron and a genre of COUNTERSTRIKE bots that annoy players so much that they relish killing them) to carefully crafted natural language–processing systems designed with strong narrative theories (for example, the FAÇADE interactive story, whose demonstration video had participants squirming at the signs of a marriage breaking up before their very eyes) to the use of multimodal spoken dialogue to interact with agents (for example, the CSLI UAV work, which uses a game-like simulator as a test bed).

A number of special events marked the symposium. Doug Church, a game designer whose credits include THIEF, SYSTEM SHOCK 2, and DEUS EX, gave a superb invited talk on the first day. On the second day, a demonstration session enabled participants to show off their systems. Finally, a rousing debate on whether virtual humans in interactive entertainment should be viewed as participants or as actors was held.

This series of meetings has succeeded in its goals. Many productive collaborations have been started, and we hope in the next few years to see a steady stream of both solid research papers in standard AI publication venues and more interesting products in the marketplace as a consequence. Other groups are now engaged in similar efforts: This year was the first academic summit at the annual Computer Game Designer's Conference (the de rigueur conference for the industry), and the Computers and Games Conference (an academic venue) is now building bridges to industry. Because we have accomplished what we set out to do, the organizers decided that this symposium would be the last in this series. There will be other game-related symposium topics in the future to be sure, and there are discussions about creating a workshop to provide a more intense forum for interaction than conferences can provide. However, we are pleased to report that this symposium has done its job, and we are all eagerly working on the next stages of this exciting line of work.

– Kenneth D. Forbus
Northwestern University

Collaborative Learning Agents

Recent advances in the multiagent systems field have generated optimism that widely applicable solutions to large, distributed problems might be at hand. However, before the field can deliver on this promise, the challenge of how to control such systems to address a prespecified goal (for example, minimize throughput of packets in data routing, win the game in soccer) in a decentralized, adaptive manner with minimal detailed hand-tuning needs to be met.

In this symposium, we focused on two crucial properties that would allow a multiagent system to meet these challenges: (1) the agents need to work collectively so that as a group, their behavior solves the overall problem and (2) both the agents and their collaborative structure need to be adaptive.

The first property is crucial in large problems (for example, internet routing), and inherently distributed problems (for example, planetary exploration rovers, constellations of satellites) in that it enables a modular approach to the problem. The importance of the second property lies in how the agents interact with one another and the environment. Because both the environment and the response of other agents to changes in this environment will modify the "background" state one agent perceives before choosing its actions, it is imperative that adaptivity be built into these agents.

In the course of the symposium, several topics surfaced repeatedly, leading to spontaneous and active discussion sessions. For example, many of the talks used game-theoretic models and relied on the assumption that all agents are fully rational. Thus, we examined the scope and realism of this assumption. Other discussion topics included the trade-offs between increasing the number of agents in a system versus computational cost, the different approaches to multiagent system design in the presence of a well-defined global objective function, and the relative merits of model-free and model-based methods.

Two invited talks highlighted crucial aspects of collaborative learning: Michael Littman (Stowe Research) presented collaborative reinforcement learning approaches in a game-theoretic context, and David Wolpert (NASA Ames Research Center) discussed the behavior of large collections of agents striving to perform a well-defined global task.

– Kagan Tumer
NASA Ames Research Center
– Peter Stone
AT&T Labs–Research

Information Refinement and Revision for Decision Making: Modeling for Diagnostics, Prognostics, and Prediction

This symposium explored selected new issues within the diagnostics and prognostics arena. These issues arise as companies collect an increasingly

large amount of process data and business information. This collection is accelerated by the use of advanced and less expensive sensors, massive information storage, internet-facilitated access, and a thrust to preserve and maintain corporate knowledge. In addition, companies are interested in being informed about potential problems earlier with higher reliability and under tighter budgetary constraints. Diagnostic decision makers faced with the daunting task of extracting, refining, and using critical information from this hodgepodge must aggregate heterogeneous information, resolve conflicting information, remove or discount stale and outdated information, and evaluate the merits of alternate decisions.

At the core of the diagnostic problem lies the question about which technologies to choose. Regrettably, the answer is not always clear at the onset and should, in part, be determined by requirements such as the acceptable false positive and false negative rates, allowable computational burden, and the degree of transparency desired. In addition, the nature of the data and resources available (including expertise of the developer) can eliminate a number of approaches. The remaining techniques should undergo a formal down-select procedure for final selection.

Automated decision-making systems also need to deal with the increased complexity of systems that can be dealt with in either a coupled or decoupled fashion. The *decoupled approach* breaks down the diagnostic problem into a number of manageable subproblems. However, it potentially fails to properly capture interactions between the individual subdomains. In addition, system designs that depend on large arrays of distributed sensors and computing must address critical issues associated with security and privacy. The acceptability of using either approach depends on the system complexity, fault interactions, and the trade-offs between centralized and local control.

Temporal effects play a major role in the decision-making process, not only because information integrity fades over time but also because new information needs to be factored in. Although this new information does not exist at the time of the system design, a system maintenance plan must be developed to account for its inclusion during operation. Ways to judge the relevance of this new information and optimize its value in decision making needs to be addressed in this context. Finally, a methodology is needed to evaluate the trade-offs in the costs, uncertainties, and quality metrics in the diagnostic system and the resulting impact of the associated decisions.

– Kai Goebel
GE Corporate R&D

Intelligent Distributed and Embedded Systems

This symposium focused on the emerging challenges associated with large-scale distributed embedded systems for sensing and acting. A central goal of this meeting was to bring together three communities to facilitate a cross-fertilization of ideas: (1) distributed AI, (2) distributed robotics, and (3) networking and communications. The first two are well represented at other AAAI Spring Symposia, but the third is traditionally not. This symposium was attended by several researchers from the third area.

Overall, we had 18 exciting talks on topics ranging from the coordination and control of distributed agents, including robots, data dissemination and fusion in embedded sensor networks, and distributed systems for observing and modeling.

The symposium closed with a lively panel discussion. The panelists included Kurt Konolige, Peter Stone, and Ian Horswill.

– Gaurav S. Sukhatme
University of Southern California
– Tucker Balch
Georgia Institute of Technology

Logic-Based Program Synthesis: State of the Art and Future Trends

Automatic programming has long been considered a core AI task. C. Green's and R. Waldinger's papers at the First International Joint Conference on Artificial Intelligence (1969) put it on a firm logical basis and spawned a long line of research. However, logic-based program synthesis is not widely used for software development. Why not? The purpose of this symposium was to survey the current state of the art, identify barriers, and discuss directions that can help make synthesis more feasible. The symposium attracted more than 30 participants. There were 19 contributed presentations that stimulated lively discussions. In addition, three invited talks were given.

One of the highlights of this symposium was the three outstanding invited speakers. The opening presentation was given by Cordell Green (Kestrel Institute). Cordell presented "A Thread of History and Progress in Program Synthesis," a very informative and witty personal view of the field and its future directions from one of its founders. On Monday afternoon, Elaine Kant (SciComp, Inc.) presented "A Commercial Program Synthesis System for Computational Finance," which Elaine's company is developing and actually selling. The system allows financial analysts to quickly specify intricate option pricing models and generate highly optimized C code that solves these models using state-of-the-art numeric methods. On Tuesday morning, Richard Waldinger (SRI) gave a talk entitled "Deductive Chat Lines for Multiple Agents." Waldinger uses deductive synthesis to generate semantic mediators that allow useful communication between agents that have been designed independently of each other. He demonstrated his approach with a geographic question-answering system and a smart personal organizer.

The contributed presentations covered a surprisingly wide range of topics and showed that there is still a diverse and viable research community active in logic-based program synthesis. The talks were approximately evenly divided between theoretical foundations and practical applications. The theoretical talks dealt with complexity-theoretic issues of program synthesis, the use of proof plan-

… many of us had quite different conceptions about what "safe learning" was. These conceptions differed based on which part of the agent was of primary concern: execution, planning, or learning.

ning and incomplete proofs, and aspects of various nonclassical logics. Practical applications were shown for security protocols for electronic commerce, optimization of network protocols, network service composition, software design synthesis, statistical data analysis, and high-performance schedulers, to name just a few.

From the feedback received from many participants, this was a very successful symposium. All talks sparked lively discussions, which were often continued throughout the breaks. Many talks were accompanied by system demonstrations. Overall, the symposium proved that logic-based program synthesis is still an active research area that finds surprising and exciting new applications, for example, semantic mediation for web-based services.

The chairs would like to thank everyone who contributed to the success of this symposium. Special thanks go to the invited speakers, Mike Lowry for his plenary presentation, and the members of the program committee.

– Bernd Fischer
NASA Ames Research Center
– Doug Smith
Kestrel Institute

Mining Answers from Texts and Knowledge Bases

This symposium provided a forum for researchers involved in building question-answering systems either by processing massive text collections or reasoning from knowledge bases to identify common ground, relevant experiences, applications, open problems, and possible future experiments.

The opening talk was given by an invited speaker, John Prange, director of the AQUAINT program in ARDA, which highlighted the many challenges imposed by advanced question answering. Prange surveyed the significance of question answering as a technique in the service of the professional information analyst and pointed out that the first challenge faced by researchers sponsored in the AQUAINT program is determined by the scenarios and problematics that depart from factual, independent questions studied until now in the natural language–processing and information-retrieval communities.

Throughout the first day of the symposium, several language-based methods of mining answers from texts stressed the combination of diverse forms of knowledge with text-processing mechanisms. In her presentation, Sanda Harabagiu discussed several forms of abductive reasoning that are developed for justifying the correctness of answers extracted from texts. A different intersection of interests was determined by the semantic processing of questions and its impact on the accuracy of mining answers from texts. Marius Pasca described an elaborate procedure for resolving definition questions that was proven successful in the NIST TREC evaluations.

Several speakers were interested in approaches of mining answers by using information-extraction techniques. In her presentation, Un Yong Nahm described the DISCOTEXT system that uses a learned information-extraction system to transform text in a more structured data, which is then mined for interesting relations. Information extraction was also considered by Paul Morarescu and Finley Lacatusu, who used this technique for creating multidocument summaries as answers for complex, synthesis-requiring answers.

The second day of the symposium started off with a panel entitled "Large Knowledge Bases," led by Adam Pease. Reinhard Stolle gave an illuminating presentation on overlaying an analogical reasoning engine that operates on the output from a natural language–processing engine. Xerox PARC uses this technology to find similarities between reports from different technicians.

Alon Halevey showed a model of question answering over multiple web sites. He presented a question-answering technique that exploits the XML annotations and the instance data in the web site to determine the schema of the information.

Bruce Porter and Ken Barker argued that an optimum way to combine natural language techniques and the knowledge-based techniques is to identify domains where it makes sense to invest significant manual effort in creating a domain model. Michael Whitbrock presented CYC's approach to combining natural language–based knowledge entry and question answering with their large commonsense knowledge base.

The symposium concluded on an enthusiastic note that even though the natural language processing, information retrieval, and knowledge representation communities have had fairly disjoint focus, we should continue to create synergy and cross-fertilization across them.

– Sanda Harabagiu
University of Texas at Dallas
– Vinay Chaudri
SRI International

Safe Learning Agents

Although there have been workshops and symposia that have discussed safe agents, this symposium was the first that specifically focused on safe learning agents. Learning agents have both challenges and opportunities

not associated with nonlearning agents. Learning allows agents to be situated in environments that are unknown or are likely to be changing in unanticipated ways. However, the task of validating and verifying the safety of such systems is just beginning to be explored.

The symposium had three invited speakers and eight accepted papers. The small size of the group led to many informal discussions that often continued into dinner. The papers and discussions often raised far more questions than they answered. Although there was some consensus on some issues, there was much more often an unexpected diversity of opinions and views.

We all agreed that "safety" was not the same as "utility" and that although safety might be made an aspect of utility, it seemed more useful at this time to keep them separate. We felt that a course of action could both have high utility and high risk (that is, be "unsafe"). Although there has been a lot of work on incorporating the notion of utility into the agent framework, there has been little work to date on incorporating the notions of risk and safety.

However, many of us had quite different conceptions about what "safe learning" was. These conceptions differed based on which part of the agent was of primary concern: execution, planning, or learning. For those interested in execution, their definition of safe learning was that the actions executed by the agent after learning were still safe actions. For those interested in planning, safe learning was that the planning process was still safe (that is, that the planner's performance did not degrade as a result of the learning). For those interested in the learning itself, safe learning was that the agent was still safe while it was in the process of learning (for example, that learning was postponable). Finally, for those who were focusing on what was being learned, safe learning was learning what was safe (for example, situated in an unknown and possible hostile environment, an agent might learn which actions could be dangerous).

There were two basic approaches to safe learning agents presented in the papers. The most popular approach was to "prove" that the system was safe. There were two categories of safety being proved. One category was that the system could not learn anything that would be unsafe. The other category was that the system would converge to safeness. The other approach was to identify what could make a system unsafe and avoid them.

From the papers presented and the discussions that occurred during the symposium, it seemed obvious that agents are being delegated more and more tasks and that safety is going to become a critical issue in the general acceptance of these agents. Given the increasingly integrated nature of fielded agents, learning and adaptation are going to be components in most future intelligent agent systems.

– Mike Barley and Hans Guesgen
University of Auckland

Sketch Understanding

As computation becomes pervasive and embedded, people are becoming increasingly dissatisfied with WIMPy interfaces. If it's absurd to suggest to people that they should interact with one another by sitting at a computer and silently typing, clicking, and dragging, why then do we interact with our software this way? Researchers in a variety of contexts are working to enable more natural forms of interaction in a variety of modalities. This symposium focused on sketch understanding as one such form of natural interaction.

By *sketch*, we mean an informal drawing created with pen strokes. By *understanding,* we mean reliably identifying the objects or concepts suggested by the pen strokes, despite the inaccuracies and ambiguities inherent in the medium. One measure of understanding is the ability to answer questions about the things depicted. Understanding a sketch of a physical device, for example, means being able to answer questions about how the device operates, what it might be useful for, and how it might be constructed.

These issues seemed to strike interest that was widespread, attracting participants, both from academia and industry, from around the world, including Belgium, China, England, Japan, Portugal, and the United States.

The primary objectives of the symposium were the dissemination of initial research results from the small but growing community of researchers working in this area and the establishment of a research agenda for the field. The format included invited speakers, paper presentations, software demonstrations, and panel discussions. The first invited speaker, Robert E. Horn, spoke about his work in a presentation entitled "Visual Language: Global Communication for the 21st Century." Bert Keely, architect of eBooks and the tablet pc at Microsoft, offered an intriguing look at the forthcoming tablet pc.

The panel discussion exploring the application of speech understanding and vision research to sketch understanding identified a number of the techniques from those fields that will have direct application to sketch understanding. The panel on multimodal interfaces (that is, those combining sketch input with speech, vision, gestures, and other modes of input) identified a number of unresolved research issues, such as how to use multiple modes to handle errors and ambiguities.

The symposium culminated in a brainstorming session aimed at mapping out a research agenda. Several specific research objectives were identified, such as the establishment of a corpus of sketches to be used for algorithm testing. Another outcome was the establishment of a web site for the sketch understanding research community.[1]

– Thomas Stahovich
Carnegie Mellon University
– Randall Davis
Massachusetts Institute of Technology
– James Landay
University of California at Berkeley

Note

1. This web site will be an extension of the symposium web site, www.me.cmu.edu/faculty1/stahovich/sketchsymposium.htm.