Report on the First International Conference on Knowledge Capture (K-CAP)

Yolanda Gil, Mark Musen, and Jude Shavlik

The First International Conference on Knowledge Capture (K-CAP) was held from 21 to 23 October 2001 in Victoria, British Columbia (Gil, Musen, and Shavlik 2001). This new conference series promotes multidisciplinary research on tools and methodologies for efficiently capturing knowledge from a variety of sources and creating representations that can be (or eventually can be) useful for reasoning. The conference attracted researchers from diverse areas of AI, including knowledge representation, knowledge acquisition, intelligent user interfaces, problem solving and reasoning, planning, agents, text extraction, and machine learning.

Knowledge acquisition has been a challenging area of research in AI, with its roots in early work to develop expert systems. Driven by the modern internet culture and knowledge-based industries, the study of knowledge capture has a renewed importance. Although there has been considerable work over the years in the area, activities have been distributed across several distinct research communities. In machine learning, learning apprentices acquire knowledge by nonintrusively watching a user perform a task. In the human-computer interaction community, programming-by-demonstration systems learn to perform a task by watching a user demonstrate

how to accomplish it. In knowledge engineering, modeling techniques and design principles have been proposed for knowledge-based systems, often exploiting commonly occurring domain-independent inference structures and reusable domain-specific ontologies. In planning and process management, mixed-initiative systems acquire knowledge about a

■ The International Conference on Knowledge Capture (K-CAP) is a new forum for multidisciplinary research on capturing knowledge from a variety of sources and creating representations that are useful for reasoning. This article describes the first conference series, held in October 2001, and presents an invitation to the AI community to participate in K-CAP 2003.

user's goals by taking commands or accepting advice regarding a task. In natural language processing, tools can process text and create representations of its knowledge content for uses such as question answering. All these approaches are related in that they acquire information and organize it in knowledge structures that can be used for reasoning. They are complementary in that they use different techniques and approaches to

capture different forms of knowledge.

The conference started off with three invited tutorials and a workshop program that represented these different approaches. Frank van Harmelen, Dieter Fensel, and Heiner Stuckenschmidt talked about the semantic web as a vehicle for knowledge capture, representation, and reasoning. They described ongoing work on standards and technologies for semantic markup. A workshop on the same topic included presentations on tools for capturing such semantic annotations of content and services on the web. Henry Lieberman surveyed successful techniques for *programming by example,* an approach where end users teach procedures to computers by demonstrating a sequence of actions on concrete examples as they use an interface. The computer can generalize from these examples through a combination of machine learning techniques and interface task representations. This tutorial included practical exercises and illustrated the concepts with applications, including creation of animation and games, automation of text editing and graphic editing tasks, and computer-aided design and manufacturing. A workshop on interactive tools for knowledge capture discussed techniques for acquiring procedural knowledge as well as ontologies. Guus Schreiber and Hans Akkermans gave a tutorial on COMMONKADS, a widely known methodology for knowledge engineering and management. It covered some of the main ideas behind the methodology, including reusable task models, inference and domain knowledge, and organizational context. The tutorial included a valuable discussion on practical experiences with the methodology in industrial settings as well as knowledge-based techniques for non-AI audiences. A related workshop discussed industrial applications of knowledge-oriented approaches relevant to electronic business.

Three invited speakers pointed the attendees to promising directions for knowledge capture. Ken Forbus, of Northwestern University, talked about the need for large knowledge

bases to build intelligent systems that have the flexibility and breadth of human reasoning and discussed analogy and sketching as promising knowledge capture techniques for bootstrapping intelligent systems. Steve Lawrence, of NEC Research Institute, described the design, implementation, and operation of RESEARCHINDEX (also known as CITESEER), a digital library of scientific literature built automatically through text-extraction techniques over the web that is used daily by many researchers and institutions. John McCarthy, of Stanford University, discussed the need for data mining that goes beyond finding relations among data and instead finds relations between the phenomena that give rise to the data. In this phenomenal data mining, relations are established among entities and facts that arise from those observations in a process that requires common sense and domain knowledge.

The K-CAP 2001 proceedings are available online from the ACM Digital Library.[1] They include papers on many important topics for the conference, including ontologies and knowledge representation, interactive acquisition tools, collaborative and distributed knowledge acquisition, information extraction, knowledge management, semantic markup, task-oriented knowledge and problem solving, adaptive user interfaces, and learning from examples.

The future of knowledge capture is vibrant and well funded. There is a large community of researchers and practitioners interested in the vision of a semantic web that will contain annotations that can be used to reason about web content and services. This vision, initially brought forward by the World Wide Web consortium, is motivating many AI researchers to look at the technology that will enable the expression and capture of such annotations and is prompting new challenges for knowledge representation and capture. Funding agencies actively support a wide range of relevant work on knowledge base development, user-centered planning and decision aids, agent-based assistants, and knowledge-based machine

learning. The needs and application of knowledge capture are many and continue to grow in science, government, and industry. The first edition of the conference had over 80 attendees from 5 continents. The conference attracted a large number of researchers that traditionally attended the Banff series of knowledge acquisition workshops that was held from October 1986 to 1999.

As conference cochairs, we would like to acknowledge our sponsors and other people that contributed to the conference. The Association of Computing Machinery (ACM) provided financial sponsorship and produced the proceedings. ACM's SIGART, the American Association for Artificial Intelligence, and IFIP TC 12 supported the event. Funding for the conference was provided by the Air Force Office of Scientific Research, the Defense Advanced Research Projects Agency, the Office of Naval Research, and STRICOM. John Gennari handled our finances and coordinated with ACM's sponsorship program. Rob Kremer proposed a wonderful location for the conference, the relaxing Laurel Point Inn, and handled the local arrangements together with Tim Menzies. Fensel organized an excellent workshop program.

The Second International Conference on Knowledge Capture will be held 23 to 25 October 2003 at Sanibel Island in Florida. It will be collocated with the Second International Semantic Web Conference.[2] We look forward to the future of the knowledge capture conference and its research community.

Note

1. portal.acm.org.
2. A call for papers and other conference details will be posted at www.k-cap.org. For any inquiries, please e-mail info@k-cap.org.

Reference

Gil, Y.; Musen, M.; and Shavlik, J., eds. Proceedings of the First International Conference on Knowledge Capture. New York: Association of Computing Machinery.

Yolanda Gil is associate division director for research for the Intelligent Systems Division at the Information Sciences Institute and a research assistant professor in the Computer Science Department at the University of Southern California. She is principal investigator for the EXPECT Project, with a research focus on interactive acquisition tools that help end users develop and maintain large knowledge bases, apply them in problem-solving contexts, and exploit knowledge through the semantic web. She received her Ph.D. in computer science from Carnegie Mellon University and her undergraduate degree from the Polytechnic University of Madrid. She recently cochaired the new conference on knowledge capture and was program chair of the 2002 conference on intelligent user interfaces.

Mark A. Musen is an associate professor of medicine (medical informatics) and computer science at Stanford University and is head of the Stanford Medical Informatics Laboratory. He conducts research related to knowledge acquisition for intelligent systems, knowledge system architecture, and medical decision support. He has directed the PROTÉGÉ project since its inception in 1986, emphasizing the use of explicit ontologies and reusable problem-solving methods to build robust knowledge-based systems. He has an M.D. from Brown University and a Ph.D. from Stanford.

Jude Shavlik is a professor of computer sciences and of biostatistics and medical informatics at the University of Wisconsin at Madison. His research focuses on machine learning, especially its application to bioinformatics and computer security. He received his Ph.D. in 1988 from the University of Illinois at Urbana/Champaign for his work on explanation-based learning. From 1996 to 1999, he served as editor in chief of the *AI Magazine*. Besides helping organize KCAP-2001, he was the chair of the 1998 International Conference on Machine Learning, was the cochair of the 1993 International Conference on Intelligent Systems for Molecular Biology, and is the conference chair of the 2003 IEEE Data-Mining Conference.