

The Constraint-Based Knowledge Broker System

J.-M. Andreoli, U. M. Borghoff, P.-Y. Chevalier, B. Chidlovskii, R. Pareschi, J. Willamowski
Rank Xerox Research Centre, Grenoble Laboratory, F-38240 Meylan, France

System Overview

The amount of information available from electronic sources on the World Wide Web and other on-line information repositories is highly heterogeneous and increases dramatically. Tools are needed to extract relevant information from these repositories. The *Constraint-Based Knowledge Brokers* project (CBKB) at RXRC Grenoble realizes sophisticated facilities for efficient information retrieval, schema integration, and knowledge fusion [1-5]. The current implementation of the CBKB research prototype involves three kinds of agents: a) *users*, who input queries and process answers (i.e., ranking, fusion) through a GUI; b) *wrappers*, capable of interrogating heterogeneous information sources, which can provide answers to elementary queries (essentially various public bibliographic catalogues available on the Web, as well as preprint archives and opera information repositories); c) *brokers*, which can manage complex queries (i.e., decompose a complex query, re-compose the partial answers, synthesize a full answer) and which mediate between the GUI and the different wrappers.

The core of the system is given by the brokers, which provide various important services such as intelligent caching, filtering and knowledge combination. Requests, intermediate information, and results are internally represented via feature constraints. Requests do not need to be fully defined; they may correspond to partial specifications of the requested information. Furthermore, requests that cannot be fully satisfied may still obtain results in the form of partial objects that refine the initial requests by instantiating some of its attributes or by adding new attributes. Efficient constraint solvers allow for local filtering/sifting of information.

CBKB-related projects include: *COIN* (MIT), *Harvest* (Univ. Colorado), *Garlic* (IBM Almaden), *HERMES* (Univ. Maryland), *Information Manifold* (AT&T), *InfoSleuth* (MCC), *KRAFT* (Univ. Aberdeen / Wales / Liverpool, BT), *SIMS* (ISI), and *TSIMMIS* (Stanford Univ.).

Besides information retrieval from the World Wide Web, another possible application of the CBKB model is dynamic document composition from on-line repositories: the widespread availability of new electronic sources of informa-

tion such as E-mail, the Web, and other on-line information repositories also multiplies the number of electronic documents available. Documents can now be built dynamically by accessing and combining information existing over distributed sources. Hierarchical mark-up languages like SGML can be used to define document templates that can be dynamically filled in with heterogeneous components. These documents can in turn be made permanent by storing them in document management systems, thus entering the normal document lifecycle. The CBKB project is well-aligned with existing standardization efforts in document management and provides an implementation of document search as specified by DMA (Document Management Alliance), an industry standard concerned with the search, retrieval, storage and conversion of documents on heterogeneous document management systems. DMA is the result of the merging of two previously existing standardization efforts: Shamrock (backed by IBM and Saros) and DEN (backed by Xerox and Novell).

With any Java-enabled Web browser you may connect to the CBKB system on-line via <http://www.rxrc.xerox.com/cbkb-cgi/main>.

For further questions concerning the project, contact Uwe.Borghoff@grenoble.rxrc.xerox.com.

References

- [1] J.-M. Andreoli, U. M. Borghoff, and R. Pareschi. The constraint-based knowledge broker model: Semantics, implementation and analysis. *J. Symbolic Computation*, 22, 1996.
- [2] J.-M. Andreoli, U. M. Borghoff, R. Pareschi, and J. H. Schlichter. Constraint agents for the information age. *J. UCS*, 1(12):762-789, Dec. 1995. <http://www.iicm.edu/jucs>.
- [3] U. M. Borghoff, P.-Y. Chevalier, and J. Willamowski. Adaptive refinement of search patterns for distributed information gathering. In A. Verbraeck, editor, *Proc. Int. Conf. EuroMedia/WEBTEC '96*, pages 5-12, London, U. K., Dec. 1996.
- [4] U. M. Borghoff, R. Pareschi, F. Arcelli, and F. Formato. Constraint-based protocols for distributed problem solving. *Science of Computer Programming*, 29, Oct. 1997.
- [5] U. M. Borghoff and J. H. Schlichter. On combining the knowledge of heterogeneous information repositories. *J. UCS*, 2(7):515-532, July 1996. <http://www.iicm.edu/jucs>.