Case-Based Reasoning: Business Applications

BRADLEY P. ALLEN

Case-based reasoning (CBR) is an approach to problem solving based on the retrieval and adaptation of *cases,* or episodic descriptions of problems and their associated solutions. In theory, CBR can be considered as a five-step problem-solving process (adapted from [6]):

Presentation: a description of the current problem is input to the system.
Retrieval: the system retrieves the closest-matching cases stored in a *case base* (i.e., a database of cases).
Adaptation: the system uses the current problem and closest-matching cases to generate a solution to the current problem.
Validation: the solution is validated through feedback from the user or the environment.
Update: if appropriate, the validated solution is added to the case base for use in future problem solving.

In practice, both the implementation or inclusion of each of the five steps and the representation used for cases can vary widely. For example, case retrieval has been implemented using nearest-neighbor algorithms, decision trees, or connectionist associative memories; case representations range from free-text documents to database records to semantic networks. CBR technology shares much in common with instance-based approaches to machine learning [2] and techniques for intelligent information filtering and retrieval [3].

Researchers have claimed that CBR provides the potential for developing knowledge-based systems (KBS) more easily than with rule- or model-based approaches [9, 16]. They argue a) that the concrete examples provided by cases are easier for users to understand and apply in various problem-solving contexts than complex chains of reasoning generated by rules or models and b) that record-like representations of cases used in some CBR systems allow for straightforward storage in relational databases and entry and update by end users [16]. Another benefit is that the presence of the validation and update steps provides a framework for learning from experience, thus incorporating knowledge acquisition as part of the day-to-day use of a CBR application [10].

Though anecdotal accounts of CBR development efforts are optimistic that these benefits can be achieved, empirical studies are needed to validate these claims. CBR may not be as effective as rule- or model-based approaches for applications where theory, not experience, is the primary guide to problem solving, and where solutions are unique to a specific problem instance and not easily reusable. In response to this, hybrid problem-solving architectures combining cases, rules, and models have been developed to allow the strengths of each form of reasoning to complement those of the others [4, 15]. Open problems in the application of CBR include the management of the size and granularity of case representations [15], the validation of case bases as a whole [10], and the enforcement of consistent use of terminology by case authors [1].

Much of the foundation for the CBR approach was laid through research performed by Roger Schank and his students at Yale University in the 1980s. During the late 1980s, the Defense Advanced Research Projects Agency (DARPA) began to fund efforts toward the application of CBR technology. This, in turn, attracted the attention of organizations interested in commercial applications of KBS technology. The past four years have begun to see the use of CBR in deployed business applications.

Current Commercial Use

Although a wide range of potential applications has been explored, commercial CBR applications have focused for the most part on using case retrieval for decision support [13]. Case retrieval avoids the step of case

adaptation by simply retrieving the closest-matching cases given a problem description, allowing decisions to be based on the most similar available precedents.

Problem resolution in customer service help desks is an example of the commercial application of case retrieval [17]. Customer service is an area where problem-solving knowledge is volatile, especially given the complex and rapidly evolving products manufactured by high-tech industries. The knowledge-engineering effort associated with traditional expert systems is too costly for the average customer service organization to undertake. This argues for a CBR approach where knowledge acquisition and distribution can be part of standard operations.

An example of a deployed CBR customer service application is Compaq Computer's SMART system [1]. SMART is an integrated call-tracking and problem resolution system that contains hundreds of cases related to diagnostic problems arising in the use of Compaq products. The system is in use by Compaq's Customer Service Department in handling calls to the central toll-free number. Incoming customer problems are presented to SMART, which retrieves the most similar cases from its case base and presents them to the customer service analyst, who then uses them to resolve the problem. Evaluation of the initial version of SMART indicated the percentage of customer problems resolved on the first call rose from 50% without the system to 87% with the use of the system. In terms of productivity benefits, Compaq estimates that SMART has paid for itself within a year. Compaq is also distributing a case base directly to its customers as a part of the PageMarq printer line: over 3,000 copies of QuickSource, an application containing 500 diagnostic cases, have already been shipped [14]. Both applications run on 386-based personal computers running the Microsoft Windows operating environment.

The use of case retrieval in customer service can be viewed as a special instance of the application of case retrieval to the automation and reengineering of business processes, through the wide-scale distribution of technical or managerial expertise. An example of this type of application is the Prism telex classification system developed by Cognitive Systems, Inc. [9]. Prism is used in several banks to route incoming international telex communications to appropriate recipients, a task relying on the accurate classification of the telex to determine the appropriate routing. Prism increased classification accuracy from 75% to 90% and reduced the average time taken to route a telex from several minutes (for human telex operators) to 30 seconds. Prism also demonstrated improvements in accuracy and speed over a previous rule-based implementation of the same type of system. Prism runs on Macintosh-II workstations.

Another example of business process reengineering using CBR is NEC Corporation's SQUAD system [12]. SQUAD is an ambitious long-term effort to provide a corporate-wide system for the capture and distribution of software quality control knowledge. Some 3,000 cases per year have been added to the system since 1982, and the developers estimate the productivity savings due to the use of SQUAD at over $100 million per year. SQUAD runs on Unix workstations.

Systems such as SMART, Prism, and SQUAD provide examples of how simple CBR techniques can be effectively integrated with daily business operations. Although some workers in the field have exploited massively parallel hardware for the implementation of case retrieval systems [5], the systems mentioned here demonstrate that CBR can be effectively implemented using widely available computing environments.

Applications of CBR that go beyond case retrieval to incorporate adaptation for problem solving have not received as much attention in the commercial world. An exception is in the area of design and configuration, where parameter-driven design techniques and reusable components make the problem of adaptation simpler to address. Case-based design and configuration systems that support the reuse and modification of standard designs have been implemented by manufacturing organizations interested in solutions specific to their operations. Two notable examples are Nippon Steel's QDES [11] and Lockheed's Clavier [10]. Extending the CBR approach to automate complex synthetic tasks such as planning and scheduling is an active area of research that is driving the development of improved representations of procedural cases [18].

Method and Tools for Building Systems

Methodologies to guide CBR system development efforts are now beginning to emerge. Compaq, NEC, Inference, and Cognitive Systems have each developed methodologies for case base development. These approaches share the following three phases:

Case-Base Design: a general representation for cases is developed using source materials at hand (e.g., documentation, database records, and written accounts by experts). This is accomplished by a coordinated effort involving users, managers, and system developers. Case-base design can involve the compilation of a lexicon of terms used to describe problem features, the selection of appropriate features for indexing cases, the specification of database schemas used to store cases, and the definition of case base authoring standards.

Initial Case-Base Development: a "seed" case base is developed to provide a baseline for the application. The initial case base is extensively reviewed by the developers and users and is iteratively refined until a valid case base covering a large portion of the application area is complete.

Ongoing Development and Maintenance: the initial case base is used and refined through the validation and storage steps, performed concurrently throughout the user base. Part of the organization manages the case base and case representation in the same manner as a database administrator in a traditional data processing organization. Statistical quality control techniques can be used to monitor case accuracy and utility.

Commercial tools for developing CBR systems tend to be simpler than

experimental research systems, relying primarily on case retrieval and record-like representations, with minimal support (at best) for adaptation. This has enhanced the technology transfer into commercial use, leveraging the widespread understanding of issues related to relational database design and implementation. A typical CBR development environment in 1994 provides default database schemes for case representation and a predefined problem-solving flow for case-based decision support. Forms are used for editing cases, features, and solutions. Utilities are provided for manual and automated indexing of cases, the automatic import of case from records in relational database tables, conceptual clustering of cases for analysis and efficient retrieval, and the annotation of cases and actions with text and pointers to other applications and processes. Examples of currently available CBR tools include Inference Corporation's CBR Express and CasePoint, Cognitive Systems' REMIND, and Esteem Systems' ESTEEM.

Future Directions

In the near future, CBR system development methods and support tools will continue to evolve providing more sophisticated approaches to case representation and indexing [7, 18]. They will also allow the creation of hybrid problem-solving applications combining cases, rules, and models [4, 15]. Techniques from research in information filtering and retrieval [3] may be brought to bear in the automatic generation and indexing of cases from unstructured text. Work along the lines of the ASK systems at Northwestern University [8] will be applied to the development of CBR systems using multimedia source materials such as full-motion video clips.

As businesses begin to build large-scale case bases as part of enterprise-wide information systems, newer issues of case acquisition and distribution will begin to arise, such as the design of business processes to support case-base acquisition and maintenance [12], distributed case-base maintenance, and support for multiple national languages.

Support for CBR may eventually be available in future operating systems, where system-level primitives for concept-based information retrieval could form a foundation for the large-scale distribution of case retrieval applications. This would enable the evolution of CBR application development into a general practice of knowledge publishing, where standard case representations and retrieval interfaces would allow developers to focus solely on case base content. These systems, together with next-generation refinements of current case retrieval front-ends, may form the basis for "corporate memories" [13] capable of storing and distributing the experience of a large organization to all of its members, in support of the entire scope of its activities. ◘

References

1. Acorn, T. and Walden, S. SMART: Support Management Automated Reasoning Technology for Compaq Customer Service. In *Proceedings of the 4th Innovative Applications of Artificial Intelligence Conference.* 1992.
2. Aha, D.W., Kibler, D., and Albert, M.K. Instance-based learning algorithms. *Mach. Learn. 6,* 1 (Jan. 1991).
3. Belkin, N.J. and Croft, W.B. Information filtering and information retrieval: Two sides of the same coin. *Commun. ACM 35,* 12 (Dec. 1992).
4. Branting, L.K. and Porter, B.W. Rules and precedents as complementary warrants. In *Proceedings of the 9th AAAI Conference,* AAAI, Anaheim, CA 1991.
5. Creecy, R.H., Masand, B.M., Smith, S.J., and Waltz, D.L. Trading MIPS and memory for knowledge engineering. *Commun. ACM 35,* 8 (Aug. 1992).
6. DARPA. Case-based reasoning from DARPA: Machine learning program plan. In *Proceedings of the 1989 DARPA Case-Based Reasoning Workshop.* DARPA, Washington, D.C., 1989.
7. Domeshek, E.A. What Abby cares about. In *Proceedings of the 1991 DARPA Case-Based Reasoning Workshop.* DARPA, Washington, D.C., 1991.
8. Ferguson, W., Bareiss, R., Birnbaum, L., and Osgood, R. ASK Systems: An approach to the realization of story-based teachers. Tech. Rep. 22, Inst. for Learning Sciences, Northwestern Univ., April, 1992.
9. Goodman, M. Prism: A case-based Telex classifier. In *Proceedings of the 2nd Innovative Applications of Artificial Intelligence Conference.* 1990.
10. Hennessy, D. and Hinkle, D. Applying case-based reasoning to autoclave loading. *IEEE Exp. 7,* 5 (Oct. 1992).
11. Iwata, Y. and Obama, N. QDES: Quality-Design Expert System for steel products. In *Proceedings of the 3rd Innovative Applications of Artificial Intelligence Conference.* 1991.
12. Kitano, H., Shibata, A., Shimazu, H., Kajihara, J., and Sato, A. Building large-scale and corporate-wide case-based systems: Integration of organizational and machine executable algorithms. In *Proceedings of the 10th AAAI Conference.* AAAI, San Jose, CA 1992.
13. Kolodner, J.L. Improving human decision making through case-based decision aiding. *AI Mag. 12,* 2 (Summer 1991).
14. Ngyuen, T., Czerwinski, M., and Lee, S.D. COMPAQ QuickSource: Providing the consumer with the power of artificial intelligence. In *Proceedings of the 5th Innovative Applications of Artificial Intelligence Conference.* 1993.
15. Pearce, M., Goel, A.K., Kolodner, J.L., Zimring, C., Sentosa, L., and Billington, R. Case-based design support: A case study in architectural design. *IEEE Exp. 7,* 5 (Oct. 1992).
16. Riesbeck, C.K. An interface for case-based knowledge acquisition. In *Proceedings of the DARPA Workshop on Case-Based Reasoning.* DARPA, Washington, D.C., 1988.
17. Simoudis, E. Using case-based retrieval for customer technical support. *IEEE Exp. 7,* 5 (Oct. 1992).
18. Zito-Wolf, R. and Alterman, R. A framework and an analysis of current proposals for the case-based organization and representation of procedural knowledge. In *Proceedings of the 11th AAAI Conference.* AAAI, Washington, D.C. 1993.

About the Author:

BRADLEY P. ALLEN is Vice President of Applications Development at Inference Corp. His current research involves the use of machine learning techniques to improve information access and discovery. **Author's Present Address:** Inference Corp., 550 N. Continental Blvd., El Segundo, CA 90245; allen@inference.com