A Review of
Case-Based Reasoning

Andrew R. Golding

Case-Based Reasoning, Janet Kolodner, Morgan Kaufmann Publishers, San Francisco, California, 1993, 668 pp., $54.95, ISBN 1-55860-237-2.

Case-based reasoning (CBR) is becoming a viable real-world technology. Case-based systems, such as the Lockheed CLAVIER system for autoclave layout, have successfully been deployed, and CBR shells, analogous to expert system shells, have made their way into the marketplace. As the field grows, so does the need for a good, comprehensive source on this area. Janet Kolodner's book is the first to fulfill this need. Not only does the book describe the state of the art, it also presents the material in terms of a unified framework for CBR, thereby improving the coherence of the literature and our understanding of it. CBR is analyzed as a general sequence of steps grouped into three phases: (1) building the case library, (2) retrieving cases, and (3) applying the retrieved cases to the target situation. Each step is presented in its abstract form together with the range of approaches that have been tried for it and concrete examples from systems. The only other comparable reference is Christopher Riesbeck and Roger Schank's 1989 book *Inside CBR* (Lawrence Erlbaum, Hillsdale, N.J.), which is primarily a series of vignettes of individual systems. Progress in the field since 1989 has enabled Kolodner to expand on this earlier effort in terms of both the diversity of systems covered and the generality of the treatment.

Although the presentation in terms of a unified framework is a major contribution, it also, unfortunately, has

inherent drawbacks. First, it fragments each CBR system across many chapters, making it difficult to get the big picture of how the system works and obscuring the interrelatedness of the system's parts. In addition, having each chapter draw its examples from multiple systems adds a certain context-switching overhead: Each time a system is introduced (or reintroduced), the book must set the context anew, and the reader must recall the details of the system. A second drawback to the unified framework is that although it has fairly broad coverage, it is still biased toward those systems that fit it best. As a result, important work sometimes gets only a cursory mention in the book. Examples include Carbonell and Veloso's seminal research on derivational analogy, Kitano's use of CBR as a corporate memory in the SQUAD system, and the work by Waltz and others on memory-based reasoning.

Other issues that I would like to have seen addressed further include evaluation and a discussion of related fields. In regard to evaluation, methods for performing the different steps of CBR are often presented without any experimental results—what size case library the method works for, what degree of speedup it affords, and so on. This makes it hard to judge the applicability of the method, which is perhaps more a critique of the field, circa 1993, than of Kolodner's reporting of it. However, since 1993, there has been increased interest in evaluation—in particular, at the 1994 AAAI CBR Workshop—which might well be worth including in future editions.

As for related fields, CBR is com-

pared with traditional rule-based systems, whose rules are extracted from experts, but not with systems that learn their own rules (or other representation) from examples. The latter comparison would be instructive because such learning systems, by working from examples, are arguably close competitors to CBR. The book also casts CBR as a cognitive theory and, in this connection, mentions SOAR and PRODIGY, two systems embodying theories of learning from experience. Some explicit comparisons would be useful here, however, for example, between the methods used by these systems and by CBR for indexing experiences for reuse.

The book is generally well organized; each chapter has extremely lucid and crisp introductory and summary sections, which are particularly helpful. The organization in the interior portions of the chapters is not always as tight as it could be however. Sometimes a section will introduce a list of points to be covered, but the subsections will not directly correspond to this list; a subsection might cover more than one of the points, or it might not cover any of them and instead introduce alternate material. Sections also tend to have long, monolithic stretches of text. These factors can make it difficult for the reader to look things up easily or skim through a chapter to get the main ideas.

The book is divided into five parts and an appendix. Part 1 covers background: an overview of CBR, a well-balanced set of case studies of CBR systems, and a review of dynamic memory (the cognitive model behind Kolodner's unified framework).

Parts 2, 3, and 4 work through the steps of the unified framework. Part 2 covers the case library: the content and representation of cases and the assignment of indexes to each case. Part 3 deals with retrieving cases from the case library: data structures and algorithms for efficient retrieval and schemes for obtaining an ordered list of the best cases for a target situation. Part 4 covers what to do with cases once they have been retrieved: how to adapt them to the target situation, how to use them to interpret or eval-

uate the target situation, and how to apply multiple source cases to a single target problem. The only disappointment here was Chapter 13 on interpretation and evaluation: It merely describes illustrative systems rather than lays out the abstract ideas plus the range of approaches tried, as is done in the rest of the book.

After parts 2 through 4 present the individual steps of CBR, part 5 pulls them together into a coherent whole. It discusses the issues involved in actually sitting down and designing a CBR system, ending with some conclusions and open problems. This part of the book really sparkles, particularly Chapter 15, which conveys Kolodner's obviously extensive experience and insight in building case-based systems. Topics covered include extracting cases from experts, maintaining the case library in a fielded system, and designing an effective user interface for a case-based system.

The final part of the book is an appendix of short summaries of CBR systems. This list is useful, although because it was acquired as the result of a query passed around on e-mail, it is far from complete. It lacks many systems that have been reported, for example, in CBR workshop notes.

In summary, Janet Kolodner's book fulfills an important need for a comprehensive reference on CBR. The main weaknesses of the book are its fragmentation of CBR systems across chapters and its underrepresentation of certain systems; it is not for the reader who seeks a complete, integrated description of some particular system. Despite these shortcomings, however, it is the best general treatment available on CBR, covering a broad range of the literature and situating everything within a common framework. I would recommend it for the graduate student, researcher, or CBR practitioner who wants to understand the process and underpinnings of CBR.

Andrew R. Golding is a research scientist at Mitsubishi Electric Research Laboratories. His research interests are in using machine-learning and case-based techniques to develop high-accuracy computer systems, particularly in the area of natural language.