Artificial Intelligence Review (2005) 24: 103–108
DOI 10.1007/s10462-005-4606-8

Introduction to the Special Issue on Explanation in Case-Based Reasoning

DAVID LEAKE[1] & DAVID MCSHERRY[2]
[1]*Computer Science Department, Indiana University, Lindley Hall 215, 150 S. Woodlawn Avenue, Bloomington, IN 47405, USA (E-mail: leake@cs.indiana.edu);*
[2]*School of Computing and Information Engineering, University of Ulster, Coleraine BT52 1SA, Northern Ireland, UK; (E-mail: dmg.mcsherry@ulster.ac.uk)*

The importance of intelligent systems having the ability to explain their reasoning is well recognized (e.g., Buchanan and Shortliffe 1984; Southwick 1991; Swartout et al., 1991; Leake 1996; Herlocker et al., 2000; McSherry 2001). Users are more likely to accept an intelligent system's conclusions if they can see for themselves the arguments or reasoning steps on which they are based. In situations where the solution is not clear-cut, it is reasonable for users to expect the system to explain the pros and cons of suggested alternatives. In domains such as fault diagnosis, users may also wish to query the relevance of information requested by the system, for example when asked to perform a test that carries high risk or cost. Likewise, explanations may play an important role in the internal processes of an intelligent system. For example, explanations of problem circumstances may help the system to choose appropriate responses and may also help to guide learning (DeJong and Mooney 1986; Mitchell et al., 1986).

In case-based reasoning (CBR), previous problems and their solutions are stored as *cases* in a case library and retrieved in response to a query describing a target problem. In the basic CBR model of problem solving (e.g., Aamodt and Plaza 1994; Kolodner and Leake 1996), the solution to the most similar prior problem, adapted if necessary to take account of differences in problem descriptions, is presented as the proposed solution to the target problem.

The use of explanation to support internal CBR processes has been a focus of research interest since the earliest days in the history of the field. Schank's (1982) Dynamic Memory Theory, a theory of human memory organization and learning that provided early foundations for CBR, views explanations as playing a fundamental role in the indexing of human memories. Early CBR research investigated the use of explanation to guide and support internal system tasks such as retrieval, determining and critiquing similarity, generalizing indices, and case adaptation (e.g., Ashley and Rissland 1987; Barletta and Mark 1988; Branting 1988; Bareiss 1989; Hammond 1989; Kolodner and Simpson 1989; Aamodt 1994). Early work also explored the use of CBR itself to generate explanations for internal system use (Koton 1988) sometimes to serve any of a range of explanation goals (Schank and Leake 1989; Leake 1992).

It has long been recognized that examples can play a central role in the effectiveness of explanations provided to people (e.g., Rissland et al., 1984). In fact, one of the attractions of CBR is that the experience captured in cases provides a natural basis for *precedent-based* explanations of CBR solutions. While rule-based explanations of an intelligent system's conclusions remain an important legacy from expert systems research, recent research supports the view that an explanation based on experience may be more convincing than an explanation based on rules (Cunningham et al., 2003). However, simply presenting examples to a user is not a panacea. CBR researchers have recently questioned the effectiveness of precedent-based explanations in which the user is simply shown one or more of the most similar cases (Doyle et al., 2004; Maximini et al., 2004; McSherry 2004; Roth-Berghofer 2004). This has led to renewed interest in how to provide the support needed to make better use of the explanatory potential of stored cases.

Interest in CBR explanation has also intensified as a result of a recent focus on providing more effective support for interactive problem solving in CBR. Advances in conversational, mixed-initiative and personalized CBR have highlighted the importance of CBR systems having the ability to explain the relevance of questions the user is asked as well as the need for explanations that fit the user's level of expertise (Aha et al., 2001; McSherry 2001, 2002; Leake 2002). Limitations of existing approaches to explanation that have come to light with the emergence of new applications of CBR have also stimulated a renewal of research interest. For example, explanation in recommender systems is an important challenge because of the need to take account of trade-offs with respect to the user's preferences (Shimazu 2002; McSherry 2003). In an intelligent tutoring system, communicating the problem-solving process to students may be as important as finding the right solution (Sørmo and Aamodt 2002).

That explanation is currently a vibrant research topic in CBR is evident from the number of related papers in recent conferences and workshops. This special issue presents extended versions of selected papers from the ECCBR-04 Workshop on Explanation in CBR, an event that attracted the participation of more than 40 CBR researchers and practitioners. Some of the papers accepted for the special issue report developments that advance the state of the art in explanation engineering, while others explore potential new roles of explanation in CBR.

In the first paper, Sørmo et al. provide a foundation for other papers in the special issue, discussing theories of explanation and lessons learned from expert systems research and proposing a classification of explanation types according to the goals supported by the explanations. In light of this framework, they present an analysis of the explanation capabilities of existing systems, review recent developments, and identify challenges that remain to be addressed.

In the second paper, Plaza et al. present an approach by which CBR systems can use explanations of similarity to improve their performance. Their method exploits explanations in the form of symbolic descriptions of similarity for tasks such as selecting a *best* symbolic description of similarity and retrieval of the subset of cases that satisfy this description. The authors also demonstrate the use of these explanations to justify CBR solutions in classification tasks, and to support justification-based learning in multi-agent CBR systems.

The remaining papers in the special issue focus on the role of explanations in increasing user understanding. Nugent and Cunningham present a CBR approach to explaining the predictions of *black box* algorithms such as neural networks or support vector machines. Local feature weights derived from a model of the black box algorithm's behavior in the region of a target problem are used to guide the retrieval of a case that is used to explain the value predicted for the target problem. Local feature weights are also used to highlight important differences between the target problem and the explanation case and explain their effects on the prediction.

The last two papers present new techniques for explaining aspects of the performance of CBR systems. McSherry demonstrates the explanatory power of a conversational CBR approach to product recommendation in a mixed-initiative recommender system. For example, recommendations based on incomplete queries can be *justified* on the grounds that the recommendation will be the same no matter how the user extends her query, and the *relevance* of questions the user is asked can be explained in terms of their ability to discriminate between competing cases. Also in the context of recommender systems, Reilly et al. demonstrate an approach to critiquing in which multi-feature critiques help to explain the retrieval opportunities that remain, relative to the current recommendation. In this way, the user is informed in advance of trade-offs associated with desired improvements, thus reducing the need for back-tracking from dead ends. The authors also show that such critiques have the potential to dramatically improve recommender performance and usability.

We hope that the collection of papers in this special issue will be useful to CBR researchers, practitioners, and graduate students, and others interested in intelligent systems that can explain their reasoning, as a guide to current research on explanation in CBR and signpost to challenges that lie ahead.

Acknowledgements

Our thanks to the program committee of the ECCBR-04 Workshop on Explanation in CBR, and to everyone who participated in the workshop or otherwise contributed to its success. We would also like to thank the reviewers for the special issue for their thoughtful and timely contributions to the selection process. Finally, our thanks to the journal editor Paul McKevitt and desk editors Eric Pieren, Fiona Routley, and Wendel van der Sluis for their cooperation in this project.

References

Aamodt, A. (1994). Explanation-Driven Case-Based Reasoning. In Wess, S., Althoff, K.-D. & Richter, M. (eds.), *Topics in Case-Based Reasoning*, 274–288. Springer-Verlag: Berlin, Heidelberg.

Aamodt, A. & Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications* **7**: 39–59.

Aha, D. W., Breslow, L. A. & Muñoz Avila, H. (2001). Conversational Case-Based Reasoning. *Applied Intelligence* **14**: 9–32.

Ashley, K. & Rissland, E. (1987). Compare and Contrast: A Test of Expertise. In *Proceedings of the Sixth Annual National Conference on Artificial Intelligence*, 273–278. Morgan Kaufmann: San Mateo, CA.

Bareiss, R. (1989). *Exemplar-Based Knowledge Acquisition: A Unified Approach to Concept Representation, Classification, and Learning*. Academic Press: San Diego.

Barletta, R. & Mark, W. (1988). Explanation-Based Indexing of Cases. In Kolodner, J. (ed.), *Proceedings of a Workshop on Case-Based Reasoning*, 50–60. Morgan Kaufmann: San Mateo, CA.

Branting, L. K. (1988). The Role of Explanation in Reasoning from Legal Precedent. In Kolodner, J. (ed.), *Proceedings of a Workshop on Case-Based Reasoning*, 94–103. Morgan Kaufmann: San Mateo, CA.

Buchanan, B. & Shortliffe, E. (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Addison-Wesley: Reading, MA.

Cunningham, P., Doyle, D. & Loughrey, J. (2003). An Evaluation of the Usefulness of Case-Based Explanation. In Ashley, K. D. & Bridge, D. G. (eds.), *Case-Based Reasoning Research and Development*, 122–130. Springer-Verlag: Berlin, Heidelberg.

DeJong, G. & Mooney, R. (1986). Explanation-Based Learning: An Alternative View. *Machine Learning* **1**: 145–176.

Doyle, D., Cunningham, P., Bridge, D. & Rahman, Y. (2004). Explanation Oriented Retrieval. In González-Calero, P. A. & Funk, P. (eds.), *Advances in Case-Based Reasoning*, 157–168. Springer-Verlag: Berlin, Heidelberg.

Hammond, K. (1989). *Case-Based Planning: Viewing Planning as a Memory Task*. Academic Press: San Diego.

Herlocker, J. L., Konstan, J. A. & Riedl, J. (2000). Explaining Collaborative Filtering Recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, 241–250. ACM Press: New York, NY.

Kolodner, J. & Leake, D. (1996). A Tutorial Introduction to Case-Based Reasoning. In Leake, D. (ed.), *Case-Based Reasoning: Experiences, Lessons & Future Directions*, 31–65. AAAI Press: Menlo Park, CA.

Kolodner, J. & Simpson, R. (1989). The MEDIATOR: Analysis of an Early Case-Based Problem Solver. *Cognitive Science* **13**: 507–549.

Koton, P. (1988). Reasoning about Evidence in Causal Explanations. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, 256–263. Morgan Kaufmann: San Mateo, CA.

Leake, D. (1992). *Evaluating Explanations: A Content Theory*. Lawrence Erlbaum: Hillsdale, NJ.

Leake, D. (1996). CBR in Context: The Present and Future. In Leake, D. (ed.), *Case-Based Reasoning: Experiences, Lessons & Future Directions*, 3–30. AAAI Press: Menlo Park, CA.

Leake, D. (2002). Personalized CBR: Challenges and Illustrations. In Göker, M. & Smyth, B. (eds.), *Proceedings of the ECCBR-02 Workshop on Case-Based Reasoning and Personalization*, 22–23. Robert Gordon University: Aberdeen, Scotland.

Maximini, R., Freßmann, A. & Schaaf, M. (2004). Explanation Service for Complex CBR Applications. In González-Calero, P. A. & Funk, P. (eds.), *Advances in Case-Based Reasoning*, 302–316. Springer-Verlag: Berlin, Heidelberg.

McSherry, D. (2001). Interactive Case-Based Reasoning in Sequential Diagnosis. *Applied Intelligence* **14**: 65–76.

McSherry, D. (2002). Mixed-Initiative Dialogue in Case-Based Reasoning. In Aha, D. W. (ed.), *Proceedings of the ECCBR-02 Workshop on Mixed-Initiative Case-Based Reasoning*, 1–8. Robert Gordon University: Aberdeen, Scotland.

McSherry, D. (2003). Similarity and Compromise. In Ashley, K. D. & Bridge, D. G. (eds.), *Case-Based Reasoning Research and Development*, 291–305. Springer-Verlag: Berlin, Heidelberg.

McSherry, D. (2004). Explaining the Pros and Cons of Conclusions in CBR. In González Calero, P. A. & Funk, P. (eds.), *Advances in Case-Based Reasoning*, 317–330. Springer-Verlag: Berlin, Heidelberg.

Mitchell, T., Keller, R. & Kedar-Cabelli, S. (1986). Explanation-Based Generalization: A Unifying View. *Machine Learning* **1**: 47–80.

Rissland, E., Valcarce, E. & Ashley, K. (1984). Explaining and Arguing with Examples. In *Proceedings of the Fourth National Conference on Artificial Intelligence*, 288–294. Morgan Kaufmann: San Mateo, CA.

Roth-Berghofer, T. R. (2004). Explanations and Case-Based Reasoning: Foundational Issues. In González Calero, P. A. & Funk, P. (eds.), *Advances in Case-Based Reasoning*, 389–403. Springer-Verlag, Berlin, Heidelberg.

Schank, R. C. (1982). *Dynamic Memory: A Theory of Learning in Computers and People*. Cambridge University Press: Cambridge, UK.

Schank, R. & Leake, D. (1989). Creativity and Learning in a Case-Based Explainer. *Artificial Intelligence* **40**: 353–385.

Shimazu, H. (2002). ExpertClerk: A Conversational Case-Based Reasoning Tool for Developing Salesclerk Agents in E-Commerce Webshops. *Artificial Intelligence Review* **18**: 223–244.

Sørmo, F. & Aamodt, A. (2002). Knowledge Communication and CBR. In González-Calero, P. A. (ed.), *Proceedings of the ECCBR-02 Workshop on Case-Based Reasoning for Education and Training*, 47–59. Robert Gordon University: Aberdeen, Scotland.

Southwick, R. W. (1991). Explaining Reasoning: An Overview of Explanation in Knowledge-Based Systems. *Knowledge Engineering Review* **6**: 1–19.

Swartout, W., Paris, C. & Moore, J. (1991). Explanations in Knowledge Systems: Design for Explainable Expert Systems. *IEEE Intelligent Systems*, **6**(3): 58–64.