The Knowledge Engineering Review, Vol. 20:3, 267–269. © 2006, Cambridge University Press
doi:10.1017/S0269888906000579 Printed in the United Kingdom

Soft case-based reasoning

WILLIAM CHEETHAM[1], SIMON SHIU[2] and ROSINA O. WEBER[3]

[1]*General Electric Global Research, 1 Research Circle, Niskayuna, NY 12309, USA;*
e-mail: cheetham@research.ge.com
[2]*Department of Computing, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China;*
e-mail: csckshiu@comp.polyu.edu.hk
[3]*College of Information Science & Technology, Drexel University, Philadelphia, PA 19104, USA;*
e-mail: rosina.weber@drexel.edu

Abstract

The aim of this commentary is to discuss the contribution of soft computing—a consortium of fuzzy logic, neural network theory, evolutionary computing, and probabilistic reasoning—to the development of case-based reasoning (CBR) systems. We will describe how soft computing has been used in case representation, retrieval, adaptation, reuse, and case-base maintenance, and then present a brief summary of six CBR applications that use soft computing techniques.

1 Introduction

Soft computing can be used for building CBR systems that can exploit a tolerance for imprecision, uncertainty, approximate reasoning, and partial truth, in order to achieve tractability, robustness, low solution cost, and closer resemblance to human decision making. Soft computing encourages combining multiple techniques into a hybrid system. A hybrid of soft computing and CBR is now being called *soft CBR* (Pal *et al*., 2001; Pal & Shiu, 2004). Soft computing techniques can be useful in most, if not all, of the CBR cycle. We will describe how soft computing has been used in case representation, retrieval, adaptation, reuse, and case base maintenance.

2 Soft case-based reasoning

Case representation is the task of choosing the conceptual data model(s) that are used to represent cases. The use of fuzzy sets allows a flexible encoding of case characteristics as linguistic terms, fuzzy numbers, and fuzzy objects (Plaza & Lopez de Mantaras, 1990; Bonissone & Ayub, 1992). Case retrieval involves selecting the cases from the case base that are most similar to an input query. During retrieval the fuzzy similarity of a case can be calculated using a fuzzy membership function for each feature that specifies the desired similarity for any possible difference in the feature values (Dubois *et al*., 1988; Bonissone & Cheetham, 1998). Weighted fuzzy pattern matching allows us to take into account the relative level of importance of each attribute in the comparison process (Weber *et al*., 1995), thus limiting the penalty derived from the matching of cases that differ on rather unimportant attributes. Case adaptation transforms the retrieved solution(s) into a more appropriate one for the input query. The adaptation process consists of assessing the difference between the retrieved case and the input query, and applying some pre-compiled adaptation knowledge to modify the old solution(s). Adaptation knowledge could be in the form of a set of fuzzy adaptation rules (Shiu *et al*., 2000), a neural network (Corchado & Lees, 2000), or a fuzzy decision tree (Shiu *et al*., 2001). Reuse is the step of using the retrieved case(s) to construct a solution for the new problem. Soft computing techniques can be used to estimate a confidence value in each solution that is generated by the CBR system (Bonissone & Cheetham, 1998). Case base maintenance refers to the constant updating and maintaining of the knowledge in the CBR system to ensure correct system performance in response to changes in task or environment. All numeric parameters of the CBR system (e.g., feature weights, value of k in k-NN, shape of fuzzy similarity membership functions) can be maintained using a genetic algorithm (Bonissone *et al*., 2002; Craw *et al*., 2001). Inductive methods can also be used to cluster case bases and find representative and redundant cases (Shiu *et al*., 2001), which can be used to direct deletion strategies.

The use of fuzzy logic in CBR systems goes back to the early 1990s, when researchers started to use attributes with fuzzy values and a fuzzy pattern matcher for case retrieval. Below we provide a brief description of a representative sample of fuzzy techniques present in several historical CBR systems. For a more detailed description, see Bonissone & Lopez de Mantaras (1998).

3 Soft CBR applications

The ARC system (Plaza & Lopez de Mantaras, 1990) is a case-based apprentice that learns from fuzzy examples. The memory organization of ARC is a hierarchy of classes and cases. Each class is represented by a fuzzy prototype, which describes the features common to most of the cases belonging to the class. The retrieval step selects the most promising classes by means of a fuzzy pattern-matching algorithm based on common features. The certainties of the fuzzy prototypes and the matching degree are expressed by means of linguistic values.

BOLERO (Lopez & Plaza, 1993) is a system that integrates rule-based and case-based representations of strategic knowledge. The object level knowledge of BOLERO is represented by rules and the meta-knowledge is the solved instances of problems, conveniently organized in the memory of cases. Since these solved instances can contain uncertain and imprecise values, linguistic values represented by fuzzy sets are used. Moreover, the pattern-matching algorithm at the retrieval step is adapted to deal with such linguistic values. An added value of such hybrid systems is the capability of learning meta-knowledge by experience. BOLERO has been successfully applied in a complex medical diagnosis problem for diagnosing pneumonias.

CAREFUL (Jaczynski & Trousse, 1994) uses fuzzy logic for the design of CBR assistants. It has an object-oriented representation based on a hierarchy of fuzzy classes. Fuzzy logic introduces flexibility (e.g., imprecision, incompleteness, preferences) for describing the target, the user requests, the cases, and the retrieval process. As with the other systems, the fuzzy sets represent imprecise values of case attributes. Furthermore, in CAREFUL, fuzzy sets are also used to represent vague constraints on the problem description.

In CARS (Bonissone & Ayub, 1992, 1994), cases and problems are defined in an object-oriented environment. The representation of cases and problems uses surface and abstract fuzzy attributes. The abstract attributes are computed using plausible inference rules. An interesting aspect of CARS is the technique used to select the nearest cases: it uses a fuzzy similarity measure between attributes based on a fuzzy algebra. The obtained similarity, which is a fuzzy set, is compared, by means of a measure of inclusion, to reference fuzzy sets labeled NO-MATCH, PARTIAL-MATCH or COMPLETE-MATCH, which represent the matching degrees for each attribute.

In FLORAN (Salotti, 1992), the cases and problems are represented within an object-oriented environment where cases and problems are instances of classes whose slots can have fuzzy values. The classes of FLORAN are linked to dependency contexts (i.e. objects that represent, for a given class, a specific goal, a list of relevant attributes, their importance, and a set of fuzzy restrictions on the attributes). This is similar to the CAREFUL system except that the contexts in CAREFUL are hierarchically organized.

PROFIT (Bonissone & Cheetham, 1998) is a fuzzy CBR system that was developed to estimate residential property values for real estate transactions. The system enhances CBR techniques with fuzzy predicates expressing preferences for determining similarities between subject and comparable properties. These similarities guide the selection and aggregation process. Fuzzy techniques are also used to generate a confidence value qualifying such estimates. PROFIT has been successfully tested on thousands of real estate transactions.

4 Conclusion

These applications show how soft computing techniques can be used to implement the CBR cycle. Soft CBR appears more often in deployed applications than in research projects because the deployed applications need to address the uncertainty of the real world. The utilization of soft computing techniques produce a CBR system that is robust, tolerant of noise, and more closely resembles the human decision-making process.

References

Bonissone, P and Ayub, S, 1992, Similarity measures for case-based reasoning systems. In *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Majorca, Spain*, pp. 483–487.

Bonissone, P and Ayub, S, 1994, Representing cases and rules in plausible reasoning systems. In *Proceeding of the 1994 ARPA/RL Planning Initiative, Tucson, AZ*, pp. 305–316.

Bonissone, P and Cheetham, W, 1998, Fuzzy case-based reasoning for residential property valuation. *Handbook on Fuzzy Computing*, vol. G 15.1. Oxford: Oxford University Press.

Bonissone, P and Lopez de Mantaras, R, 1998, Fuzzy case-based reasoning systems. *Handbook on Fuzzy Computing*, vol. F 4.3. Oxford: Oxford University Press.

Bonissone, P, Subbu, R and Aggour, KS, 2002, Evolutionary optimization of fuzzy decision systems for automated insurance underwriting. In *Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, Honolulu, Hawaii*, vol. 2, pp. 1003–1008.

Corchado, JM and Lees, B, 2000, Adaptation of cases for case based forecasting with neural network support. In Pal, SK *et al.* (eds) *Soft Computing in Case-Based Reasoning*. London: Springer, pp. 293–320.

Craw, S, Jarmulak, J and Rowe, R, 2001, Maintaining retrieval knowledge in a case-base reasoning system. *Computational Intelligence* **17**(2), 346–363.

Dubois, D, Godo, L, Prade, H and Zapico, A, 1998, Making decision in a qualitative setting: from decision under uncertainty to case-based decision. In *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning, Trento, Italy*, pp. 594–605.

Jaczynski, M and Trousse, B, 1994, Fuzzy logic for the retrieval step of a case-based reasoner. *Proceedings Second European Workshop on Case-Based Reasoning*, pp. 313–322.

Lopez, B and Plaza, E, 1993, Case-based learning of strategic knowledge. In Kodratoff (ed.) *Machine Learning EWSL-91 (Lecture Notes in Artificial Intelligence)*. Berlin: Springer, pp. 398–411.

Pal, SK and Shiu, SCK, 2004, *Foundations of Soft Case-Based Reasoning*. New Jersey: Wiley and Sons.

Pal, SK, Dillon T and Yeung, D (eds), 2000, *Soft Computing in Case Based Reasoning*. London: Springer.

Plaza, E and Lopez de Mantaras, R, 1990, A case-based apprentice that learns from fuzzy examples. In Ras, Zemankova and Emrich (eds) *Methodologies for Intelligent Systems*, vol. 5. Elsevier, pp. 420–427.

Salotti, S, 1992, Filtrage flou et representation centree objet pour raisonner par analogie: le systeme FLORAN, PhD thesis.

Shiu, SCK, Sun, CH, Wang, XZ and Yeung, DS, 2000, Maintaining case-based reasoning systems using fuzzy decision trees. *Proceedings of the 5th European Workshop on Case-based Reasoning—EWCBR2K, Trento, Italy*, pp. 285–296.

Shiu, SCK, Sun, CH, Wang, XZ and Yeung, DS, 2001, Transferring case knowledge to adaptation knowledge: an approach for case-base maintenance. *Computational Intelligence* **17**(2), 295–314.

Weber, R, Barcia, R and Khator, S, 1995, Case-based reasoning for cash flow forecasting using fuzzy retrieval. In Veloso, M and Aamodt, A (eds) *Case-Based Reasoning Research and Development (Lecture Notes in Artificial Intelligence, 1010)*. Berlin: Springer, pp. 510–519.