The Knowledge Engineering Review, Vol. 20:3, 261–265. © 2006, Cambridge University Press
doi:10.1017/S0269888906000683 Printed in the United Kingdom

Distributed case-based reasoning

ENRIC PLAZA[1] and LORRAINE MCGINTY[2]

[1]*IIA—Institut d'Investigació en Intel ligència Artificial, CSIC—Spanish Scientific Research Council, Campus UAB, Bellaterra, Catalonia, Spain;*
e-mail: enrich@iiia.csic.es
[2]*The School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland;*
e-mail: lorraine.mcginty@ucd.ie

Abstract

Distribution of resources within case-based reasoning (CBR) architectures is beneficial in a variety of application contexts. This article briefly discusses some of the approaches that fall under the heading of distributed CBR, and their general impact.

1 Introduction

Conventional case-based reasoning (CBR) systems have taken what could be called a *single agent, single case base* problem-solving approach where one, usually well-maintained, case base functions as the central knowledge resource. Classic examples include 'Entree' (Hammond *et al*., 1997) and 'FAQ-Finder' (Burke *et al*., 1997) (see also related articles Bridge *et al*. (2005) and Cheetham & Watson (2005)). Research efforts in the area of distributed CBR concentrate on the distribution of resources with the intent of improving the performance of CBR systems. Although the phrase *distributed CBR* can be used in a number of different contexts, research efforts in the this area can be categorized by using two key criteria: (1) how knowledge is organized/managed within the system (i.e. single versus multiple case bases), and (2) how knowledge is processed by the system (i.e. single versus multiple processing agents). Figure 1 displays the categories for each of the contributions we reference.

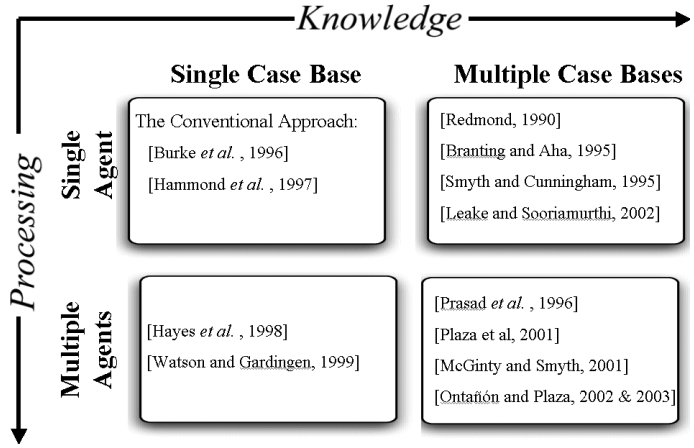


Figure 1 The distributed CBR landscape

https://doi.org/10.1017/S0269888906000683 Published online by Cambridge University Press

2 Distributed CBR approaches

2.1 Single agent/multiple case base

Several researchers have focused on *single-agent* CBR architectures where case experience is distributed across multiple case bases, but where there is a single problem-solving agent with access to these case bases. For example, Branting & Aha (1995) and Smyth & Cunningham (1996) focus on CBR systems with access to multiple, hierarchical case bases that contain problem cases at different levels of abstraction. New problems are solved in a top-down fashion by selecting and combining multiple cases to solve the sub-parts of the target problem; similar approaches are described by Redmond (1990). Taking full advantage of these multiple case bases requires *multi-case-base reasoning* (Leake & Sooriamurthi, 2002), which reasons not only about how to apply cases but also about when and how to draw on particular case bases. Leake & Sooriamurthi (2001) consider the trade-offs of merging individual case bases into a single source versus retaining them individually.

2.2 Multi-agent/single case base

In contrast to single-agent CBR systems, multi-agent systems distribute the case base itself and/or some aspect(s) of the CBR cycle (Retrieve, Reuse, Revise, Retain) among several (more than one) agents. A multi-agent CBR system in this context is a distributed system of cooperative agents, where individual agents draw on the previous experience resource(s) to solve new problems. For instance, it can refer to the situation where a multiple number of problem-solving agents draw on the information held by a single case base (for example, helpdesk applications). Watson & Gardingen (1999a) describe a good example of a fielded distributed CBR application (i.e. 'Really Cool Air') that supports engineering sales staff by taking a *multi-agent/single-case-base* approach. The authors also describe how system efficiency can be further improved by making some adjustments to its architecture and retrieval techniques using introspective reasoning (Watson & Gardingen, 1999b).

Hayes *et al.* (1998) present a distributed architecture motivated by the need to move processing to the client to improve interactive response times for CBR systems where the client dialogue is long lived. They propose that some of the case base processing be distributed to the client once a sufficiently small number of potentially relevant cases have been narrowed down using incremental case retrieval techniques (López de Mántaras *et al.*, 2005). In addition, they introduce CBML (Case Based Markup Language) as a protocol for enabling distributed CBR over any network. This XML format for case representation enables granular updating. Thus, in a distributed CBR environment, servers do not have to dispatch an entire subsection of the case base to the client every time there is a change. The server holds a profile of what has already been sent to the client and only resends the changed element of that dispatch. Moreover, CBML allows centralized case data to be updated quickly.

Systems that rely on using a single-case-base source for problem solving are limited by the scope of this knowledge resource. While this is appropriate in certain application domains (for example, online product catalogues where there is a definite and distinctive range of options), there are many others that would benefit from the improved coverage offered by multiple case bases. Much of the work in distributed CBR assumes multi-case-base architectures involving multiple processing agents (i.e. architectures where problem-solving experience and responsibility is distributed across multiple CBR agents). For this reason, we will next review some of the key contributions in this area.

2.3 Multi-agent/multiple case base

A major theme in *multi-agent, multi-case-base* systems is the *autonomy* of the agents. Two key factors that govern agent autonomy are (1) its capability to determine whether it is competent to

https://doi.org/10.1017/S0269888906000683 Published online by Cambridge University Press

solve a problem, and (2) its capability to interact with other agents (for example, through negotiation, cooperation/collaboration) to obtain a global solution for a given problem. Thus, this section summarizes (and differentiates) related research in this area, firstly, according to how an agent decides *when* to interact with other agents, and secondly, in terms of how to identify *what* agents are best suited to the problem-solving task. For instance, Prasad *et al*. (1996) examine issues pertaining to cooperative retrieval and composition of a case in which sub-cases are distributed across different agents in a multi-agent system. Their *negotiated retrieval* approach assumes that each agent has a partial and imperfect view of the problem-solving situation. This requires agents to cooperatively access their case bases and retrieve the best 'sub-cases', which can be combined in a consistent manner to provide a complete solution. Thus, the CBR-TEAM system (Prasad *et al*., 1996) uses negotiated retrieval in a parametric design task. It assumes a team of heterogenous cooperative agents whereby individual agents are responsible for a particular sub-part of the design task. Agents cooperate to resolve design conflicts as each new design is constructed from the contribution of the distributed agents. However, the negotiated retrieval approach assumes that groups of specialized agents *always* work as a cooperative team when solving new problems. This is in contrast to other research that has looked at distributed, multi-agent frameworks whereby individual agents are capable of directly solving problems that fall within their own area of expertise, but need to draw on the experience of other agents for problems that do not (Plaza & Ontañón, 2001; McGinty & Smyth, 2001).

Several types of cooperation among CBR agents (CoopCBR) have been proposed by Plaza and Ontañón. In *ensemble CBR* (Plaza & Ontañón, 2001) they study the 'ensemble effect' where a collection of agents with uncorrelated case bases improve the accuracy of any individual. The ensemble effect diminishes when individual agents have case bases with a biased sample of the case space, but Ontañón & Plaza (2002) show that agents can barter cases using individual policies that achieve individual and collective improvement in performance. Moreover, the ensemble effect can be improved if the CBR agents can provide a justification of their predictions (Ontañón & Plaza, 2003a); justified predictions can also be used to improve the distributed process of case retention involving several CBR agents, since justifications support an informed bargaining among the agents to determine which agent benefits more from retaining a particular case (Ontañón & Plaza, 2004). Finally, proactive learning can be used for acquiring a *competence model* that allows an agent to learn how to identify which other agents are best suited as cooperation partners (Ontañón & Plaza, 2003b). Competence models can improve both the size and the robustness of the ensembles of CBR agents, since using the most competent agents in an ensemble reduces its size while maintaining the ensemble effect.

McGinty & Smyth (2001) describe the *Collaborative Case-Based Reasoning* (C-CBR) architecture which, like CoopCBR, consists of a collection of homogeneous CBR agents, each having the same CBR capabilities but differing in their problem solving experiences. While CoopCBR addresses classification tasks on distributed data, C-CBR is about distribution in relation to planning tasks. In particular, McGinty & Smyth (2001) demonstrate how C-CBR can be used to solve an important challenge in the area of personalized route planning: how to generate routes that conform to a user's implicit travel preferences in both familiar and unfamiliar map territories (McGinty, 2002; McGinty & Smyth, 2002).

3 Concluding remarks

Distributed CBR strategies can improve both the performance and maintainability of CBR systems. As mentioned earlier, up until recently the vast majority of CBR systems have taken a single-agent, single-case-base approach to problem solving. However, while this approach has proven to be effective in certain application settings, there are two important factors that may render this approach unsuitable: (1) privacy and (2) scalability. *Privacy* refers to the situation where cases are owned by different users, who may not always be willing to give them to a centralized 'case

https://doi.org/10.1017/S0269888906000683 Published online by Cambridge University Press

repository'. *Scalability* concerns the impracticability of processing a centralized case base when dealing with very large amounts of data.

Subsequently, the distributed CBR approach has been proposed as an appropriate alternative, and research efforts in this area have demonstrated a number of additional benefits for favoring this over the more conventional approach. For example, any centralized CBR system will have limited coverage characteristics, so by drawing on extra knowledge that may be available but scattered throughout the system, coverage can be improved as well as the overall competence of the system. Moreover, the more obvious benefit in terms of performance is that efficiency can be improved by distributing the workload. Finally, there are also potential gains in terms of system maintenance since it may be easier to adapt local case bases independently of each other.

References

Branting, K and Aha, DW, 1995, Stratified case-based reasoning: reusing hierarchical problem solving episodes. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, San Francisco, CA: Morgan Kaufmann, pp. 384–390.

Bridge, D, Göker, M, McGinty, L and Smyth, B, 2005, Case-based recommender systems. *The Knowledge Engineering Review* this issue.

Burke, R, Hammond, K, Kulyukin, V, Lytinen, S, Tomuro, N and Schoenberg, S, 1997, Question answering fromfrequently-asked question files: experiences with the FAQ Finder system. *AI Magazine* **18**(2), 57–66.

Cheetham, B and Watson, I, 2005, Fielded applications of CBR. *The Knowledge Engineering Review* this issue.

Hammond, K, Burke, R and Young, B, 1997, The FindMe approach to assisted browsing. *IEEE Expert* **12**(4), 32–40.

Hayes, C, Cunningham, P and Doyle, M, 1998, Distributed CBR using XML. In *Proceedings of the UKCBR Workshop: Intelligent Systems and Electronic Commerce*.

Leake, D and Sooriamurthi, R, 2001, When two case bases are better than one: exploiting multiple case bases. In Aha, DW and Watson, I (eds) *Proceedings of the Fourth International Conference on Case-Based Reasoning*. Berlin: Springer, pp. 321–335.

Leake, D and Sooriamurthi, R, 2002, Automatically selecting strategies for multi case base reasoning. In Craw, S and Preece, A (eds) *Proceedings of the Fifth European Conference on Case-Based Reasoning*. Berlin: Springer, pp. 204–219.

López de Mántaras, R, McSherry , D, Bridge, D, Leake, D, Smyth, B, Craw, S, Faltings, B, Mahar, ML, Cox, M, Forbus, K, Keane, M, Aamodt, A and Watson, I, 2005, Retrieval, reuse and retention in CBR. *The Knowledge Engineering Review* this issue.

McGinty, L and Smyth, B, 2001, Collaborative case-based reasoning: applications in personalized route planning. In Aha, DW and Watson, I (eds) *Proceedings of the Fourth International Conference on Case-Based Reasoning*. Berlin: Springer, pp. 362–376.

McGinty, L and Smyth, B, 2002, Shared experiences in personalized route planning. In Haller, S and Simmons, G (eds) *Proceedings of the Fifteenth International FLAIRS Conference*. San Mateo, CA: AAAI Press, pp. 111–115.

McGinty, L, 2002, Collaborative case-based reasoning: applications in personalized route planning. PhD thesis, Computer Science Department, University College Dublin, Ireland.

Ontañón, S and Plaza, E, 2002, A bartering approach to improve multi-agent learning. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 386–393.

Ontañón, S and Plaza, E, 2003a, Justification-based multi-agent learning. In *International Conference on Machine Learning*. San Francisco, CA: Morgan Kaufmann, pp. 576–583.

Ontañón, S and Plaza, E, 2003b, Learning to form dynamic committees. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 504–511.

Ontañón, S and Plaza, E, 2004, Justification-based case retention. In Gonzalez, PA and Funk, P (eds) *Proceedings of the Seventh Conference on Case-Based Reasoning*. Berlin: Springer, pp. 346–360.

Plaza, E and Ontañón, S, 2001, Ensemble case-based reasoning: collaboration policies for multi-agent cooperative CBR. In Aha, DW and Watson, I (eds) *Proceedings of the Fourth International Conference on Case-Based Reasoning*. Berlin: Springer, pp. 437–451.

Prasad, MVN, Lesser, VR and Lander, SE, 1996, Retrieval and reasoning in distributed case bases. *Journal of Visual Communication and Image Representation, Special Issue on Digital Libraries* **7**(1), 74–87.

Redmond, M, 1990, Distributed cases for case-based reasoning: facilitating use of multiple cases. In Dietterich, T and Swartout, W (eds) *Proceedings of the Eighteenth National Conference on Artificial Intelligence*. San Mateo, CA: AAAI Press, pp. 304–309.

https://doi.org/10.1017/S0269888906000683 Published online by Cambridge University Press

Smyth, B and Cunningham, P, 1996, The utility problem analysed: a case-based reasoning perspective. In Smith, I and Faltings, B (eds) *Proceedings of the Third European Workshop on Case-Based Reasoning*. Berlin: Springer, pp. 392–399.

Watson, I and Gardingen, D, 1999a, A distributed case-based reasoning application for engineering sales support. In Dean, T (ed.) *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, vol. 1. San Francisco, CA: Morgan Kaufmann, pp. 600–605.

Watson, I and Gardingen, D, 1999b, Improving a distributed case-based reasoning system using introspective learning. *Research and Development in Intelligent Systems* **XVI**, 178–190.

https://doi.org/10.1017/S0269888906000683 Published online by Cambridge University Press