

## Retrieval in Case-Based Reasoning: An Overview

Ramon López de Mántaras<sup>1</sup>, Derek Bridge<sup>2</sup>, David McSherry<sup>3</sup>

<sup>1</sup>*Artificial Intelligence Research Institute, CSIC  
Campus UAB, 08193 Bellaterra, Spain  
[mantaras@iiia.csic.es](mailto:mantaras@iiia.csic.es)*

<sup>2</sup>*Computer Science, University College Cork, Cork, Ireland  
[d.bridge@cs.ucc.ie](mailto:d.bridge@cs.ucc.ie)*

<sup>3</sup>*School of Computing and Information Engineering, University of Ulster,  
Coleraine BT52 1SA, Northern Ireland, UK  
[dmg.mcsherry@ulster.ac.uk](mailto:dmg.mcsherry@ulster.ac.uk)*

**Abstract.** An important step in the solution of a target problem in case-based reasoning (CBR) is the retrieval of similar previous cases that can be used to solve the target problem. We review a selection of papers from the CBR literature on aspects of retrieval, such as approaches to the assessment of surface and structural similarity and techniques for automating the construction and maintenance of similarity measures. We also examine a number of retrieval techniques that have been developed to address the limitations of retrieval based purely on similarity.

### 1 Introduction

In case-based reasoning (CBR), problems are solved by adapting the solutions of similar previous cases stored in a case memory. The retrieval of cases that are most likely to be useful in the solution of a target problem relies on an accurate assessment of their similarity to the target problem. Improving retrieval performance through the development of more effective approaches to similarity assessment has been the focus of a considerable amount of research in CBR. While retrieval based purely on similarity remains the most widely used technique, important limitations of similarity-based retrieval have come to light with the emergence of new and more demanding applications of CBR. For this reason, similarity is increasingly being combined with other criteria to guide the retrieval process, such as how effectively the solution space is *covered* by the retrieved cases (McSherry, 2003c), how easily their solutions can be *adapted* to solve the target problem (Smyth & Keane, 1995), or how easily the proposed solution can be *explained* (Doyle *et al.*, 2004).

In Section 2, we examine the distinction between surface and structural similarity and review approaches to the assessment of these different types of similarity. In Section 3, we examine factors that affect retrieval performance and techniques for improving retrieval performance such as the use of machine learning techniques to refine the similarity measures used to guide the retrieval process. In Section 4, we examine new approaches to retrieval that have been developed to address the limitations of similarity-based retrieval. Our concluding remarks are presented in Section 5.

## 2 Similarity Assessment

In some applications of CBR, it may be adequate to assess the similarity of the stored cases in terms of their *surface* features. The surface features of a case are those that are provided as part of its description and are typically represented using attribute-value pairs. In other applications, it may be necessary to use *deep* or *derived* features obtained from the case description by inference based on domain knowledge. In yet other applications, cases are represented by complex structures (such as graphs or first-order terms) and retrieval requires an assessment of their *structural* similarity. As might be expected, the use of deep features or structural similarity is computationally expensive; however, the advantage is that relevant cases are more likely to be retrieved.

### 2.1 Assessment of Surface Similarity

In approaches to retrieval based on surface features, the similarity of each case to the target problem, typically represented as a real number in the range from 0 to 1, is computed according to a given similarity measure. Usually the retrieved cases are the  $k$  cases that are most similar to the target problem, an approach often referred to as “ $k$  nearest neighbour” retrieval or simply  $k$ -NN. Alternatively, the retrieved cases may be those whose similarity to the target problem exceeds a predefined threshold.

There are many ways of measuring similarity and different approaches are appropriate for different case representations. For example, it is common in practice for each case to be represented as a simple feature vector (or set of attribute-value pairs). With this representation, it is usual to define a *local* similarity measure for each attribute. A *global* similarity measure aggregates the degrees of similarity that are computed by the local measures. This is most easily done when the degrees of local similarity are represented by real numbers normalised to some fixed interval such as  $[0,1]$ . Then, global similarity can be computed as a weighted average of the local similarities. The weights allow different attributes to have different importance; they may be assigned by a domain expert or a user, or as discussed in Section 3, they may be determined by an adaptive learning process.

The soundness and completeness of retrieval is ensured by computing the similarity of the target problem to every case in memory. However, this process has complexity  $O(n)$  where  $n$  is the number of cases in memory, which may be an unacceptable overhead if  $n$  is very large. One approach to reducing retrieval time, as exemplified by the pioneering work of Stanfill and Waltz (1986), involves the use of massively parallel computers. While the requirement for expensive hardware is an obvious drawback, the approach still guarantees the soundness and completeness of retrieval by performing an exhaustive memory search. Stanfill and Waltz describe the implementation of a memory-based reasoning algorithm on a fine-grained SIMD parallel machine. Their *Connection Machine* performs a highly parallel search for similar cases and has been applied to the problem of pronouncing English words using a case memory containing thousands of examples of words that are correctly pronounced.

Another approach to reducing retrieval time relies on the organization of cases in memory. For example, Wess, Althoff and Derwand (1993) propose a sound and complete approach to retrieval in which the organization of the case memory is based on similarities *between* cases. A binary tree called a *k-d* tree is used to split the case base into groups of cases in such a way that each group contains cases that are similar to each other according to a given similarity measure. To ensure that the most similar cases are retrieved, the retrieval algorithm computes similarity bounds to determine which groups of cases should be considered first.

Simoudis and Miller (1990) argue that retrieval based only on surface similarity may not be sufficiently discriminating when applied to large case memories, and needs to be combined with other techniques in order to reduce the number of cases to be considered for adaptation. They present an approach called *validated retrieval* that is capable of dramatically reducing the number of potentially relevant cases. Retrieval based on surface similarity is combined in the approach with validation of the retrieved cases to determine if they are applicable to the target problem. Associated with each case in memory is a *validation procedure* consisting of a set of domain-specific tests and their results for that case. In order to validate a retrieved case, the tests associated with the case are applied to the target problem. The retrieved case is considered to be relevant to the target problem only if all the tests give the same results for the target problem as they do for the retrieved case.

The *validation* phase that follows the initial retrieval of cases in validated retrieval resembles the *justification* phase in CASEY, a CBR system for medical diagnosis (Koton, 1988). The goal of CASEY's justification component is to determine whether the causal explanation of a retrieved case applies to the target problem. Often this enables CASEY to avoid invoking its causal model when creating an explanation for the target problem. Other systems that combine retrieval based on surface similarity with an additional filter to improve retrieval performance include CHEF (Hammond, 1986), SWALE (Kass, Leake & Owens, 1986), and KRITIK (Goel & Chandrasekaran, 1989).

## 2.2 Assessment of Structural Similarity

Because it makes extensive use of domain knowledge, retrieval based on structural similarity is computationally expensive, but as previously mentioned, the advantage is that relevant cases are more likely to be retrieved. One way of mitigating the extra cost is to use a two-stage retrieval algorithm. Forbus *et al.* (1994) present a computational model called MAC/FAC in which surface and structural similarity are combined in a two-stage process. The first stage uses a surface matcher to filter out a set of potentially relevant items from long-term memory (those that are most similar to the target problem). The selected candidates are processed in the second stage by a *structure-mapping engine* that computes structural matches between them and the target problem and selects one or more of the best matches. Experiments based on human assessment of similarities and analogies have confirmed that both surface and structural similarity assessment are necessary for sound retrieval (Forbus *et al.*, 1994).

Inspired by previous work by Gentner and Forbus (1991), Börner (1993) proposes an approach to retrieval in which fast retrieval of candidate cases based on their surface similarity to the target problem is followed by a more expensive assessment of their structural similarity. She defines structural similarity as the most specific graph structure that the target problem has in common with a stored case, and a set of transformation rules, given as background knowledge, needed to determine this common structure.

Object-oriented case representations generalise simple attribute-value representations. Cases are represented by sets of objects. Objects belong to classes, which are organised in a class hierarchy. An object's class determines the attributes it may have. Attributes may be *relational*, which means that their values will themselves be further objects. It seems obvious that the class hierarchy must contain useful similarity knowledge. For example, objects that are close to each other in the hierarchy are likely to be more similar than objects that are far apart. However, Bergmann and Stahl (1998) suggest that there is no clear view about how the similarity between objects of different classes should be determined, with the result that the assessment of similarity is often restricted to objects of the same class. To address this issue, they present a new framework for computing object similarities for object-oriented case representations that enables objects of different classes to be compared and takes account of the knowledge that is implicit in the class hierarchy.

Another way of dealing with relations between attributes is based on the concept of *generalised cases*. (Bergmann *et al.*, 1999). A generalised case covers a subspace of the problem-solution space, providing solutions to a set of closely related problems, rather than just a single problem (Mougouie & Bergmann, 2002). Dependencies between attributes are explicitly represented in a way that supports the extension of similarity measures. For example, Bergmann (2002) defines the similarity between a query and a generalised case as the similarity between the query and the most similar case contained in the generalised case. Mougouie and Bergmann (2002) formulate the similarity assessment problem for generalised cases, described by continuous attributes, as a nonlinear programming problem and introduce an optimization-based retrieval method. Tartakovski *et al.* (2004) extend the case representation to support mixed, discrete, and continuous attributes. They also formulate similarity assessment as a special case of a mixed integer nonlinear optimization problem, and propose an optimization-based retrieval method operating on a given index structure.

Bunke and Messmer (1993) propose one of a number of structural similarity measures for domains in which cases are represented as graph structures. Another such proposal (Börner, 1993) was mentioned earlier in our discussion. Bunke and Messmer's proposed measure is based on *graph editing* operations (inserting, deleting, and substituting nodes and edges in the graph). To improve the practical efficiency of the approach, they introduce a *subgraph matching* algorithm that works on a compactification of the case memory in which subgraphs which are common to multiple cases may be stored only once. In a similar vein, Champin and Solnon (2003) propose a similarity measure, based on graph editing

operations within a modification of Tversky's (1977) *contrast model*, to compare cases represented by labelled graphs where vertices and edges can have more than one label. To deal with the intractability of this representation, a heuristic greedy algorithm is proposed.

Arcos and López de Mántaras (1997) describe a retrieval mechanism called *Perspectives* for structured case representations. Cases and degrees of similarity are represented as feature terms, which are equivalent to first-order terms and can also be viewed as directed acyclic graphs labelled by features and values (Plaza, 1995). Their knowledge-intensive approach to retrieval uses a *subsumption* mechanism between the feature terms to obtain an order relation between case descriptions on the basis of a set of user-defined relevant aspects of the target problem. The system is implemented in an object-oriented language (Arcos, 1997) based on feature terms and has been applied to the problem of synthesizing expressive music (Arcos & López de Mántaras, 2001; López de Mántaras & Arcos, 2002).

Emde and Wettschereck (1996) propose an alternative way of measuring the similarity of first-order terms. They also present a generalization of a propositional instance-based learner (distance-weighted *k*-NN) to first-order representations. Issues addressed in the approach, which the authors refer to as *relational* instance-based learning (RIBL), include the generation of cases from the knowledge base, assessment of similarity between arbitrarily complex cases, and estimation of the relevance of predicates and attributes. Empirical results are presented which suggest that RIBL is capable of achieving high levels of classification accuracy in a variety of domains.

### 2.3 Similarity Frameworks

With so many ways of measuring similarity, it is unsurprising that some researchers have looked at similarity in a general way, independent of any specific algorithm. For example, Richter (1992) discusses the notion of similarity in the context of a formal mathematical framework. He describes approaches to modelling similarities with increasing complexity and informativeness, ranging from simple predicates (least informative) to relations and functions (most informative). General forms of distance functions and similarity measures are discussed, including a generalization of Tversky's (1977) *contrast model*. The contrast model is based on a set-theoretic approach that expresses the similarity between objects as a linear combination of their numbers of matching and mismatching features. One limitation of Tversky's model is that all features are assumed to be equally important, whereas Richter's generalization allows different weights to be assigned to features. However, Richter emphasises that to allow for changes in the problem-solving environment, the parameters of a similarity measure should be the result of an adaptive learning process, an idea we explore further in Section 3.

Osborne and Bridge (1996) present another general framework. They distinguish *ordinal* and *cardinal* similarity measures. Ordinal measures use a description of the target problem to induce a partial ordering over the cases in the case memory. No information about the degree of similarity is given; the

cases are merely ordered, with the implication that cases higher in the ordering should be retrieved prior to any that are lower in the ordering. Cardinal measures, on the other hand, are functions that score the cases, returning real numbers to denote degrees of similarity. Osborne and Bridge present a set of operators that allows the flexible and systematic construction of both ordinal and cardinal similarity measures. In later work (e.g. Osborne and Bridge, 1997), the framework is generalised further to accommodate not only similarity measures in which the degree of similarity is denoted numerically (which is most common) but also similarity measures in which the degree of similarity is instead denoted by any value drawn from an ordered set. With this extension, the framework accounts for similarity measures in which the degree of similarity is denoted by common subgraphs (Börner, 1993) or feature terms (e.g. Jantke, 1994; Plaza, 1995).

### **3 Evaluating and Improving Retrieval Performance**

Several techniques for improving the *speed* of retrieval were mentioned in our discussion of similarity assessment in Section 2. Another important aspect of retrieval performance is its impact on *solution quality*. Measures used to evaluate retrieval performance in terms of solution quality are likely to depend on the type of problem-solving task (e.g. classification, recommendation, planning) for which the system is designed. For example, evaluation in terms of *classification accuracy* is possible only if the outcome classes to be predicted in the *test set* are represented in the *training set*. This is not the case in domains such as product recommendation in which each outcome class (a unique product or service) is represented by a single case in the case memory (McSherry, 2001a). Evaluation of retrieval performance in terms of classification accuracy is similarly compromised in *conversational CBR* (Aha *et al.*, 2001), where it is typical for most cases to have unique solutions. Appropriate measures of retrieval performance for datasets of this type include precision, recall, and the average length of problem-solving dialogues (Aha *et al.* 2001; McSherry, 2001a, 2001b, 2003b).

Problems likely to affect retrieval performance in terms of solution quality include the use of inadequate similarity measures, missing values in cases, unknown values in the description of the target problem, and the so-called *heterogeneity* problem that arises when different attributes are used to describe different cases (Aha *et al.*, 2001; Bogaerts & Leake, 1994; McSherry, 2001b, 2003c; Stahl and Gabel, 2003). Bogaerts and Leake (2004) propose and evaluate a variety of possible strategies for handling missing information in similarity assessment. Retrieval based on incomplete information is an important challenge in conversational CBR, where a description of the target problem is incrementally (and often incompletely) elicited in an interactive dialogue with the user. Aha *et al.* (2001) evaluate an approach to incremental query elicitation that takes account of the heterogeneity that is typically found in domains such as fault diagnosis. McSherry (2003b) proposes a conversational CBR approach to product recommendation that includes a mechanism for ensuring that the dialogue is terminated only when it is certain that a more similar case will not be found if the dialogue is allowed to continue.

Retrieval performance can often be improved by making the similarity measure the subject of an adaptive learning process. Focusing on variants of  $k$ -NN that automatically learn the weights assigned to features, Wettschereck and Aha (1995) propose a multi-dimensional framework for the categorization and comparison of feature weighting methods in CBR. The proposed framework can be used to categorise new methods, thus facilitating their comparison with existing methods. However, it cannot be applied to  $k$ -NN methods that incorporate domain-specific knowledge and complex representations. Noting that most feature weighting methods are designed to optimise classification accuracy, Wilke and Bergmann (1996) argue that decision *costs* should also be considered in many applications. Experimental results are presented which support the hypothesis that classification based on weights learned using cost optimisation leads to lower decision costs than classification based on weights learned by accuracy optimisation.

Improving the adaptability of retrieved cases can also be the subject of an adaptive learning process. For example, working in the field of case-based planning, Muñoz-Avila and Hüllen (1996) extend the *foot-printed* similarity metric used in PRODIGY/ ANALOGY (Velo, 1992) by incorporating feature weights in a new metric which counts the weights of relevant features that match features in the target problem. A feature is considered relevant to a planning goal with respect to a solution if it contributes to achieving the goal in the solution. The authors also present an algorithm for analyzing the performance of retrieved cases to identify features whose weights need to be recomputed. The algorithm provides a bridge between the new similarity metric and a feature weighting model based on incremental optimisers. Experimental results are presented which show that integration of the proposed similarity metric and analysis algorithm in the feature weighting model improves the adaptability of the retrieved cases by convergence to best weights over a period of multiple problem-solving episodes.

Many CBR applications rely on domain knowledge encoded in the similarity measures used by the system to guide the retrieval of relevant cases. Such a *knowledge-intensive* approach to similarity assessment typically relies on knowledge acquired from a domain expert. In a recent series of papers, Stahl and Gabel propose the use of machine learning techniques to reduce the knowledge-acquisition overheads associated with the construction and maintenance of domain-specific similarity measures (Stahl, 2001, 2002; Stahl & Gabel, 2003; Gabel & Stahl, 2004). A distinctive feature of their approach is the use of feedback about the quality of retrieval results provided by a domain expert to guide the automatic refinement of similarity measures.

#### **4 Alternatives to Similarity-Based Retrieval**

Much of the research on which we comment in this section has been motivated by an increased awareness of the limitations of retrieval based purely on similarity. Though continuing to play a prominent role in most approaches to retrieval, similarity is increasingly being combined with other criteria to guide the retrieval process more effectively.

#### 4.1 Adaptation-Guided Retrieval

While many factors may contribute to the performance of the retrieval component of a CBR system, what matters most is whether the retrieved cases can be used to solve the target problem. The issue is not finding cases that are similar *per se*; it is a question of finding cases that are *usefully* similar. This view is perhaps most formally expressed in (Bergmann *et al.*, 2001), where it is explicitly stated that problem similarity is used as a proxy for solution *utility*. This raises the question of what to do when similarity is a poor proxy for utility.

For example, Smyth and Keane (1995) question the assumption that the most similar case is the one that is easiest to adapt. In fact, the most similar case may even be impossible to adapt. To address this issue, they introduce the notion of *adaptation-guided retrieval* in which the adaptation requirements of cases are assessed during retrieval by means of domain-specific adaptation knowledge. In contrast to previous approaches that relied on heuristics to predict the ease with which a given case could be adapted, adaptation-guided retrieval combines local and global measures of adaptability to ensure that the most adaptable case is always selected. Empirical results are presented which show that the approach can also reduce adaptation costs by performing preliminary adaptation work during retrieval.

#### 4.2 Diversity-Conscious Retrieval

In CBR approaches to product recommendation, descriptions of the available products are stored as cases in a case memory and retrieved in response to a *query* representing the user's requirements. An important advantage of the approach is that if there is no case that exactly matches the user's requirements, she can be shown the cases that are *most similar* to her query. However, one limitation of the approach is that the most similar cases are often very similar to each other, with the result that the user is offered a very limited choice. Recently there has been considerable research interest in algorithms that combine measures of similarity and *diversity* in the retrieval process to achieve a better balance between these often conflicting characteristics of the retrieved cases (e.g. McSherry, 2002; Smyth & McClave, 2001). Experimental results have shown that major gains in diversity can often be achieved at the expense of relatively small reductions in similarity.

#### 4.3 Compromise-Driven Retrieval

McSherry (2003c) proposes a *compromise-driven* approach to retrieval in recommender systems inspired by the observation that the cases that are most similar to the user's query are often not sufficiently representative of compromises that the user may be prepared to accept. For example, if the *k* most similar products are made by a manufacturer that the user is not prepared to accept, then the system has failed to deliver an acceptable recommendation. An underlying assumption of similarity-based retrieval (or *k*-NN) is that a given case (or product) is more acceptable than another if it is more similar to the user's query. Compromise-driven retrieval is based on the weaker assumption that a

given case is more acceptable than another if it is more similar to the user's query *and* it involves a subset of the compromises that the other case involves. As well as being less likely to be contradicted by user behaviour, this weaker assumption provides the basis of a more principled approach to deciding which cases are included in the retrieval set than setting an arbitrary threshold on the number of retrieved cases as in  $k$ -NN. For example, no case is included in the retrieval set if there is a more similar case that involves a subset of the compromises it involves.

Though not relying on diversity as an additional measure of recommendation quality, compromise-driven retrieval shares with diversity-conscious retrieval (McSherry, 2002; Smyth & McClave, 2001) the aim of offering users a better *choice* of alternatives. Another important benefit of the approach is that the retrieval set is guaranteed to provide full *coverage* of the available cases in the sense that for any case that is not included in the retrieval set, one of the recommended cases is at least as good in terms of its similarity to the user's query and the compromises it involves. While the size of the retrieval set required to provide full coverage cannot be predicted in advance, experimental results suggest that retrieval-set sizes tend to remain within reasonable limits even for queries of realistic complexity (McSherry, 2003c).

#### 4.4 Order-Based Retrieval

Order-based retrieval is another new approach with particular application to recommender systems (Bridge & Ferguson 2002a). Rather than scoring the cases, order-based retrieval offers an expressive query language for defining and combining ordering relations; the result of query evaluation is to partially order the cases in the case base. The claims made for this new approach include: it is more expressive than similarity-based retrieval because it allows queries that naturally combine not just the user's preferred value (one for which similar values are sought) but also dispreferred values, minimum values, maximum values, and so on; and it returns inherently diverse result sets (Bridge & Ferguson 2002b).

#### 4.5 Explanation-Oriented Retrieval

It is important that CBR systems be able to explain their reasoning (e.g. Cunningham *et al.*, 2003; Doyle *et al.*, 2004; Leake, 1996; McSherry, 2004a). Explanations serve many different *goals*, such as teaching the user about the domain or explaining the relevance of a question the user is asked (Sørmo & Cassens, 2004). For example, McSherry (2003b, 2004b) proposes a conversational CBR approach to product recommendation in which the system can explain why a question is being asked in terms of its ability to discriminate between competing cases. Explaining the *retrieval failures* that occur when no case exactly matches the user's query is another potential role of explanation in CBR recommender systems (e.g. McSherry, 2003a).

More commonly, the goal is to explain how the system reached its conclusions. In applications such as classification and diagnosis, an attractive feature of CBR is the ability to explain the predicted outcome by showing the user one or more of the target problem's nearest neighbours. As noted by Leake (1996), "... the results of CBR systems are based on actual prior cases that can be presented to the user to provide compelling support for the system's conclusions". Such explanations are known as *precedent-based* explanations. An empirical study by Cunningham *et al.* (2003) has shown that they are often more compelling than alternative forms of explanation. However, a number of authors have recently questioned the effectiveness of precedent-based explanations in which the user is simply shown the case that is most similar to the target problem.

For example, McSherry (2004a) argues that such explanations are often less informative than might be expected, and should ideally supported by an analysis of the *pros* and *cons* of the proposed solution. Doyle *et al.* (2004) have found that the most compelling explanation case may not necessarily be the one that is most similar to the target problem. Their realisation that a case that lies between the target problem and the decision boundary may often be more useful for explanation motivated the development of an approach to retrieval which they refer to as *explanation-oriented* retrieval. The approach is still precedent-based: once a classification or diagnosis has been reached on the basis of the nearest neighbours, the system performs an additional retrieval step, using an explanation utility metric, to obtain the explanation case. Doyle *et al.* (2004) also report the results of an empirical study that show their explanation cases to be generally more compelling than the nearest neighbour.

## 5. Concluding Remarks

Our aim in this paper has been to provide a concise overview of retrieval in CBR, rather than a comprehensive survey, by commenting on a representative selection of papers from the CBR literature over the past couple of decades. We have tried to strike a balance between research that can be seen as laying the foundations of retrieval in CBR and more recent contributions. The fact that a considerable number of the selected papers have been published in the last few years is evidence of a significant amount of ongoing research activity. It will be clear from our discussion in Section 4 that much of this research has been motivated by an increased awareness of the limitations of traditional approaches to retrieval. This is a trend that seems likely to continue with the emergence of new and more demanding applications of CBR, and we look forward to the challenges and opportunities that lie ahead.

## References

- D.W. Aha, L.A. Breslow, and H. Muñoz-Avila (2001), Conversational Case-Based Reasoning. Applied Intelligence 14 (1), 9-32.
- J.L. Arcos (1997), The NOOS Representation Language. PhD Thesis, Artificial Intelligence Research Institute and Technical University of Catalonia. Published in the IIIA-CSIC Monographies Series, Bellaterra, Barcelona, Spain.
- J.L. Arcos and R. López de Mántaras (1997), Perspectives: A Declarative Bias Mechanism for Case

- Retrieval. Proceedings of the 2<sup>nd</sup> International Conference on Case-Based Reasoning (Eds. D. B. Leake and E. Plaza), Springer-Verlag, 279-290.
- J.L. Arcos and R. López de Mántaras (2001), An Interactive Case-Based Reasoning Approach for Generating Expressive Music. *Applied Intelligence* 14 (1), 115-129.
- R. Bergmann (2002), *Experience Management*. Springer-Verlag.
- R. Bergmann, M.M. Richter, S. Schmitt, A. Stahl, and I. Vollrath (2001), Utility-oriented Matching: A new research direction for case-based reasoning. Proceedings of the 9<sup>th</sup> German Workshop on Case-Based Reasoning, [http://wm2001.aifb.uni-karlsruhe.de/ws/programm\\_detail\\_ws4\\_frame.htm](http://wm2001.aifb.uni-karlsruhe.de/ws/programm_detail_ws4_frame.htm)
- R. Bergmann and A. Stahl (1998), Similarity Measures for Object-Oriented Case Representations. Proceedings of the 4<sup>th</sup> European Workshop on Case-Based Reasoning (Eds. B. Smyth and P. Cunningham), Springer-Verlag, 25-36.
- R. Bergmann, I. Vollrath, and T. Wahlmann (1999), Generalized Cases and their Application to Electronic Design. Proceedings of the 7<sup>th</sup> German Workshop on Case-Based Reasoning (Ed. E. Melis).
- S. Bogaerts and D. Leake (2004), Facilitating CBR for Incompletely-Described Cases: Distance Metrics for Partial Problem Descriptions. Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning (Eds. P. Funk and Pedro A. González-Calero), Springer-Verlag, 62-76.
- K. Börner (1993), Structural Similarity as Guidance in Case-Based Design. Proceedings of the 1<sup>st</sup> European Workshop on Case-Based Reasoning (Eds. S. Wess, K-D. Althoff, and M.M. Richter), Springer-Verlag, 197-208}.
- D.G. Bridge and A. Ferguson (2002a), An Expressive Query Language for Product Recommender Systems. *Artificial Intelligence Review* 18 (3-4), 269-307.
- D. Bridge and A. Ferguson (2002b), Diverse Product Recommendations using an Expressive Language for Case Retrieval. Proceedings of the 6<sup>th</sup> European Conference on Case-Based Reasoning (Eds. S. Craw and A. Preece), Springer-Verlag, 43-57.
- H. Bunke and B.T. Messmer (1993), Structural Similarity as Guidance in Case-Based Design. Proceedings of the 1<sup>st</sup> European Workshop on Case-Based Reasoning (Eds. S. Wess, K-D. Althoff, and M.M. Richter), Springer-Verlag, 106-118.
- P.A. Champin and C. Solnon (2003), Measuring the Similarity of Labeled Graphs. Proceedings of the 5<sup>th</sup> International Conference on Case-Based Reasoning (Eds. K.D. Ashley and D.G. Bridge), Springer-Verlag, 80-95.
- P. Cunningham, D. Doyle and J. Loughrey (2003), An Evaluation of the Usefulness of Case-Based Explanation. Proceedings of the 5<sup>th</sup> International Conference on Case-Based Reasoning (Eds. K.D. Ashley and D.G. Bridge). Springer, 122-130.
- D. Doyle, P. Cunningham, D. Bridge, and Y. Rahman (2004), Explanation Oriented Retrieval. Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning (Eds. P. Funk and P.A. González-Calero). Springer, 157-168.
- W. Emde and D. Wettschereck (1996), Relational Instance-Based Learning. Proceedings of the 13<sup>th</sup> International Conference on Machine Learning (Ed. L. Saitta), Morgan Kaufmann, 122-130.
- K.D. Forbus, D. Gentner, and K. Law (1994), MAC/FAC: A Model of Similarity-based Retrieval, *Cognitive Science* 19 (2), 141-205.
- T. Gabel and A. Stahl (2004), Exploiting Background Knowledge when Learning Similarity Measures. Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning (Eds. P. Funk and Pedro A. González-Calero), Springer-Verlag, 169-183.

- D. Gentner and K.D. Forbus (1991), MAC/FAC: A Model of Similarity-based Retrieval. Proceedings of the 13<sup>th</sup> Cognitive Science Conference, Erlbaum. Hillsdale, 504-509.
- A. Goel and B. Chandrasekaran (1989), Use of Device Models in Adaptation of Design Cases. Proceedings of the 5<sup>th</sup> National Conference on Artificial Intelligence, Morgan Kaufmann, 100-109.
- K. J. Hammond (1986), CHEF: A Model of Case-based Planning. Proceedings of the 5<sup>th</sup> National Conference on Artificial Intelligence, AAAI Press, 267—271.
- K.P. Jantke (1994), Non-standard Concepts in Similarity in Case-Based Reasoning. Proceedings der Jahrestagung der Gesellschaft für Klassifikation (Eds. H.H. Block *et al.*), Springer-Verlag, 29-44.
- A. Kass, D. Leake, and C. Owens (1986), SWALE: A program that explains. In: Explanation Patterns: Understanding Mechanically and Creatively (Ed. R. Schank), Lawrence Erlbaum Assoc. 232-254.
- P. Koton (1988), Reasoning about Evidence in Causal Explanations. Proceedings of the 7<sup>th</sup> National Conference on Artificial Intelligence. AAAI Press, 256-261.
- D.B. Leake (1996), CBR in Context: The Present and Future, Case-Based Reasoning: Experiences, Lessons and Future Directions (Ed. D.B. Leake). MIT Press, 3-30.
- R. López de Mántaras, and J.L. Arcos (2002), AI and Music: From Composition to Expressive Performance. AI Magazine 23 (3), 43-57.
- D. McSherry (2001a), Minimizing Dialog Length in Interactive Case-Based Reasoning. Proceedings of the 17<sup>th</sup> International Joint Conference on Artificial Intelligence (Ed. B. Nebel), Morgan Kaufmann, 993-998.
- D. McSherry (2001b), Precision and Recall in Interactive Case-Based Reasoning. Proceedings of the 4<sup>th</sup> International Conference on Case-Based Reasoning (Eds. D.W. Aha and I. Watson, Springer-Verlag, 392-406.
- D. McSherry (2002), Diversity-Conscious Retrieval. Proceedings of the 6<sup>th</sup> European Conference on Case-Based Reasoning (Eds. S. Craw and A. Preece) Springer-Verlag, 219-233.
- D. McSherry (2003a), Explanation of Retrieval Mismatches in Recommender System Dialogues. Proceedings of the Workshop on Mixed Initiative Case-Based Reasoning, Workshop Programme, Fifth International Conference on Case-Based Reasoning (Ed. L. McGinty), Norwegian University of Science and Technology, 191-199.
- D. McSherry (2003b), Increasing Dialogue Efficiency in Case-Based Reasoning without Loss of Solution Quality. Proceedings of the 18<sup>th</sup> International Joint Conference on Artificial Intelligence, Morgan Kaufmann (Eds. G. Gottlob and T. Walsh), 121-126.
- D. McSherry (2003c), Similarity and Compromise. Proceedings of the 5<sup>th</sup> International Conference on Case-Based Reasoning (Eds. K.D. Ashley and D.G. Bridge), Springer-Verlag, 291-305.
- D. McSherry (2004a), Explaining the Pros and Cons of Conclusions in CBR, Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning, (Eds. P. Funk and Pedro A. González-Calero), Springer-Verlag, 317-330.
- D. McSherry (2004b), Explanation in Recommender Systems, Proceedings of the Workshop on Explanation in Case-Based Reasoning, Workshop Programme, Seventh European Conference on Case-Based Reasoning (Eds. P. Gervás and K. M. Gupta), Universidad Complutense Madrid, 125-134.
- B. Mougouie and R. Bergmann (2002), Similarity Assessment for Generalized cases by Optimization Methods. Proceedings of the 6<sup>th</sup> European Conference on Case-Based Reasoning (Eds. S. Craw and A. Preece), Springer-Verlag, 249-263.
- H. Muñoz-Avila and J. Hüllen (1996), Feature Weighting by Explaining Case-Based Planning

- Episodes. Proceedings of the 3<sup>rd</sup> European Workshop on Case-Based Reasoning (Eds. I. Smith and B. Faltings), Springer-Verlag, 280-294.
- H.R. Osborne and D.G. Bridge (1996), A Case Base Similarity Framework. Proceedings of the 3<sup>rd</sup> European Workshop on Case-Based Reasoning (Eds. I. Smith and B. Faltings), Springer-Verlag, 309-323.
- H. Osborne and D.G. Bridge (1997), Similarity Metrics: A Formal Unification of Cardinal and Non-Cardinal Similarity Measures. Proceedings of the 2<sup>nd</sup> International Conference on Case-Based Reasoning (Eds. D. B. Leake and E. Plaza), Springer-Verlag, 235-244.
- E. Plaza (1995), Cases as Terms: A feature term approach to the structured representation of cases. Proceedings of the 1<sup>st</sup> International Conference on Case-Based Reasoning (Eds. M. Veloso and A. Aamodt), Springer-Verlag, 265-276.
- M.M. Richter (1992), Classification and Learning of Similarity Measures. Proceedings der Jahrestagung der Gesellschaft für Klassifikation (Eds. Opitz, Lausen, and Klar), Springer-Verlag, 1-8.
- E. Simoudis and J. Miller (1990), Validated Retrieval in Case-Based Reasoning. Proceedings of the 8<sup>th</sup> National Conference on Artificial Intelligence, AAAI Press, 310-315.
- B. Smyth and M.T. Keane (1995), Experiments on Adaptation-Guided Retrieval in Case-Based Design. Proceedings of the 1<sup>st</sup> International Conference on Case-Based Reasoning (Eds. M. Veloso and A. Aamodt), Springer-Verlag, 313-324.
- B. Smyth and P. McClave (2001), Similarity vs. Diversity. Proceedings of the 4<sup>th</sup> International Conference on Case-Based Reasoning (Eds. D.W. Aha, and I. Watson, Springer-Verlag, 347-361.
- F. Sørmo and J. Cassens (2004), Explanation Goals in Case-Based Reasoning. Proceedings of the Workshop on Explanation in Case-Based Reasoning, Workshop Programme, Seventh European Conference on Case-Based Reasoning (Eds. P. Gervás and K. M. Gupta), Universidad Complutense Madrid, 165-174.
- A. Stahl (2001), Learning Feature Weights from Case Order Feedback. Proceedings of the 4<sup>th</sup> International Conference on Case-Based Reasoning, Springer-Verlag (Eds. D.W. Aha and I. Watson), Springer-Verlag, 502-516.
- A. Stahl (2002), Defining Similarity Measures: Top-Down vs. Bottom-Up. Proceedings of the 6<sup>th</sup> European Conference on Case-Based Reasoning (Eds. S. Craw and A. Preece) Springer-Verlag, 406-420.
- A. Stahl and T. Gabel (2003), Using Evolution Programs to Learn Local Similarity Measures. Proceedings of the 5<sup>th</sup> International Conference on Case-Based Reasoning (Eds. K.D. Ashley and D.G. Bridge), Springer-Verlag, 537-551.
- T. Gabel and A. Stahl (2004), Exploiting Background Knowledge when Learning Similarity Measures. Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning (Eds. P. Funk and Pedro A. González-Calero), Springer-Verlag, 169-183.
- C. Stanfill and D. Waltz (1986), Toward Memory-Based Reasoning, Communications of the ACM 29 (12), 1213-1228.
- A. Tartakovski, M. Schaaf, R. Maximini, and R. Bergmann (2004), MINLP Based Retrieval of Generalized Cases. Proceedings of the 7<sup>th</sup> European Conference on Case-Based Reasoning (Eds. P. Funk and Pedro A. González-Calero), Springer-Verlag, 404-418.
- A. Tversky (1977), Features of Similarity. Psychological Review 84 (4), 327-352.
- M. Veloso (1992), Learning by Analogical Reasoning in General Problem Solving. PhD Thesis, Carnegie Mellon University, USA.

S. Wess, K-D. Althoff, and G. Derwand (1993), Using K-D Trees to Improve the Retrieval Step in Case-Based Reasoning. Proceedings of the 1<sup>st</sup> European Workshop on Case-Based Reasoning (Eds. S. Wess, K-D. Althoff, and M.M. Richter), Springer-Verlag, 167-181.

D. Wettschereck and D. Aha (1995), Weighting Features. Proceedings of the 1<sup>st</sup> International Conference on Case-Based Reasoning (Eds. M. Veloso and A. Aamodt), Springer-Verlag, 347-358.

W. Wilke and R. Bergmann (1996), Considering Decision Cost During Learning of Feature Weights. Proceedings of the 3<sup>rd</sup> European Workshop on Case-Based Reasoning (Eds. I. Smith and B. Faltings), Springer-Verlag, 460-472.