rules with much too narrow a scope to be applicable in all but a few retrieval circumstances. Additionally it takes no account of providing a learning mechanism whereby the knowledge can be maintained over time. Furthermore to rely on the expert for adaptation knowledge results inevitably in biased knowledge, in that not all experts solve problems in the same way and often there is much disagreement as to how a problem is best solved. Ideally what is required is an unbiased, automated, flexible, dynamic and intelligent method of acquiring and maintaining *locally* specific adaptation knowledge (and indeed all knowledge in CBR). This perspective was proposed initially by Patterson et al [Patterson et al 1999] where a framework was put forward to develop a Case-Based Reasoning Knowledge Base Management System which would assist in the automated acquisition and maintenance of knowledge for *all* the CBR containers. It is recognised that one knowledge type can often be converted into another in CBR [Richter 1995]. What we propose here is to use the knowledge present within case knowledge, coupled with data mining techniques, to discover and apply locally relevant adaptation knowledge independently from the expert. Leake [Leake 1996] has already recognised the potential of data mining techniques as a useful complement to CBR processes. Anand et al [Anand et al 1998a, Anand et al 1998b] and Patterson et al [Patterson et al 2000, Patterson et al 2002] have empirically shown them to be useful in acquiring and maintaining knowledge for various CBR processes such as similarity and retrieval.

A number of approaches, designed to automatically acquire flexible adaptation knowledge from case-knowledge, have been presented in the literature. For example Hanney [Hanney & Keane 1996] automatically discovered adaptation knowledge in the form of rules by generalising case knowledge. A disadvantage of this approach is that it moves adaptation knowledge maintenance back in the direction of a rule base and all the associated maintenance problems this introduces. McSherry [McSherry 1999a, McSherry 1999b], in order to discover adaptation rules to adapt a retrieved case to a target, discovered what he termed as an *adaptation triplet.* This is an engaging approach due to its implicit simplicity which does discover and use specific adaptation knowledge during adaptation but, unfortunately no empirical studies have been carried out using real world case-bases to validate its competency fully. Leake [Leake et al 1996] proposed a hybrid rule based/case-based approach to adaptation knowledge discovery and reuse. This approach is novel in that they proposed a *case-based* adaptation knowledge container, which although it requires a small set of rules to be engineered into the system to begin with, it learns, stores and reuses new adaptation knowledge in the form of cases during the lifetime of the system. This approach is attractive in that adaptation knowledge is stored as cases and therefore locally relevant adaptation knowledge is available to solve problems but the newly discovered knowledge must be maintained over time and this can be a burden on the system. Lazy learning algorithms such as the nearest neighbour [Cover 1967] have also been used to adapt cases in CBR [Patterson et al 2000]. These generalise the solution through the variation of the parameter k (where k is the number of cases retrieved from an case-base). Using this as a basis for adaptation initially the most similar case (the primary case) is retrieved and then its solution is generalised (adapted) by modifying it as determined by the solutions of the other (k-1) cases (secondary cases) retrieved, based on a voting scheme. The contribution of the secondary cases solutions to the overall target solution is determined by their individual similarity to the target. For example, the k nearest neighbours could be assigned votes which are inversely proportional to their distances from the target [Dudani 1975], [Aha 1990]. The adaptation technique proposed here is a hybrid system, which combines knowledge from the nearest neighbour approach (lazy learning), with regression (eager learning), to dynamically discover specific and *locally relevant* adaptation knowledge, to adapt retrieved cases in real time. By locally relevant we mean the adaptation knowledge used is uniquely specific to the target problem in question. The competency of this approach is compared to k-nearest neighbour adaptation and shown to produce consistently better results over all four case-bases studied. The rest of the paper is organised as follows. The next section describes the hybrid adaptation strategy in detail. This is followed by a section which outlines the experiments implemented, and a section which describes and analyses the results obtained. Finally, in the last section conclusions are drawn and future research outlined.

Adaptation Strategy

The technique used for case adaptation was a hybrid approach based on the k nearest neighbour algorithm and regression analysis. The details of the approach are as follows. Ten fold cross validation [Kohavi 1995] was carried out within each case-base. Initially the k nearest (most relevant) neighbours to the target case were retrieved using the nearest neighbour algorithm. K was calculated as a percentage size of the whole case-base and was at least 30 cases (30 was experimentally found to be the minimum number of cases necessary to produce an accurate regression function). A generalised case was formed from the k cases. The input and output attribute values of the generalised case were created by combining the individual attributes of the k cases after inversely weighting them according to their Euclidean distance from the target case. Using the k nearest neighbours, a linear regression function was determined using standard methods which predicted the difference in the output attributes between two cases **C**($x1,x2,..,xn$) and **C**$'(x1',x2'..xn')$ based on a difference matrix where each entry was based on the difference between two case's attribute values[1] i.e. a function of the form shown in equation 1 was devised.

$$(y'- y) = a_1(x1'-x1) + a_2 (x2'-x2) + \ldots + a_n(xn'-xn \quad ..1)$$

Where n is the number of input attributes ,
$a_1..a_n$ are the regression coefficients and
$y'-y$ is the difference in output attributes between 2 cases.

Using the regression function the difference between the target attribute's solution field and the generalised case's solution field was then predicted ($y'-y$), using the differences in each input attribute for the target and generalised case as inputs to the function. This difference value was then added to the generalised case's output attribute field to give a final prediction to the target case's output attribute field. The mean absolute error (MAE) of the adaptation process was determined by comparing this predicted output value with the actual output value for the target case, and gave a measure of the competency of the system. It was hoped that this technique would provide competent adaptations as local regression has been shown to provide accurate approximations for diverse regression surfaces in machine learning [Torgo 2000].

Methodology

Four case-bases were used to evaluate the technique. As we were using a regression approach to adaptation case-bases were chosen which had continuous output fields.

House 1 case-base - consisted of 565 cases and ten attributes taken from a housing domain supplied by the Valuation and Lands Agency of Northern Ireland. Of these ten attributes five were numeric and five were categorical. The goal was to build a model for predicting house price.

House 2 case-base - consisted of 584 cases described using ten attributes taken from a different housing domain. Of these ten attributes eight were numeric and two were categorical. The goal was to build a model for predicting house price.

House 3 case-base obtained from the ML repository It consisted of 506 cases and 14 features, 12 of which were continuous, 1 was binary and one was categorical. The goal was to build a model for predicting house price.

Abalone case-base was obtained from the ML repository. It consisted of 4177 cases described using 8 fields. Of these 6 were continuous one was categorical and one was integer. The goal was to determine the age of the animal.

The object of the experiments was to investigate the competency of implementing the modified regression analysis as a method of automatically discovering locally specific adaptation knowledge to improve the adaptation process in CBR. A number of different experiments were carried out. Firstly the competency of using knn (k=5) to adapt cases was investigated (5nn). This was followed by using the proposed regression approach to adaptation (knnReg). As a large number of cases were retrieved and used in forming the regression function it was decided to investigate how using the same cases with knn alone would affect the competency of adaptation (knn Max). Feature irrelevancy and feature weighting are widely known to affect the competency of the nearest neighbour approach. Therefore the effects of feature subset selection (FSS) and feature weight optimisation were also investigated. This was a two-stage process whereby a genetic algorithm was initially used to reduce the feature subset and then to optimise the remaining feature weights to a value between 0 and 1. The effects of this on both 5nn (5nnOp) and the regression approach (knnRegOp) were investigated. As the regression approach inherently applies weights to features it is not expected that this will directly affect this approach greatly whereas it should improve the 5nn approach.

Results & Discussion

Figures 1 to 4 show how the MAE varies for each of the adaptation strategies across each of the four case-bases studied. From this it can be seen that when the number of cases considered for knn adaptation is large (knn Max) the results are poor. This is due to many irrelevant cases influencing the final solution. Knn Max always produced the least competent results of the five strategies investigated. Adaptation using 5nn produced significantly more competent solutions then knn Max across all case-bases. In every experiment this was the fourth most competent adaptation strategy. Feature optimisation through a process of FSS and feature weighting improved the 5nn strategy further. For three of the four case-bases 5nnOp proved to be the third most competent technique and once (House 3) it proved to be the second most competent technique, clearly demonstrating how feature optimisation improves the competency of the nearest neighbour approach. KnnRegOp proved to be the second most competent strategy with 3 of the case-bases and once (house 3) it proved to be the most competent. Finally the most competent strategy was knnReg. This provided the best results with 3 case-bases and once (house 3) it was third most competent behind knnRegOp and knn respectively. From this it can be seen that the hybrid adaptation strategy (knnReg or knnRegOp) is always more competent than knn (knn or knnOp). Therefore it can be advocated as a competent strategy for adaptation in CBR. It also negates the requirement for specific adaptation knowledge acquisition during system development as the required knowledge is obtained automatically, when required, in real time from the most appropriate cases in the case-base.

[1] For efficiency reasons, it is not necessary to calculate all the differences between all the case values; the differences between a case and random subset of other cases is sufficient.

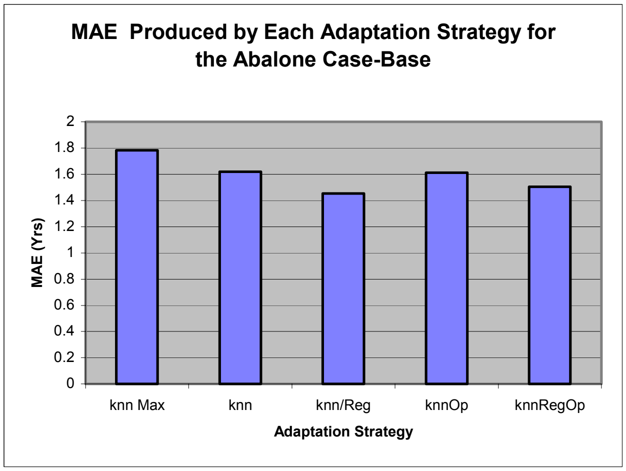

Figure 1 Affects of the adaptation strategies on MAE for Abalone case-base

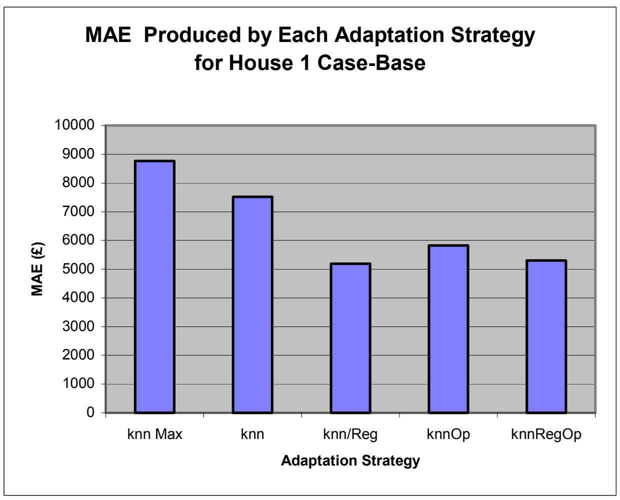

Figure 2 Affects of the adaptation strategies on MAE for House 1 case-base

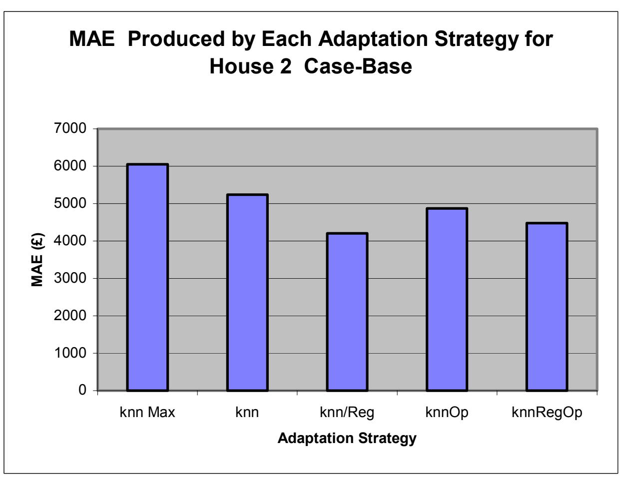

Figure 3 Affects of the adaptation strategies on MAE for House 2 case-base

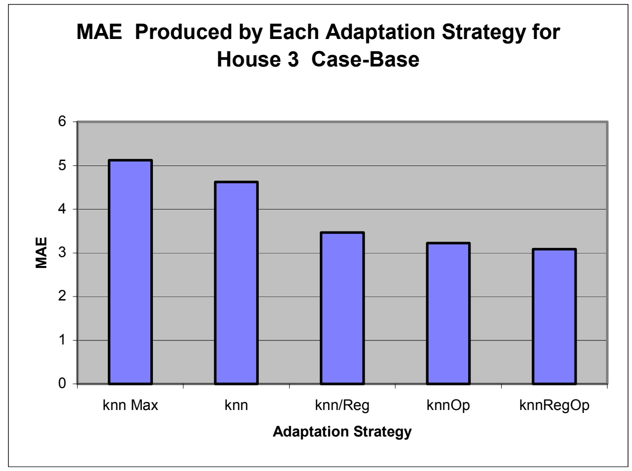

Figure 4 Affects of the adaptation strategies on MAE for House 3 case-base

This is an important benefit of the technique as human experts are no longer required during knowledge acquisition from whom to acquire knowledge. Additionally the need for adaptation knowledge maintenance is also not necessary as because no adaptation knowledge is stored (it is discovered each time it is required) then there is no requirement to maintain it. Where maintenance still has a crucial role is with case knowledge, as by carrying out case knowledge maintenance the adaptation knowledge discovered from it, is guaranteed to be optimal and current. Therefore case knowledge maintenance automatically subsumes adaptation knowledge maintenance at no additional cost to the system. Finally the most significant benefit of the approach is that the adaptation knowledge discovered is guaranteed to be locally relevant and specific to the task at hand. We believe this is why the technique provides such encouraging results. The results also confirm that FSS and weight optimisation has no positive effect on the competency of the adaptation technique.

Developing new strategies which improve one aspect of the CBR problem solving approach may cause limitations elsewhere. We have demonstrated that this approach improves problem-solving competency but how does it affect the efficiency of the problem solving process? It would be expected that the time taken to solve problems would be increased due to the extra time required to carry out the regression analysis after retrieval of the most relevant cases. From Table1 it can be seen that this is so.

Case-Base	knn sol. time (ms)	knn/Reg sol. time (ms)	Ratio
House 1	647	1421.1	2.2
House 2	705	1913.8	2.7
House 3	440.6	932.4	2.1
Abalone	21050.3	39009.1	1.9

Table 1 Solution times in ms

House 1 takes 2.2 times longer to adapt cases using the hybrid knn/reg strategy than with 5nn alone, house 2 takes 2.7 times longer, house 3 takes 2.1 times longer and abalone 1.9 times as long. These increased adaptation times reflect an improvement in competency of 31.0%, 19.7%, 25.1% and 10.2% respectively for house 1, house 2, house 3 and abalone when compared to knn and an improvement in competency of 11%, 13.6%, 4.3% (knnregOp) and 11.6% when compared to knnOp. Therefore the question arises - is the price of a decrease in efficiency worth paying for an improvement in competency? The answer to this depends on the goals of the system. If competency is paramount (as with most systems) then the decrease in efficiency is a small price to pay for such significant improvements in competency. Conversely if the system is time critical then the increased time costs associated with the improved competency may be too high and competency may have to be sacrificed to produce a less competent estimate of the solution in a faster time.

One possible solution to this trade off between competency and efficiency would be to use an indexing scheme to improve the speed cases are retrieved from the case-base. Indexes generally operate by identifying discriminatory features of cases and using these to partition the case-base into groups of cases with similar features. This is sometimes known as feature based recognition and a target case can be quickly matched with similar cases in the case-base through recognition of features they have in common. Examples of this type of indexing include k-d trees [Wess at al 1994], ID3 and C5.0 [Michalski et al 1999]. As only a selective portion of the case-base is made available during retrieval the efficiency of identifying a possible solution is increased dramatically. Unfortunately indexing cases correctly is not an easy task. The identification of a good feature for indexing is dependent on the retrieval circumstances. Therefore as circumstances change (as they inevitably will in a real world environment) the indexing structure of the case base must be maintained to reflect this. If the indexing scheme is poor or maintenance is ignored, cases with perfectly good solutions to the target problem may be overlooked as they reside in a different part of the case-base not accessible under the current indexing scheme. This can lead to the complex adaptation of less suited cases, the reduction in competency and in severe situations, problem-solving failures. Therefore, due to poor indexing and a lack of good maintenance, in an attempt to improve retrieval efficiency, competency is often sacrificed [Hunt et al 1995]. It has already been shown [Patterson et al 2002] how k-means clustering can define an efficient indexing strategy, which does not compromise on the competency of retrievals, while still providing an easily maintainable structure. Clustering is an unsupervised data mining technique, whereby groups of cases (clusters) are formed, based on their degree of similarity. The idea being that if they are similar they will have similar behaviours. When a target case, T, is presented, the cluster centroid it is closest to is identified. This thereby selects the cluster wherein T's most similar cases most probably lie. Retrieval is carried out on this identified cluster to provide an estimate of a solution. The observed efficiency improvements are because the retrieval process only considers cases in one cluster at any time, thus ignoring cases in the other clusters. Here we examine how this indexing technique affects the efficiency and competency of the knnReg adaptation strategy. Clusters were formed and cases to be used to determine the adaptation knowledge for a target were retrieved from the most relevant cluster.

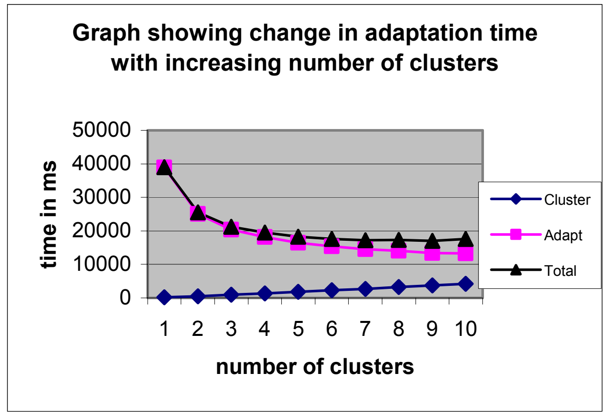

Figure 5 Time curves for the cluster based adaptation process

Figures 5 and 6 show the effect of k-means clustering on the knnReg adaptation process for the abalone case-base. From Figure 5 it can be seen that overall adaptation time (Total) decreases as the number of clusters formed increases. Also shown are times taken to form the individual clusters and the time for the regression based adaptation process itself (Adapt). Note total adaptation time is the cluster time plus the adaptation time and forming 1 cluster is equivalent to adaptation in the absence of clusters.

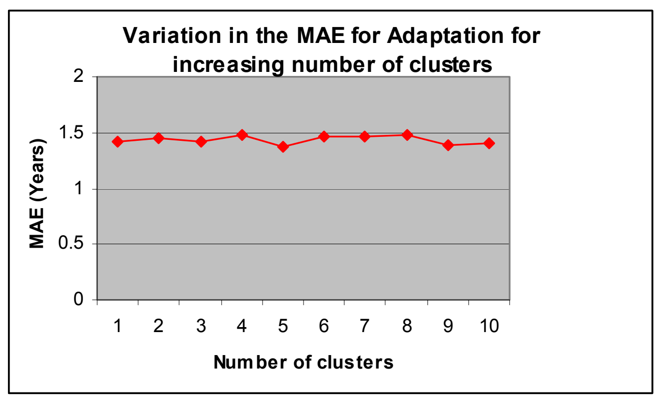

Figure 6 Showing how the MAE of adaptation varies with increasing cluster numbers

From this it can be seen how forming more than 7 clusters has little effect on efficiency of the adaptation process which takes 17174 ms at this point. If this is compared to the 5nn model, which takes 21050 ms (Table 1), it can be seen that the clustering process has in fact improved the efficiency of the adaptation strategy to the extent that it is now *more* efficient than 5nn retrieval by a factor of 1.3. Obviously this is only a useful approach if the competency of the adaptation process is not affected by clustering. From Figure 6 it can be seen that the competency of the adaptation process is stable as the number of clusters formed is increased. K-means clustering is therefore an efficient and competent indexing method to use to improve the speed of the adaptation process.

Conclusions and Future Work

In this paper a novel hybrid adaptation strategy for CBR was proposed and evaluated. It was shown to significantly improve the competency of adaptation compared to a standard knn approach. One possible limitation was the extra time required to adapt cases compared to knn. A novel indexing scheme based on k-means clustering was proposed as a way of reducing these effects and shown to provide more efficient adaptations than knn without loss of competency with one of the case-bases. This approach needs further experimentation in conjunction with the knnReg adaptation technique with more case-bases to verify its usefulness.

A drawback of this approach is that the regression strategy implemented is not very transparent, that is it is difficult to understand the adaptation knowledge the system uses to adapt the cases. A rule based adaptation knowledge base like Hanneys [Hanney & Keane 1996] would be easier to interpret but as pointed out can lead to maintenance problems which is something the hybrid approach will not cause.

In the future the technique should be extended to using nominal output features as opposed to just continuous as at present and the use of weighted regression. Additionally it could be applied to predicting missing values in data sets. Another interesting extension to this work would be to look at ways of storing the discovered adaptation knowledge for reuse in the future leading to a case-based adaptation process. This would improve the efficiency but increase the maintenance overhead. Finally we believe that this work has obvious implications for the machine learning community.

References

Aha, D. 1990. A study of instance-based algorithms for supervised learning tasks, Ph.D. diss., University of California, Irvine.

Allen, J.R.C., Patterson, D.W.R., Mulvenna, M. D. and Hughes, J.G. 1995. Integration of Case Based Retrieval with a Relational Database System in Aircraft Technical Support, 1st Intl Conference in Case Based Reasoning, pp 1-10.

Anand, S. S., Patterson, D. and Hughes, 1998. J. G. Knowledge Intensive Exception Spaces, Proceedings of 15th National Conference on Artificial Intelligence, pp 574-579.

Anand, S. S., Patterson, D., Hughes, J. G. and Bell, D. A. 1998. Discovering Case Knowledge Using Data Mining. Proceedings of 2nd Pacific-Asia Conference in Knowledge Discovery and Data Mining, pp 25- 35, LCNS Springer.

Cover, T.M. and Thomas, J.A. 1991. Elements of Information Theory. Wiley Series in Telecommunications. J. Wiley & Sons.

Dudani, S.A. 1975. Distance Weighted k-Nearest Neighbour rule. IEEE Trans on Systems, Man and Cybernetics, 6(4), pp 325-327.

Hanney, K. and Keane M. 1996. Learning Adaptation Rules from a Case-Base, Proc. Advances in Case-Based Reasoning, 3rd European Workshop, EWCBR, pp179-192.

Hanney, K. and Keane M. 1997. The Adaptation Knowledge Bottleneck. *Proc. Case-Based Reasoning: Research and Development*, *ICCBR-97*, pp359-370.

Hunt, J.E., Cooke, D.E. and Holstein, H. 1995. Case-memory and retrieval based on the immune system. 1st Int Conference on Case-Based reasoning (ICCBR-95), pp 205-216.

Kohavi, R. 1995. A study of cross validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95), pp 1137-1145, Sam Mateo, CA: Morgan Kauffmann.

Leake, D. (ed.) 1996. *Case-Based Reasoning : Experiences, Lessons and Future Directions*, MIT Press, MA.

Leake, D. B., Kinley, A. and Wilson, D. 1996. Acquiring case adaptation knowledge: A hybrid approach. Proceedings of the 13th National Conference on Artificial Intelligence, AAAI Press, Menlo Park, CA.

McSherry, D. 1999. Relaxing Similarity Criteria in Adaptation Knowledge Discovery, Workshop Automating the Construction of Case-Based Reasoners, 16th International Joint Conference on Artificial Intelligence, Stockholm, pp 56-61.

McSherry, D. 1999. Automating case selection in the construction of a case library. Proceedings of ES99, the19th SGES International Conference on Knowledge-Based Systems and Applied Artificial Intelligence, Cambridge, pp 163-177.

Michalski, R., Bratko, I. And Kubat, M. 1999. Machine Learning & Data Mining: Methods & Applications. J. Wiley and Sons.

Patterson, D., Anand, S.S., Dubitzky, D. and Hughes, J. 1999. Towards Automated Case Knowledge Discovery in the M^2 Case-Based Reasoning System, Knowledge and Information Systems:An International Journal, (1), pp 61-82, Springer Verlag.

Patterson, D., Anand, S.S., Dubitzky, D. and Hughes, J. A 2000. Knowledge Light Approach to Similarity Maintenance for Improving Case-Based Competence. Workshop on Flexible Strategies for Maintaining Knowledge Containers 14th European Conference on Artificial Intelligence ECAI.

Patterson, D., Rooney, N., Galushka, M. and Anand, S. Towards Dynamic Maintenance of Retrieval Knowledge in CBR. 2002. Proceedings Fifteenth International FLAIRS Conference.

Richter, M. 1995. The Knowledge Contained in Similarity Measures. Invited Talk, The First International Conference in Case-Based Reasoning, Sesimbra, Portugal.

Simoudis, E. and Miller, J. 1991. The application of CBR to helpdesk Aapplications. In Bareiss, R. (Ed). Proceedings of the Case-Based Reasoning Workshop, pp 25-36. San Mateo. DARPA, Morgan Kaufmann, Inc.

Torgo, L. 2000. Efficient and Comprehensible Local Regression in Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2000). Terano et al. (eds.). LNAI 1805, p. 376-379. Springer-Verlag.

Wess, S., Althoff, K-D. and Derwand, G. 1994. Using k-d trees to improve the retrieval step in case-based reasoning. In topics in case-based reasoning. Lecture notes in artificial intelligence, Vol. 837. Springer-Verlag, Berlin Heidelberg New York, pp 167-181.