

Retrieval strategies for case-based reasoning: a categorised bibliography

FRANCISCO AZUAJE,¹ WERNER DUBITZKY,² NORMAN BLACK³
and KENNY ADAMSON³

¹*Department of Computer Science, Trinity College, Dublin, Ireland.*

²*German Cancer Research Centre, Intelligent Bioinformatics Systems, Heidelberg, Germany.*

³*School of Information and Software Engineering, University of Ulster, Northern Ireland, UK.*

Abstract

The retrieval of relevant cases plays a crucial role in case-based reasoning. There are three major methods for the retrieval of relevant cases: computational approaches (based upon measures of similarity), representational approaches (based upon indexing structures) and hybrid approaches. This paper looks at recent successful implementations of case retrieval with regard to this classification framework. Similarly it emphasises computational and representational models applied to feature-vector case representations.

1 Introduction

Case-based reasoning (CBR) is one of the emerging paradigms useful for both modelling human reasoning (cognitive science) and for building intelligent computer systems (artificial intelligence). It essentially aims to develop an appropriate symbiotic relationship with human decision-makers, a process which may be developed in those areas where humans are not very good at retrieving and adapting past experiences. In this model, understanding and reasoning is viewed as a by-product of the underlying *memory processes* of memorising (storage) and reminding (retrieval).

The ability to retrieve past experiences is one of the most fundamental aspects of human cognition and is associated with the capacity of learning. The retrieval of *relevant* cases is crucial for recognition and classification and it plays an important role in scientific reasoning and creativity. Thus the idea of retrieving relevant cases is at the core of CBR systems and retrieval applications, such as in document retrieval, time series, sequence databases, data mining and bio-informatics.

In CBR, the basic processes of solving a new problem or interpreting a new situation entail the *retrieval of relevant* cases from a memory of cases (case base) followed by the *adaptation* of the past solution. Given the description of a new problem, called *query case*, the first and arguably most crucial step in a CBR system is to retrieve those cases from the case base that are most relevant to solving the query case. There are three fundamental approaches to retrieving relevant cases in CBR:

1. *Distance-based* approaches calculate the distance between cases by measuring the distances between the attributes that describe the cases. The most relevant cases are determined via a similarity/distance measure. The evaluation of distance or similarity may be performed by applying, for instance, distance functions defined ad hoc, expert knowledge acquisition techniques or automated learning methods. Within the first group, the Euclidean and Hamming distances are included, which are defined a priori without taking into account the informational *idiosyncrasy* of the case base under consideration. Knowledge acquisition-based approaches may be implemented to represent attribute- or case-relevance prior to the application of some

standard distance measure. Learning-based approaches have also been applied in order to discover relevant knowledge or improve the quality and efficiency of the retrieval strategies.

2. *Indexing* approaches encode relevant cases into the structure of the case base itself. The cases are organised by indexing structures, which are processed to search for relevant cases. Approaches that implement some type of distance measure during the construction or processing of a retrieval structure may be considered as indexing approaches if the retrieval of relevant cases is mainly coded and performed by indexing structures. Retrieval strategies based upon decision trees, supervised and unsupervised neural learning, and clustering algorithms are some of the approaches included in this category.
3. *Hybrid* approaches are the combination of the above two approaches. In this case, both indexing structures and distance measures are fundamental to performing the process of retrieval of relevant cases. The search for relevant cases may be performed by implementing indexing structures, and similarity measures may be generated to indicate the relevance of cases. Even when this type of approach has not been formally defined in the literature, they are considered as part of this classification framework for their conceptual significance and the existence of projects under development.

This classification framework and some of its dimensions included in this paper are depicted in Figure 1.

This paper aims to facilitate the work of researchers from the CBR and information retrieval communities in the identification of significant contributions to the retrieval of relevant cases. We concentrate on those publications which emphasise

1. strategies for the retrieval of relevant cases based on distance measures and index approaches since these are the most performed methods in different application domains;
2. retrieval of relevant cases in the context of learning – we are particularly interested in approaches that discover relevant knowledge based upon the characteristics of a particular case base;
3. those systems that are able to perform efficient retrieval in different application domains.

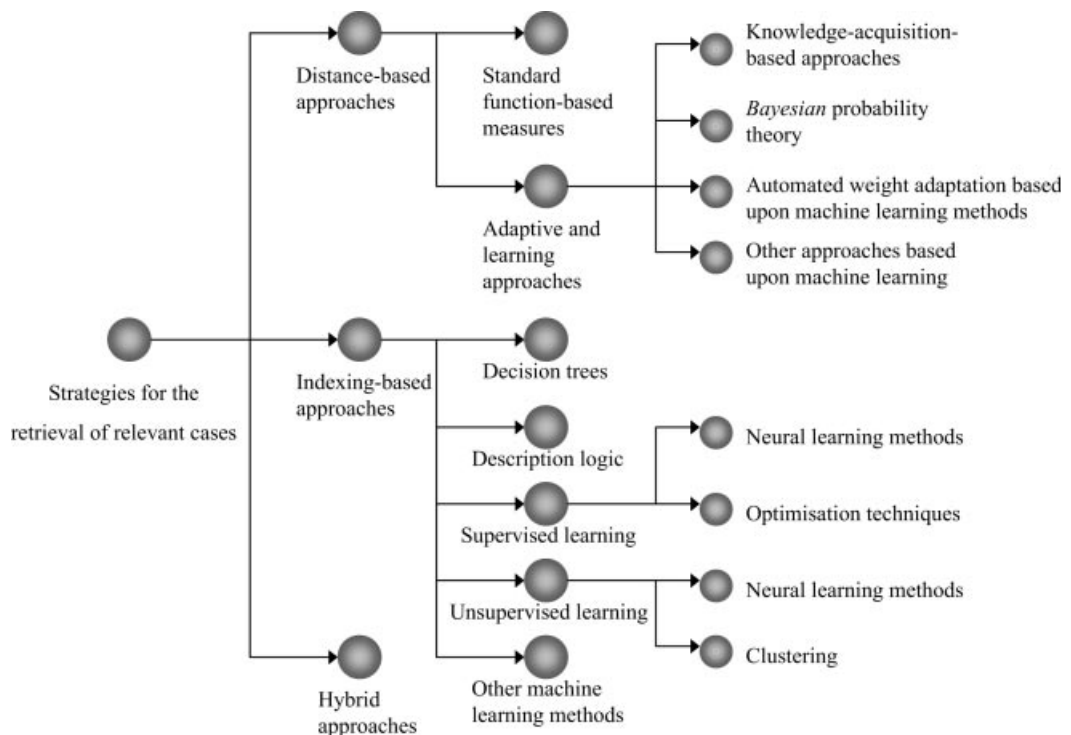


Figure 1 A framework of strategies for the retrieval of relevant cases

The remainder of this paper is organised into two main parts: section 2 is devoted to papers that refer to a first generation of distance-based approaches. Section 3 focuses on distance- and indexing-based approaches within the context of learning and adaptation.

2 A first generation of distance-based approaches

2.1 Standard quantitative metrics

This section contains references that describe retrieval approaches based on standard quantitative metrics. These models represent cases as points in some coordinate space, such that the observed similarities (or dissimilarities) between cases correspond to the metric distances between the respective points. These methods are based on the minimality, symmetry and triangle inequality axioms described by Tversky (1977).

- Aamodt, A, 1989, "Towards robust expert systems that learn from experience – an architectural framework" in: J Boose, B Gaines and J Ganascia (eds.) *Proceedings of EKAW89 Third European Knowledge Acquisition for Knowledge-Based Systems Workshop* 311–326.
- Aamodt, A and Plaza, E, 1994, "Case-based reasoning: foundational issues, methodological variations and system approaches" *AI Communications* 7(1) 39–59.
- Aamodt, A, 1991, "A knowledge-intensive approach to problem solving and sustained learning" Ph.D. dissertation, University of Trondheim, Norwegian Institute of Technology.
- Agre, G, 1995, "KBS Maintenance as learning two-tiered domain representation" *Proceedings of the First International Conference on CBR (ICCBR-95)* 109–120.
- Bath, PA, Poirrette, AR, Willett, P and Allen FH, 1994, "Similarity searching in files of three-dimensional chemical structures: comparison of fragment-based measures of shape similarity" *Journal of Chemical Information and Computer Sciences* 34 141–147.
- Cain, T, Pazzani, M and Silverstein, G, 1991, "Using domain knowledge to influence similarity judgments" *Proceedings of the Case-Based Reasoning Workshop Morgan Kaufmann* 23–27.
- Chang, CL, 1974, "Finding prototypes for nearest neighbor classifiers" *IEEE Transactions on Computers* 23(11) 1179–1184.
- Cover, TM and Hart, PE, 1967, "Nearest neighbor pattern classification" *IEEE Transactions on Information Theory* 13(1) 21–27.
- Diday, E, 1974, "Recent progress in distance and similarity measures in pattern recognition" *Proceedings Of the Second International Joint Conference on Pattern Recognition*, 534–539.
- Dudani, SA, 1976, "The distance-weighted k -nearest-neighbor rule" *IEEE Transactions on Systems, Man and Cybernetics* 6(4) 325–327.
- Gates, GW, 1972, "The reduced nearest neighbor rule" *IEEE Transactions on Information Theory* 18 431–433.
- Hart, PE, 1968, "The condensed nearest neighbor rule" *IEEE Transactions on Information Theory* 14 515–516.
- Hunt, JE, Cooke, DE and Holstein, H, 1995, "Case memory and retrieval based upon the immune system" *Proceedings of the First International Conference on CBR (ICCBR-95)* 205–216.
- Kolodner, J and Leake, D, 1996, "A tutorial introduction to case-based reasoning" in D Leake (ed.) *Case-Based Reasoning: Experiences, Lessons, & Future Directions* MIT Press.
- Kolodner, J, 1993, *Case-based reasoning* Morgan Kaufmann.
- Kolodner, J, 1995, "Improving human decision making through case-based decision aiding" in G Luger (ed.) *Computation & Intelligence*, AAAI Press/MIT Press.
- Liao, T, Zhang, Z and Mount, C, 1998, "Similarity measures for retrieval in case-based reasoning systems" *Applied Artificial Intelligence* 12 267–288.
- Pal, K and Campbell, JA, 1995, "A hybrid system for decision-making about asset in English divorce cases" *Proceedings of First United Kingdom Workshop on CBR* 152–156.
- Randall, D and Martinez, T, 1997, "Improved heterogeneous distance functions" *Journal of Artificial Intelligence Research* 6 1–34.
- Ricci, F and Avesani, P, 1995, "Learning a local similarity metric for case-based reasoning" *Proceedings of First International Conference on CBR (ICCBR-95)* 301–312.
- Ritter, GL, Woodruff, HB, Lowry, SR and Isenhour, TL, 1975, "An algorithm for a selective nearest-neighbor decision rule" *IEEE Transactions on Information Theory* 21(6) 665–669.
- Stanfill, J and Waltz, D, 1986, "Toward memory-based reasoning" *Communications of the ACM* 29 1213–1228.
- Surma, J and Vanhoof, K, 1995, "Integrating rules and cases for the classification task" *Proceedings of the First International Conference on CBR (ICCBR-95)* 325–334.
- Tversky, A, 1977, "Features of Similarity" *Psychological Review* 84 327–352.

- Tversky, A and Gati, I, 1978, "Studies of similarity" in E Rosch and BB Lloyd (eds) *Cognition and Categorisation*, Hillsdale.
- Watson, I and Marir, F, 1994, "Case-based reasoning: a categorised bibliography" *Knowledge Engineering Review* **9**(4) 355–381.
- Wess, S and Globig, C, 1993, "Case-based and symbolic classification: a case study" *Proceedings of the First European Workshop on CBR (EWCBR-93)* 77–91.
- Wilson, DL, 1972, "Asymptotic properties of nearest-neighbor rules using edited data" *IEEE Transactions on Systems, Man and Cybernetics* **2**(3) 408–421.

2.2 Qualitative, asymmetric and fuzzy approaches

Tversky (1977) questions the idea of similarity measure as a pure computation of metric distance between points. He suggests that similarity may better be described and measured as a comparison of features rather than as the calculation of a mere quantitative metric. Similarity measures based on qualitative approaches are included in this section, as well as methods based on asymmetric properties and binary and fuzzy representations.

- Bridge, D, 1998, "Defining and combining symmetric and asymmetric similarity measures" In B Smyth and P Cunningham (eds) *4th European Workshop on Case-based Reasoning, EWCBR-98 (Lecture Notes in Artificial Intelligence: Advances in Case-Base Reasoning, 1488)* Springer.
- Chen, SJ and Hwang, CL, 1992, *Fuzzy Multiple Attribute Decision-Making: Methods and Applications* Springer.
- Chen, SM, Yeh, MS and Hsiao, PY, 1995, "A comparison of similarity measures of fuzzy values" *Fuzzy Sets Systems* **72** 79–89.
- Cox, E, 1994, *The Fuzzy Systems Handbook: A Practitioner's Guide* Academic Press Professional.
- Davey, BA and Priestley, HA, 1990, *Introduction to Lattices and Order* Cambridge University Press.
- Diday, E, 1974, "Recent progress in distance and similarity measures in pattern recognition" *Proceedings Of the Second International Joint Conference on Pattern Recognition*, 534–539.
- Downs, GM and Willet, P, 1995, "Similarity searching in databases of chemical structures" *Reviews in Computational Chemistry* **7** 1–66.
- Dubitzky, W, Lopes, P, White, JA, Anderson, JMCC, Dempsey, GJ, Hughes, JG and Bell, DA, 1996, "A holistic approach to coronary heart disease risk assessment using case-based reasoning" *Proceedings of the 2nd International Conference on Neural Networks and Expert Systems in Medicine and Healthcare* 97–103.
- Dubitzky, W, Schuster, A, Bell, DA, Hughes, JG and Adamson, K, 1997, "How similar is very young to 43 years of age?: on the representation and comparison of polymorphic properties" *Proceedings of the 15th International Joint Conference on Artificial Intelligence* 226–232.
- Gesu, VD, 1994, "Integrated fuzzy clustering" *Fuzzy Sets Systems* **68** 293–308.
- Hyung, LK, Song, YS and Lee, KM, 1994, "Similarity measure between fuzzy sets and between elements" *Fuzzy Sets Systems* **62** 291–293.
- Krantz, DH and Tversky, A, 1975, "Similarity of rectangles: an analysis of subjective dimensions" *Journal of Mathematical Psychology* **12** 4–34.
- Lee, RW, Barcia, RM and Khator, SK, 1995, "Case-base reasoning for cash-flow forecasting using fuzzy retrieval" *Proceedings of the First International Conference on CBR (ICCB-95)* 23–26.
- Liao, TW and Zhang, ZM, 1996, "A review of similarity measures for fuzzy systems" *Proceedings of the Fifth IEEE International Conference on Fuzzy Systems* 100–106.
- Liao, T, Zhang, Z and Mount, C, 1998, "Similarity measures for retrieval in case-based reasoning systems" *Applied Artificial Intelligence* **12** 267–288.
- Matuschek, D and Jantke, KP, 1997, "Axiomatic characterization of structural similarity for case-based reasoning" *Proceedings of The Florida AI Research Symposium* 432–436.
- Miyamoto, S, 1990, "Information retrieval based upon fuzzy associations" *Fuzzy Sets Systems* **38** 191–205.
- Pappis, CP and Karacapilidis, NI, 1993, "A comparative assessment of measures of similarity of fuzzy values" *Fuzzy Sets Systems* **56** 171–174.
- Sellers, P, 1974, "An algorithm for the distance between two finite sequences" *Journal of Combinatorial Theory* **16** 253–258.
- Shepard, RN, 1987, "Toward a universal law of generalization for psychological science" *Science* **237** 1317–1323.
- Sneath, PHA, 1974, *Numerical Taxonomy: The Principles and Practice of Numerical Classification* WH Freeman.
- Tversky, A and Gati, I, 1978, "Studies of similarity" in E Rosch and BB Lloyd (eds) *Cognition and Categorisation*, Hillsdale.
- Tversky, A, 1977, "Features of similarity" *Psychological Review* **84** 327–352.

- Wess, S and Globig, C, 1994, "Case-based and symbolic classification: a case study" *Proceedings of The Second European Workshop on Case-based Reasoning (EWCBR-94)* 165–177.
- Zadeh, LA, 1965, "Fuzzy sets" *Information and Control* **8** 338–353.
- Zwick, R, Carlstein, E and Budescu, DV, 1987, "Measures of similarity among fuzzy sets: a comparative analysis" *International Journal of Approximate Reasoning* **1** 221–242.

3 Distance and indexing-based approaches within the context of learning and adaptation

The distance measure approaches presented above are based upon static distance measures defined a priori and independent of the application domain under consideration (function-based approaches). Learning approaches are those that are oriented towards the retrieval of relevant cases in a learning context and the modelling of the data under consideration. This section focuses on case retrieval methods for CBR based on learning approaches.

3.1 Initial approaches

This section aims at presenting some initial approaches to the problem of similarity-measure learning. Works relating to the psychological and philosophical background of case retrieval learning systems are also included.

- Biberman, Y, 1994, "A context similarity measure" in F Bergadano and L De Raedt (eds) *Machine Learning: European Conference on Machine Learning (Lecture Notes in Computer Science, 784)* Springer.
- Diday, E, 1974, "Recent progress in distance and similarity measures in pattern recognition" *Proceedings of the Second International Joint Conference on Pattern Recognition*, 534–539.
- Dudani, SA, 1976, "The distance-weighted k -nearest-neighbor rule" *IEEE Transactions on Systems, Man and Cybernetics* **6**(4) 325–327.
- Gebhardt, F, Voß, A, Gräther, W and Schmidt-Belz, B, 1997, *Reasoning with Complex Cases* Kluwer.
- Morgan, JN and Messenger, RC, 1973, *Thaid* Thomson-Shore.
- Nosofsky, R, 1986, "Attention, similarity and the identification–categorisation relationship" *Journal of Experimental Psychology* **115** 39–57.
- Pazzani, M, 1991, "Learning to predict and explain: an integration of similarity-based, theory-driven and explanation-based learning" *Journal of the Learning Sciences* **1**(2) 153–199.
- Ritter, GL, Woodruff, HB, Lowry, SR and Isenhour, TL, 1975, "An algorithm for a selective nearest-neighbor decision rule" *IEEE Transactions on Information Theory* **21**(6) 665–669.
- Salzberg, S, 1991, "A nearest hyperrectangle learning method" *Machine Learning* **6** 277–309.
- Santini, S and Jain, R, 1995, *Similarity Matching* Technical Report, Visual Computing Laboratory, University of California, San Diego.
- Sebestyen, GS, 1962, *Decision-making processes in pattern recognition* Macmillan
- Simon, JC, 1974, "A formal aspect of pattern recognition and scene analysis" *Proceedings of the Second International Joint Conference on Pattern Recognition*, 660–664.
- Sneath, PHA, 1974, *Numerical Taxonomy: The Principles and Practice of Numerical Classification* WH Freeman.
- Stanfill, J and Waltz, D, 1986, "Toward memory-based reasoning" *Communications of the ACM* **29** 1213–1228.
- Tversky, A and Gati, I, 1978, "Studies of similarity" in E Rosch and BB Lloyd (eds) *Cognition and Categorisation* Hillsdale.
- Vosniadou, S and Ortony A (eds) 1989, *Similarity and analogical reasoning*, Cambridge University Press.
- Watson, I and Marir, F, 1994, "Case-based reasoning: a categorized bibliography" *Knowledge Engineering Review* **9**(4) 355–381.
- Wettschereck, D and Dietterich, T, 1995, "An experimental comparison of nearest-neighbor and nearest-hyperrectangle algorithms" *Machine Learning* **19**(1) 5–28.

3.2 Retrieval strategies for cases described as feature vectors

This section focuses on learning-based retrieval strategies for feature-vector case representations. Cases described as n -tuple database records, signals and images are included in this categorisation. Some of these case-retrieval models are implemented by applying probabilistic theory, machine

learning, signal-processing methods, common-sequence analysis algorithms, decision trees and so on.

- Aha, DW and Goldstone, RL, 1992, "Concept learning and flexible weighting" *Proceedings of the 14th Annual Conference of the Cognitive Science Society*, 534–539.
- Anand, S, Patterson, D, Hughes, JG and Bell, DA, 1998, "Hybrid data mining systems: the next generation" in X Wu, R Kotagiri and K Korb (eds) *Research and Development in Knowledge Discovery and Data Mining: 2nd Pacific-Asia Conference on Knowledge Discovery and Data Mining (Lecture Notes in Artificial Intelligence 1394)* Springer.
- Anand, S, Scotney, B, Tan, M, McClean, S, Bell, D and Hughes, JG, 1997, "Designing a kernel for data mining" *IEEE Expert* **12**(2) 65–74.
- Azuaje, F, Dubitzky, W, Black, N and Adamson, K, 2000, "Discovering Relevance Knowledge in Data: A Growing Cell Structure Approach" *IEEE Transactions on Systems, Man and Cybernetics, Part B* **30**(3) 448–460
- Azuaje, F, Dubitzky, W, Black, N and Adamson, K, 1999, "Improving clinical decision support through case-based data fusion" *IEEE Transactions on Biomedical Engineering* **46**(10) 1181–1185.
- Azuaje, F, Dubitzky, W, Lopes, P, Black, N, Adamson, K, Wu, X and White, JA, 1998, "Discovery of incomplete knowledge in electrocardiographic data" *Proceedings of the 3rd International Conference of Neural Networks and Expert Systems in Medicine and Healthcare*, 286–294
- Azuaje, F, Dubitzky, W, Lopes, P, Black, N, Adamson, K, Wu, X and White, JA, 1998, "Knowledge discovery in electrocardiographic data based upon neural clustering algorithms" *Proceedings of the 8th Mediterranean Conference on Medical and Biological Engineering and Computing* (electronic proceedings on CD).
- Bloch, G, Sirou, F, Eustache, V and Fatrez, P, 1997, "Neural intelligent control for a steel plant" *IEEE Transactions on Neural Networks* **8** 910–918.
- Bonzano, A, Cunningham, P and Smyth, B, 1997, "Using introspective learning to improve retrieval in CBR: a case study in air traffic control" in DB Leake and E Plaza (eds) *Case-Based Reasoning: Research and Development: 2nd International Conference On Case-Based Reasoning*, Springer.
- Bridge, D, 1998, "Defining and combining symmetric and asymmetric similarity measures" in B Smyth and P Cunningham (eds) *Advances in Case-based Reasoning: 4th European Workshop on CBR (EWCBR-94) (Lecture Notes in Artificial Intelligence: Advances in Case-Base Reasoning, 1488)* Springer.
- Cost, S and Salzberg, S, 1993, "A weighted nearest-neighbour algorithm for learning with symbolic features" *Machine Learning* **10** 57–78.
- Das, G, Gunopulos, D and Mannila, H, 1997, "Finding similar series" in J Komorowsky and J Zytkow (eds) *Principles of Data Mining and Knowledge Discovery: First European Symposium (Lecture Notes in Artificial Intelligence: Principles of Data Mining and Knowledge Discovery, 1263)* Springer.
- Di Gesù, V and Starovoitov, V, 1999, "Distance-based functions for image comparison" *Pattern Recognition Letters* **20** 207–214.
- Dickhaus, H and Heinrich, H, 1996, "Classifying biosignals with wavelet networks" *IEEE Engineering in Medicine and Biology* **15** 103–111.
- Dorransoro, JR, Ginel, F, Sanchez, C and Santa Cruz, C, 1997, "Neural fraud detection in credit card operations" *IEEE Transactions on Neural Networks* **8** 827–834.
- Downs, GM and Willet, P, 1995, "Similarity searching in databases of chemical structures" *Reviews in Computational Chemistry* **7** 1–66.
- Downs, J, Harrinson, R, Lee, KR and Cross, S, 1996, "Application of the fuzzy ARTMAP neural network model to medical pattern classification tasks" *Artificial Intelligence in Medicine* **8** 403–428.
- Dubitzky, W, Azuaje, F, Lopes, P, Black, N, McCullagh P and Song, P, 1999, "On local and global feature weight discovery for case-based reasoning" *Proceedings of ISCA: 14th International Conference on Computers and their Applications* 107–110,
- Fayyad, UM and Irani, KB, 1993, "Multi-interval discretization of continuous-valued attributes for classification learning" *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, 1022–1027.
- Figliola, A and Serrano, E, 1997, "Analysis of physiological time series using wavelet transforms" *IEEE Engineering in Medicine and Biology Magazine*, **16**(3) 74–79
- Jagadish, H, 1996, "Indexing for retrieval by similarity" in V Subrahmanian and S Jajodia (eds) *Multimedia Database Systems: Issues and Research Directions* Springer.
- Juristica, I, Mylopoulos, J, Glasgow, J, Shapiro, H and Casper, R, 1998, "Case-based reasoning in IVF: prediction and knowledge mining" *Artificial Intelligence in Medicine* **12** 1–24.
- Kokar, MM and Reveliotis, SA, 1993, "Reinforcement learning: architectures and algorithms" *International Journal of Intelligent Systems* **8** 857–894.
- Kontkanen, P, Myllymaki, P, Silander, T and Tirri, H, 1998, "On Bayesian case matching" in B Smyth and P Cunningham (eds) *Advances in Case-based Reasoning: 4th European Workshop on CBR (EWCBR-94) (Lecture Notes in Artificial Intelligence: Advances in Case-Base Reasoning, 1488)* Springer.

- Korn, F, Sidiropoulos, N, Faloutsos, C, Siegel, E and Protopapas, Z, 1996, *Fast Nearest-Neighbour Search in Medical Image Databases*, Technical Report CS-TR-3613 & UMIACS-TR-96-17, University of Maryland Institute for Advanced Computer Studies.
- Osborne, H and Bridge, DG, 1997, "Models of similarity for case-based reasoning" *Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation* 173–179.
- Osborne, H and Bridge, DG, 1997, "Similarity metrics: a formal unification of cardinal and non-cardinal similarity measures" in DB Leake and E Plaza (eds) *Cased-Based Reasoning Research and Development: 2nd International Conference* Springer.
- Perner, P, 1998, "Different learning strategies in a case-based reasoning system for image interpretation" in B Smyth and P Cunningham (eds) *Advances in Case-based Reasoning: 4th European Workshop on CBR (EWCBR-94) (Lecture Notes in Artificial Intelligence: Advances in Case-Base Reasoning, 1488)* Springer.
- Perner, P, 1993, "Case-based reasoning for image interpretation in non-destructive testing" *Proceedings of EWCBR-93* 2 403–409.
- Raghavan, VV and Wong, SKM, 1986, "A critical analysis of vector space models for information retrieval" *Journal of the American Society of Information Science* 37 279–287.
- Ricci, F and Avesani, P, 1995, "Learning a local similarity metric for case-based reasoning" *Proceedings of First International Conference on CBR (ICCB-95)* 301–312.
- Rioul, O and Vetterli, M, 1991, "Wavelets and signal processing" *IEEE Signal Processing Magazine* 8 14–38.
- Ruiz de Angulo, V and Torras, C, 1997, "Self-calibration of a space robot" *IEEE Transactions on Neural Networks* 8 951–963.
- Struzik, Z and Siebes, A, 1998, "Wavelet transform in similarity paradigm" in X Wu, R Kotagiri and K Korb (eds) *Research and Development in Knowledge Discovery and Data Mining: 2nd Pacific-Asia Conference on Knowledge Discovery and Data Mining (Lecture Notes in Artificial Intelligence 1394)* Springer.
- Tanaka, E and Masuda, S, 1996, "Two-dimensional largest common subpatterns between pictures" *IEICE Transactions on Information and Systems* E79 1358–1361.
- Ting, KM, 1995, "Towards using a single uniform metric in instance-based learning" in M Veloso and A Aamodt (eds) *Case-based Reasoning Research and Development: First International Conference, ICCBR-95 (Lecture Notes in Artificial Intelligence, 1010)* Springer.
- Wu, S, Manber, U and Myers, G, 1990, "An O(NP)sequence comparison algorithm" *Information Processing Letters* 35 317–323.
- Wettschereck, D, Aha, DW and Mohri, T, 1995, *A Review and Comparative Evaluation of Feature Weighting Methods for Lazy Learning Algorithms* Technical Report AIC-95-012. Washington, DC: Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence.

3.3 Distance-based learning approaches

This section presents references on distance-based approaches for case retrieval. It has been shown that the performance of a case-retrieval strategy, such as the k -nearest neighbour method, depends on the relevance weights assigned to the attributes. Therefore the automated calculation of weights that represent attribute relevance is one of the crucial tasks in a distance-based approach. Thus many of the references provided in this section represent works on the discovery of attribute weights for case retrieval.

- Aha, DW (ed.) 1997, *Lazy learning* Kluwer Academic Publishers.
- Aha, DW and Goldstone, RL, 1992, "Concept learning and flexible weighting" *Proceedings of the 14th Annual Conference of the Cognitive Science Society* 534–539
- Anand, S, Patterson, D, Hughes, J and Bell, DA, 1998, "Hybrid data mining systems: the next generation" in X Wu, R Kotagiri and K Korb (eds) *Research and Development in Knowledge Discovery and Data Mining: 2nd Pacific-Asia Conference on Knowledge Discovery and Data Mining (Lecture Notes in Artificial Intelligence 1394)* Springer.
- Anand, S, Smith, A, Hamilton, P, Anand, J, Hughes, J and Bartels, P, 1999, "An evaluation of intelligent prognostic systems for colorectal cancer" *Artificial Intelligence in Medicine* 15 193–214.
- Atkeson, CG, Moore, AW and Schaal, S, 1997, "Local weighted learning" in Aha, DW, (ed.) *Lazy Learning* Kluwer Academic Publishers.
- Azuaje, F, Dubitzky, W, Black, N, Adamson, K, 2000, "Discovering relevance knowledge in data: a growing cell structure approach" *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 30(3) 448–460
- Bonzano, A, Cunningham, P and Smyth, B, 1997, "Using introspective learning to improve retrieval in CBR: a case study in air traffic control" in DB Leake and E Plaza (eds) *Proceedings of Case-Based Reasoning: Research and Development: 2nd International Conference On Case-Based Reasoning* Berlin: Springer.

- Cost, S and Salzberg, S, 1993, "A weighted nearest-neighbor algorithm for learning with symbolic features" *Machine Learning* **10** 57–78.
- Dasarathy, V (ed.), 1991 *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques* IEEE Computer Society Press.
- Dubitzky, W, Azuaje, F, Lopes, P, Black, N, McCullagh, P and Song, P, 1999, "On local and global feature weight discovery for case-based reasoning" *Proceedings of ISCA 14th International Conference on Computers and their Applications* 107–110.
- Gebhardt, F, Voß, A, Gräther, W and Schmidt-Belz, B, 1997, *Reasoning with Complex Cases* Kluwer.
- Han, U and Chater, N, 1998, "Understanding similarity: a joint project for psychology, case-based reasoning and law" *Artificial Intelligence Review* **12** 393–427.
- Liao, T, Zhang, Z and Mount, C, 1998, "Similarity measures for retrieval in case-based reasoning systems" *Applied Artificial Intelligence* **12** 267–288.
- Ling, CX and Wang, H, 1997, "Computing optimal attribute weight settings for nearest neighbor algorithms" in Aha, DW (ed.) *Lazy Learning* Kluwer Academic Publishers.
- Ricci, F and Avesani, P, 1995, "Learning a local similarity metric for case-based reasoning" *Proceedings of The First International Conference on CBR (ICCB-95)* 301–312.
- Watson, I and Marir, F, 1994, "Case-based reasoning: a categorized bibliography" *Knowledge Engineering Review* **9**(4) 355–381.
- Wettschereck, D, Aha, DW and Mohri, T, 1995, *A Review and Comparative Evaluation of Feature Weighting Methods for Lazy Learning Algorithms* Technical Report AIC-95-012. Washington, DC: Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence.
- Wettschereck, D, Aha, DW and Mohri, T, 1997, "A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms" *Artificial Intelligence Review* **11** 273–314.

3.4 Indexing approaches

This section contains references to those methods based on the construction of indexing structures to perform the retrieval of relevant cases. They are based on splitting methods, search algorithms, supervised and unsupervised neural networks, description logic and semantic networks.

- Arya, S, 1995, "Nearest neighbour searching and applications", Ph.D. Thesis, Computer Vision Laboratory, University of Maryland, College Park, Maryland, USA.
- Azuaje, F, Dubitzky, W, Black, N, Adamson, K, 2000, "Discovering relevance knowledge in data: a growing cell structure approach" *IEEE Transactions on Systems, Man and Cybernetics, Part B*, **30**(3) 448–460
- Azuaje, F, Dubitzky, W, Black, N and Adamson, K, 1999, "Improving clinical decision support through case-based data fusion" *IEEE Transactions on Biomedical Engineering* **46**(10) 1181–1185.
- Becker, L and Jazayeri, K, 1989, "A connectionist approach to case-based reasoning" in Hammond, KJ (ed.) *Proceedings of The Case-Based Reasoning Workshop* Morgan Kaufmann.
- Caudell, TP, Smith, SDG, Escobedo, R and Anderson, M, 1994, "NIRS: large scale ART-1 neural architectures for engineering design retrieval" *Neural Networks* **7** 1339–1350.
- Cleary, JG, 1979, "Analysis of an algorithm for finding nearest neighbours in Euclidean spaces" *ACM Transactions on Mathematical Software* **5** 183–192.
- Coupey, P, Fouqueré, C and Salotti, S, 1998, "Formalizing partial matching and similarity in case-based reasoning with a description logic" *Applied Artificial Intelligence* **12** 71–112.
- Gebhardt, F, Voß, A, Gräther, W and Schmidt-Belz, B, 1997, *Reasoning with Complex Cases* Kluwer.
- Gudivada, VV, Raghavan, VV, Grosky, W and Kasanagottu, R, 1997, "Information Retrieval on the World Wide Web" *IEEE Internet Computing Magazine* **1** 58–68.
- Kamp, G, 1994, "On the use of CBR in corporate service and support" *Proceedings of the 2nd European Workshop on Case-Based Reasoning* 175–183.
- Koehler, J, 1994, "An application of terminology logics to case-based reasoning" *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning* 351–362.
- Köhle, M and Merkl, D, 1996, "Visualizing similarities in high dimensional input spaces with a growing and splitting neural network" *Proceedings of the International Conference of Artificial Neural Networks (ICANN'96)* 581–586.
- Korn, F, Sidiropoulos, N, Faloutsos, C, Siegel, E and Protopapas, Z, 1996, *Fast Nearest Neighbour Search in Medical Image Databases* Technical Report CS-TR-3613 & UMIACS-TR-96-17, University of Maryland Institute for Advanced Computer Studies.
- Malek, M, 1995, "A connectionist index approach for CBR systems" in M Veloso and A Aamodt (eds) *Case-based Reasoning Research and Development: First International Conference, ICCBR-95 (Lecture Notes in Artificial Intelligence 1010)* Springer.

- Merkel, D, 1997, "Exploration of document collections with self-organizing maps: a novel approach to similarity representation" in J Komorowsky and J Zytchow (eds) *Principles of Data Mining and Knowledge Discovery: First European Symposium (Lecture Notes in Artificial Intelligence: Principles of Data Mining and Knowledge Discovery, 1263)* Springer.
- Mohri, T and Tanaka, H, 1994, "An optimal weighting criterion of case indexing for both numeric and symbolic attributes" in Aha, E (ed.) *Case-Based Reasoning: Papers from the 1994 AAAI Workshop* AAAI Press.
- Reategui, EB, Campbell, J and Borghetti, S, 1995, "Using a neural network to learn general knowledge in a case-based reasoning" in M Veloso and A Aamodt (eds) *Case-based Reasoning Research and Development: First International Conference, ICCBR-95 (Lecture Notes in Artificial Intelligence 1010)* Springer.
- Reategui, EB, Campbell, J and Leao, BF, 1997, "Combining a neural network with case-based reasoning in a diagnostic system" *Artificial Intelligence in Medicine* **9** 5–27.
- Reiser, C and Kaindl, H, 1994, "Case-based reasoning for multi-step problems and its integration with heuristic search" in Aha, E (ed.) *Case-Based Reasoning: Papers from the 1994 AAAI Workshop* AAAI Press.
- Ricci, F and Senter, L, 1998, "Structured cases, trees, and efficient retrieval" in B Smyth and P Cunningham (eds) *Advances in Case-based Reasoning: 4th European Workshop on CBR (EWCBR-94) (Lecture Notes in Artificial Intelligence: Advances in Case-Base Reasoning, 1488)* Springer.
- Rissland, E, Skalak, D and Friedman, T, 1996, "BankXX: supporting legal arguments through heuristic retrieval" *Artificial Intelligence and Law* **4** 1–71.
- Smith, SD, Escobedo, R, Anderson, M and Claudell T, 1997, "A deployed engineering design retrieval system using neural networks" *IEEE Transactions on Neural Networks* **8** 847–851.
- Sproull, RF, 1991, "Refinements to nearest-neighbor searching in *k*-dimensional trees" *Algorithmica* **6** 579–589.
- Ultsch, A, Farsch, S and Li, H, 1995, "Automatic acquisition of medical knowledge from data sets with neural networks" *Proceedings of KI'95* 258–260.
- Watson, I and Marir, F, 1994, "Case-based reasoning: a categorized bibliography" *Knowledge Engineering Review* **9**(4) 355–381.
- Wess, S, Althoff, KD and Derwand, G, 1994, "Using *k*-d trees to improve the retrieval step in case-based reasoning" in KD Wess, S Althoff and MM Richter (eds) *Topics in Case-Based Reasoning*, 167–181, Berlin: Springer.
- White, D and Jain, R, 1996, "Similarity indexing: algorithms and performance" *Proceedings of the SPIE: Storage and Retrieval for Image and Video Databases IV* 62–75.
- Wolverton, M and Hayes-Roth, B, 1994, "Retrieving semantically distant analogies with knowledge-directed spreading activation" *Proceedings of the 12th National Conference on Artificial Intelligence* 56–61.
- Yuanhui, Z, Yuchang, L and Chunyi, S, 1997, "Using neural networks to extract knowledge from databases" in J Komorowsky and J Zytchow (eds) *Principles of Data Mining and Knowledge Discovery: First European Symposium (Lecture Notes in Artificial Intelligence: Principles of Data Mining and Knowledge Discovery, 1263)* Springer.
- Zavrel, J, 1996, "Neural navigation interfaces for information retrieval: are they more than an appealing idea?" *Artificial Intelligence Review* **10** 477–504.