A Value Supplementation Method for Case Bases with Incomplete Information

Kerstin Bach, Meike Reichle, and Klaus-Dieter Althoff

Intelligent Information Systems Lab
University of Hildesheim
Marienburger Platz 22, 31141 Hildesheim, Germany
{bach,reichle,althoff}@iis.uni-hildesheim.de

Abstract. In this paper we present a method for supplementing incomplete cases with information from other cases within a case base. The acquisition of complete and correct cases is a time-consuming task, but nevertheless crucial for the quality and acceptance of a case-based reasoning system. The method introduced in this paper uses association rules to identify relations between attributes and, based on the discovered relations we are able to supplement values in order to complete cases. We argue that using these related attributes when retrieving supplementation candidates will yield better results than simply picking the case with the highest global similarity. The evaluation of the method is carried out using four different publicly available case bases.

1 Introduction

Incomplete information in cases is a problem often encountered in various areas of Case-Based Reasoning (CBR). For instance, Bogaerts and Leake [1] discuss how to assess the similarity of incomplete problem descriptions in Conversational CBR applications, while Selvamani and Khemani [2] discuss how missing information can be completed using decision tree induction. Further on when cases are collected from WWW sources, like blogs, websites or web communities as in [3], completing cases or dealing with incomplete information is one of the major challenges.

Missing attributes can happen for a number of reasons. For instance, considering products as cases as within our example case bases, attributes can be empty either because the attribute's value is unknown or because a certain attribute doesn't apply to a certain case/product. Obviously only the first group should be substituted, so a substitution strategy should ideally also include a set of rules or constraints that controls when an empty attribute is substituted and when a substitution is not applicable. Such constraints could for example be based on other attribute's values ("If a PC is a Desktop do not substitute the battery attribute") or a similarity-based comparison with other products and their missing values.

In this paper we present a method for supplementing incomplete cases in Structured CBR (SCBR) applications [4] using only the case base itself, the

L. McGinty and D.C. Wilson (Eds.): ICCBR 2009, LNAI 5650, pp. 389–402, 2009.

knowledge already included in the cases and the similarity model. In SCBR a case can be represented as attribute-value tables, in an object-oriented manner, trees or graphs as well as in predicate logic [5]. We focus on the representation in attribute-value tables, because it is one of the most common kind of case representation in CBR. Cases in SCBR are represented by a predefined set of attributes and the range of the attributes' values (mostly nominal and numerical) is given in a vocabulary [6].

This paper picks up on previous experiments on the subject, raises them to a more general level, evaluates the results of the method and identifies constraints on its applicability.

The aim of this paper is to demonstrate that – given the necessity to supplement cases – our method leads to more accurate supplementations than doing a standard CBR retrieval on the case to be supplemented and simply supplementing it with the attribute values from the most similar case. Our method provides a result set which is optimized towards retrieving the best fitting supplementation candidates, even if they are not actually the most similar cases.

The work in this paper is structured as follows: Section 2 gives a short overview on the preliminary work already carried out on the subject and the first practical results achieved within our project docQuery [7]. Section 3 presents and evaluates the method in a more generalized way: in subsection 3.1 we describe the experiments used to evaluate the general applicability of our method, followed by the presentation of the results of the individual experiments in subsection 3.2 and the interpretation of the results as well as a concluding estimation of its general applicability in section 3.3. Section 4 presents related work on comparable topics. The paper concludes with a brief summary and an outlook on future work in section 5.

2 Preliminary Work

We initially presented this approach in [8] as an improvement to adaptation in case bases consisting of a complete set of cases but with individual cases suffering from incomplete information. In that first scenario we dealt with a geographic case base, which was used in the context of travel medicine. This case base included cases for all known countries but with a heavily varying information quality, i.e. some attributes were always present, such as the vaccinations that are obligatory in order to enter a country, others were often empty. Since the application required a complete case for its next steps to work out we had to develop a method for filling those attributes with values. A closer study of the case format revealed that there are certain attributes that are related with regard to their content, in that case for instance the necessary vaccinations and the general list of infection risks. The content of the vaccinations attributes suggested, at least partly, the contents of the general infection risk, and thus lent itself to be used in order to derive a sensible value for that attribute.

Assuming that similar vaccinations suggest similar general infection risks, we then developed a 2-step retrieval method in order to find the optimal case from which to take over the necessary values. Applied to the given example the method was carried out as follows:

1. Select the desired country from the case base. Since we identify the countries by name, which is a unique attribute, this can be done using similarity based retrieval as well as a simple selection.
2. If the country's general infection risks are not indicated, extract the content of the vaccinations attribute.
3. Send a new query, this time using the country's name and vaccinations as query input.
4. Take the result set, remove the best hit (which will again be the country in question), randomly pick one of the remaining countries with the highest similarity.
5. Extract the randomly picked country's general infection risks and use them to supplement the original country's information. If the picked country also has an empty general infection risks attribute pick another case of the same similarity.

In order to evaluate the quality of the resulting supplementations we manually prepared a test case base (countries of South East Asia) with complete information and then subsequently took every country, emptied its general infection risks and restored them using once the 2-step retrieval method and once using only a geographic taxonomy as the similarity measure in order to pick out a supplementation candidate.

Using the 2-step retrieval method we were able to significantly reduce the number of supplementation candidates in 90% of the test. Evaluating the quality of the supplementations done with the respective remaining candidates we found that in a total of 90 supplementations using the taxonomy based retrieval 62% of the supplemented cases contained all of the expected infection risks. Using the 2-step retrieval method on the same cases amounted in 76% of the supplemented cases containing all expected infection risks. Although both retrieval variants also returned false positives in most of the tests, the solutions of the 2-step retrieval method were generally more reliable, especially with respect to false negatives, which were the more serious problem in this particular application scenario.

Our application scenario and its underlying architecture SEASALT [3] uses modularized knowledge bases. For the combination of information retrieved from these knowledge bases we do a subsequent retrieval and the more information one solution contains, the more possibilities the algorithm has for further retrieval steps. If we would only have a single retrieval step, a more precise similarity measure including specialised domain knowledge would probably work as well.

3 Generalization

Phrased more generally, according to the 4-R model [9] our supplementation steps in after the *retrieve* step, but before a potential *reuse* step. Thus in our point of view we replace the general *retrieve* step, with two steps: first we retrieve the desired case as usual, but then we check it for completeness, and, if the values of a related attribute are missing, we do a second retrieval in order to get a supplementation candidate case. We then use its values to supplement the desired case and finally pass it on to *reuse*, *revise*, and *retain*.

In order to see if the method is generally suited for supplementing incomplete cases we identified the following research hypotheses:

1. *There are pairs of attributes that are related with respect to their content, i.e. the value of one attribute determines – with a certain confidence – the value of the other attribute. If a case format includes such relations between attributes they can ...*
   (a) be identified automatically and
   (b) be used to supplement missing values of related attributes.
2. *The supplementation candidates retrieved using only related attributes will be the most fitting, i.e. the results of the supplementation will be better than the results when using a global similarity measure to retrieve the supplementation candidates.*

3.1 Experiments towards an Extensions of the Initial Scenario

In order to be as representative as possible we used publicly available test case bases from UCD, namely the PC and Whiskey case bases [10], the camera case base [11] and AI-CBR's travel case base[1]. These case bases each provide a set of cases as well as a case format and the associated similarity measures. We used the models and measures as indicated.

The camera case base consists of 210 case described with four nominal and six numerical attributes. The numerical attributes have a predefined range depending on their values. The global similarity measure is calculated using each attribute with the same weight. The whiskey case base consists of 553 cases and each case is represented by ten attributes, five nominal and five numerical attributes. The global similarity measure is calculated using each attribute with the same weight. The PC case base contains 120 cases which are represented by eight attributes and the global similarity measure is calculated using each attribute with the same weight. Three attributes are nominal and five are numerical with a defined range of possible values. The travel case base contains 1024 cases and the case representation consists of six nominal and three numerical attributes. The global similarity measure is calculated using each attribute with the same weight.

As a first step we calculated the related attributes. For this purpose we used a simple 1R algorithm [12] for finding association rules within the cases of one case base. Then we compared for each attribute combination the respective rules with a confidence $\geq 67\%$ against the total number of all value combinations for those attributes within the case base, resulting in a final correlation score.

[1] We gathered the XML files describing the case bases from the Case-Based Reasoning Wiki at http://cbrwiki.fdi.ucm.es/wiki/index.php/Case_Bases

$$\text{c-score}_{(A1,A2)} = \frac{\text{\# of rules from A1 to A2 with a confidence} \geq 67\%}{\text{\# of combinations from A1 to A2 within the case base}} * 100$$

For example considering the Whiskey case base there are 193 different combinations of values of the attributes *Proof* and *Finish*. 77 of these combinations could serve as an association rule with a confidence ≥ 67%. Thus the c-score for *Proof* → *Finish* amounts to 39.9.

We then iterated over each case base and, for each case, we tried to supplement the identified attributes using the respective related attributes. We did this twice for each case, once using the whole case for the second query, once using only the identified related attribute. In order to simulate an incomplete case base we each time randomly removed three attribute values from the complete cases, but never the related attribute, so the supplementation in the first test was based on the retrieval results using all remaining attributes. The second test supplementation was based on a query using only the related attributes.

To illustrate this with an example let's consider the camera case base. In this case base one of the highest rating attribute pairs was *Weight* → *Format*, we will thus try to supplement the *Format* attribute. We consider the case listed in table 1 and try to supplement the value of *Format*. In the first test we simply use the whole case in order to find the most similar camera and use its *Format* for supplementation. The result set naturally lists the original case first, then, with the second best amount of similarity, there are two supplementation candidate cases. One of these cases has the desired value "SLR" and the other one "Compact", so there would have been a 50% chance of a correct supplementation. In

Table 1. Example case from the camera case base. *Format* is the attribute to be supplemented, *Manufacturer*, *Model* and *Storage Included* have been randomly removed.

Attribute	**Value**
ID	Case140
Manufacturer	- removed -
Model	- removed -
Price ($)	900
Format	- missing -
Resolution (M Pixels)	1.92
Optical Zoom (X)	3.2
Digital Zoom (X)	2
Weight (grams)	630
Storage Type	Compact Flash
Storage Included (MB)	- removed -

the second test we only use the related attribute *Weight* for retrieval. The result is computed analogously. This time the chance of a correct supplementation would have been 71.42%, because 10 out of 14 supplementation candidate cases suggest the correct value.

We did this once for each case in each of the four case bases using the following attribute pairs[2]:

- Camera
  - *Camera.Weight* → *Camera.Format*
  - *Camera.Weight* → *Camera.OpticalZoom*
  - *Camera.Weight* → *Camera.StorageType*
  - Excluded attributes: *CaseId* (unique, avg. frequency 1), *Model* (avg. frequency 1.005), *Price* (avg. frequency 1.2)
- PC
  - *PC.Monitor* → *PC.Type*
  - *PC.DriveCapacity* → *PC.Type*
  - *PC.ProcessorSpeed* → *PC.ProcessorType*
  - Excluded attributes:*Price* (avg. frequency 1.5)
- Travel
  - *Travel.Hotel* → *Travel.Accommodation*
  - *Travel.Hotel* → *Travel.Region*
  - *Travel.Hotel* → *Travel.Transportation*
  - *Travel.Region* → *Travel.Transportation*
  - Excluded attributes:*Price* (avg. frequency 1.6)
- Whiskey
  - *Whiskey.Proof* → *Whiskey.Finish*
  - *Whiskey.Proof* → *Whiskey.Availability*
  - *Whiskey.Proof* → *Whiskey.Sweetness*
  - *Whiskey.Proof* → *Whiskey.Peatiness*
  - Excluded attributes: none

3.2 Results

Considering the supplementation results it can be noted that our first research hypothesis (*There are pairs of attributes that are related with respect to their content, i.e. the value of one attribute determines – with a certain confidence – the value of the other attribute. If a case format includes such relations between attributes they can (a) be identified automatically and (b) be used to supplement missing values of related attributes.*) is confirmed by the results of our experiments. We were able to reliably detect meaningful correlations between attributes and with very few exceptions the attribute pairs with the highest correlation score were also the ones with the best supplementation results. Even

[2] Attribute pairs including unique or almost unique attributes (i.e. attributes with values with an average frequency near 1) were manually excluded.

when supplementation using all available attributes performed better overall, the difference between the results is smaller the higher the correlation score is.

Our second research hypothesis (*"The supplementation candidates retrieved using only related attributes will be the most fitting, i.e. the results of the supplementation will be better than the results when using other retrieval methods to retrieve the supplementation candidates."*) only held for some of the case bases.

In detail the results were as follows: In the Camera case base (see results in table 2) supplementation based on a retrieval using all available attributes outperformed supplementation based on a search using only the related attribute in all of the tests. In the Whiskey case base (see results in table 3) supplementation

Table 2. Successful supplementations in the camera case base: based on all available attributes vs. using only the related attribute

Att Pair	**Correlation Score**	**Correct supps all atts [%]**	**Correct supps only rel att [%]**	**Improvement [%]**
Weight → *Format*	68.52	76.00	63.00	-13.00
Weight → *OpticalZoom*	56.10	63.00	47.00	-16.00
Weight → *StorageType*	50.00	72.00	30.00	-42.00
		Camera Case Base		

Table 3. Successful supplementations in the whiskey case base: based on all available attributes vs. using only the related attribute

Att Pair	**Correlation Score**	**Correct supps all atts [%]**	**Correct supps only rel att [%]**	**Improvement [%]**
Proof → *Finish*	39.90	47.00	66.66	19.66
Proof → *Availability*	38.01	18.90	18.75	-0.15
Proof → *Sweetness*	32.96	26.10	10.40	-15.70
Proof → *Peatiness*	32.60	22.70	8.30	14.40
		Whiskey Case Base		

based on a retrieval using all available attributes was outperformed by supplementation based on a search using only the related attribute in 25% of the tests. In the PC case base (see results in table 4) supplementation based on a retrieval using all available attributes was outperformed by supplementation based on a search using only the related attribute in 66% of the tests. The outperformed attribute pair was also the one with the lowest correlation score. In the Travel case base (see results in table 5) supplementation based on a retrieval using all available attributes was outperformed by supplementation based on a search using only the related attribute in 100% of the tests. The result tables each indicate the used attribute pair,

Table 4. Successful supplementations in the PC case base: based on all available attributes vs. using only the related attribute

Att Pair	**Correlation Score**	**Correct supps all atts [%]**	**Correct supps only rel att [%]**	**Improvement [%]**
Monitor → *Type*	46.67	51.00	65.00	14.00
DriveCapacity → *Type*	45.00	45.30	46.70	1.40
ProcSpeed → *ProcType*	37.50	76.00	56.00	-20.00
PC Case Base				

Table 5. Successful supplementations in the travel case base: based on all available attributes vs. using only the related attribute

Att Pair	**Correlation Score**	**Correct supps all atts [%]**	**Correct supps only rel att [%]**	**Improvement [%]**
Hotel→ *Accomodation*	99.44	64.37	99.38	36.01
Hotel → *Region*	98.04	41.68	98.57	56.89
Hotel → *Transportation*	90.21	88.22	92.76	4.54
Region → *Transportation*	77.33	89.46	93.53	10.07
Travel Case Base				

their respective correlation score and the percentage of correct supplementations in both tests as well as the difference between both tests (Improvement).

3.3 Evaluation of the Experiments and Their Results

Summarizing the results of all four case bases we think that our supplementation method shows promise also on a more general level. The very different nature of the case bases used in these experiments allows us to draw first conclusions with regard to the general applicability of our method. We are very satisfied with the results of our identification of related attribute couples. Although it uses a very simple algorithm and is overall implemented in a rather pragmatic way it achieves very good results. The method also seems to perform equally well on case bases with more and fewer numerical attributes. An additional benefit of detecting such related attributes is the possibility to use this knowledge in order to improve the system's similarity model. The determining attributes obviously possess a high information value. On the other hand the attributes that can be deduced from them are not necessarily redundant but have more of a supporting role. This knowledge can for instance be reflected in assigning a higher weight to the determining attributes and a lower one to the supporting attributes. By recalculating the correlation scores on a regular basis and automatically adapting the attribute weights accordingly the correlation score provides an easy way of improving the similarity model along with the case base's competence.

Regarding the supplementation results, the Travel case base's results are obviously best, since the identified attribute pairs are very closely related, which is also reflected in the very high correlation scores. However there are also case bases in which a retrieval with all available attributes performs better than our method, the most prominent example being the Camera case base. We assume that these bad results of our method are caused by weak attribute pairs, i.e. an inexact computation of correlation scores. The Camera case base is comparatively small (210 cases) but has a rather high number of attributes (10), many of which again have a large range of possible values. This means that the majority of possible value combinations is not covered in the case base and that the amount of covered combinations might even not be representative. As presented by MacDonald et. al. [13] a high number of attributes and possible values also requires a high number of cases. Otherwise the minimum similarity threshold has to be specified so low that result quality is no longer acceptable. We assume that the same holds for attribute correlations.

Concerning the applicability of our approach there are of course certain limits, mostly with respect to the application domain. Most of all there have to be attributes which are related with regard to their content, but it should also be kept in mind that any supplementation method comes down to more or less educated guessing and thus should not be used in domains that are safety critical or require high precision data. Also, as mentioned in the introductory notes, not all missing values require supplementation. Attributes that don't necessarily apply in any case should possess a null value that indicates a deliberately empty value. Also attributes with a high similarity weight could be treated more carefully (e.g. by requiring a higher confidence value when doing supplementations) in order to avoid too much of an effect on the retrieval results.

Finally it would be advisable to reflect the fact that a case has been supplemented in the case's description, thus allowing the ranking mechanism to favour more complete and thus more reliable cases and also creating a higher transparency towards the user.

4 Related Work

Several researchers have presented works on topics related to the work presented in this paper. Data Mining or Knowledge Discovery techniques have already been combined with CBR in the past. O'Sullivan et. al. [14] used data mining algorithms to maintain similarity knowledge in order to improve case-based collaborative filtering recommendations. Díaz-Agudo et. al. [15] used the Formal Concept Analysis (FCA) as an inductive technique to extract domain specific knowledge from cases. They used FCA in knowledge intensive applicatioons to enrich domain ontologies or change the organization of case bases.

Dubois et al. [16] also consider similarities between a case's attributes and combine that knowledge with fuzzy logic in order to deal (among other things) with incomplete cases.

The relations between attributes have also been investigated by Tawfik and Kasrin [17] who represent them using dependency graphs which are then sectioned using either d-separation [18] or multiply sectioned Bayesian networks [19]. Tawfik and Kasrin use the resulting subgraphs/-cases to generate completely new cases for the purpose of increasing case base coverage. By using the subgraphs they aim to detect dependencies between attributes and thus prevent intra-case inconsistencies [20] when generating new cases.

Redmond [21] introduced an approach for combining information from different cases. They define a case as a set of information pieces, like snippets in [22], consisting of an attribute-value-pair. Each snippet is assigned to a particular goal and holds information on how to pursuit this goal. Since the reference application originates in CBR-diagnosis, the snippets also contain information (links) that preserve the structure of the diagnosis. Further on, they use a case-based reasoning process to retrieve single snippets and based on the predefined links they put together a problem's solution. The snippet information is highly dependent on the domain and has to be modeled by hand. The approach presented in our paper focuses on a more general method that also uses information from different cases by employing knowledge contained in the different knowledge containers [23], without explicitly modeling additional case information. However, both approaches have in common that the reasoning processes are used to supplement incomplete information in order to find better solutions.

The approach of doing several, subsequent retrieval steps instead of only one can also be found in the works of several authors. Weibelzahl [24,25] present an approach based on two different case bases. On the one hand they use a specific case base to create an enriched query that uses the given information more effectively and on the other hand they do regular CBR. They evaluate the approach in a system on holiday recommendation consisting of two case bases with different knowledge models. The first case base, called customer case base, holds information on the customers' needs and desires which are mapped to attributes describing products provided in the second case base. In the first step the query containing the user's expectations on their vacation is analysed in order to fill relevant attributes creating a request which can be sent to the product case base. The second request contains especially those product attributes which the user would not request on their own, but which help to find an appropriate solution in the product case base. In comparison to our approach, we use the case base's knowledge model to enhance the query aiming at a more differentiated result while Weibelzahl points out that users cannot exactly describe their desires by framing a request. The incremental approach kind of matches the users statement to correct attribute-value pairs.

A similar approach is presented by Cunningham et al. in [26,27]. They introduce the Incremental CBR (I-CBR) mechanism for diagnosis. The I-CBR approach separates information in "free" and "expensive" features and starts the first retrieval steps based on the free features before the user is asked to give information about expensive features to narrow the set of cases. In comparison with their approach we have a different point of view. The method presented in our paper is able to indicate attributes that are determining for the retrieval and those which can be derived from the knowledge within the case base. In contrast to Cunningham's method, our approach does not classify "free" and "expensive" attributes; instead we are able to supplement missing information and thus do not require "expensive" information at all.

Another approach on how I-CBR can influence the result sets has been presented in [28], but in comparison to our approach Jurisica et. al. did not receive additional information from existing cases, they used query series and user interaction instead.

5 Conclusion and Outlook

Case acquisition is an extensive and time-consuming task, and often has to deal with incomplete information, resulting in incomplete cases. Supplementing such incomplete cases with information adapted from other cases is a relatively easy way to improve a case base. However, when carrying out such a supplementation the choice of which cases to supplement from is of paramount importance, since a wrong supplementation may actually worsen a case's information quality or even make it inconsistent. Association rules are a handy tool to identify attributes that are related with regard to their content, a knowledge which can be used well when choosing the optimal candidate for a supplementation.

In this paper we presented a method for supplementing incomplete cases using attributes from other cases that makes use of association rules and similarity based retrieval in order to pick an optimal supplementation candidate. On the one hand our method produces better supplementation candidates than using a CBR system's standard case retrieval and global similarity measure, since it focuses on the related attributes. On the other hand, it could neither be replaced by a simple rule-based approach since it makes use of the CBR system's underlying similarity model and thus, if no valid supplementation candidate can be found, it will at least come up with a most similar value instead of none.

After having tested the method in one of our projects already, we now evaluated it using publicly available test case bases. The results of these evaluations are promising in so far as that our research hypotheses hold and the method performs well in most scenarios. However not all results are good, so there is still room for further research and improvement.

This further research will for instance concern the question under which circumstances the identification of related attributes works best, i.e. if there is a minimum number of cases and/or attribute combinations necessary in order for our method to yield satisfactory results. We will also do a few experiments on other association rule learning algorithms in order to find out how far the relative simplicity of the 1R algorithm influences the overall result quality. Apart from other association rule learning algorithms we will also evaluate our method against other substitution methods, e.g. not randomly picking the substitution value from the substitution candidates but using the most frequent value or a mathematic mean.

We are also interested in trying our method on other case bases in order to gain even more insight in the conditions of its general applicability and its behavior under different conditions. It would be especially interesting to evaluate our method with a more realistic test case base, that is a case base where we don't have to randomly delete values but already have missing values that derive from an real life application nd are thus not as uniformly distributed. Alternatively, if such a test case base is not available, we will do some more experiments with varying amounts of removed values.

Another aspect that will receive greater attention in our future work are the possibilities to integrate the correlation score in automated improvement of the system's similarity model, as already sketched out in section 3.3, and possibly other areas of CBR research such as maintenance and adaptation. Finally a topic that might become more relevant in future experiments is the performance of our method and whether/how it can be improved with regard to computation time.

References

1. Bogaerts, S., Leake, D.B.: Facilitating CBR for incompletely-described cases: Distance metrics for partial problem descriptions. In: Funk, P., González-Calero, P.A. (eds.) ECCBR 2004. LNCS, vol. 3155, pp. 62–76. Springer, Heidelberg (2004)
2. Selvamani, R.B., Khemani, D.: Decision tree induction with CBR. In: Pal, S.K., Bandyopadhyay, S., Biswas, S. (eds.) PReMI 2005. LNCS, vol. 3776, pp. 786–791. Springer, Heidelberg (2005)
3. Bach, K., Reichle, M., Althoff, K.D.: A domain independent system architecture for sharing experience. In: Proceedings of LWA 2007, Workshop Wissens- und Erfahrungsmanagement, September 2007, pp. 296–303 (2007)

4. Börner, K., Pippig, E., Tammer, E.C., Coulon, C.H.: Structural similarity and adaptation. In: Smith, I.F.C., Faltings, B. (eds.) EWCBR 1996. LNCS, vol. 1168, pp. 58–75. Springer, Heidelberg (1996)
5. Bergmann, R.: Experience Management: Foundations, Development Methodology, and Internet-Based Applications. LNCS, vol. 2432. Springer, Heidelberg (2002)
6. Bergmann, R., Althoff, K.D., Minor, M., Reichle, M., Bach, K.: Case-based reasoning - introduction and recent developments. Künstliche Intelligenz: Special Issue on Case-Based Reasoning 23(1), 5–11 (2009)
7. Bach, K.: docquery - a medical information system for travellers. Internal project report (September 2007)
8. Reichle, M., Bach, K.: Improving result adaptation through 2-step retrieval. In: Nalepa, G.J., Baumeister, J. (eds.) Proceedings of the 4th Workshop on Knowledge Engineering and Software Engineering (KESE 2008) at the 31st German Conference on Artificial Intelligence (KI 2008), September 2008, pp. 73–84 (2008)
9. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Communications 1(7) (March 1994), `ftp://ftp.ifi.ntnu.no/pub/Publikasjoner/vitenskaplige-artikler/aicom-94.pdf` (letzte Verifikation 11. June 2007)
10. McGinty, L., Smyth, B.: Adaptive selection: An analysis of critiquing and preference-based feedback in conversational recommender systems. Int. J. Electron. Commerce 11(2) (06-7), 35–57
11. McCarthy, K., Reilly, J., McGinty, L., Smyth, B.: Experiments in dynamic critiquing. In: IUI 2005: Proceedings of the 10th international conference on Intelligent user interfaces, pp. 175–182. ACM, New York (2005)
12. Holte, R.C.: Very simple classification rules perform well on most commonly used datasets. Machine Learning, 63–91 (1993)
13. MacDonald, C., Weber, R., Richter, M.M.: Case base properties: A first step. In: Schaaf, M. (ed.) ECCBR Workshops, pp. 159–170 (2008)
14. O'Sullivan, D., Wilson, D.C., Smyth, B.: Improving case-based recommendation: A collaborative filtering approach. In: Craw, S., Preece, A.D. (eds.) ECCBR 2002. LNCS, vol. 2416, pp. 278–291. Springer, Heidelberg (2002)
15. Díaz-Agudo, B., González-Calero, P.A.: A declarative similarity framework for knowledge intensive CBR. In: Aha, D.W., Watson, I. (eds.) ICCBR 2001. LNCS, vol. 2080, pp. 158–172. Springer, Heidelberg (2001)
16. Dubois, D., Esteva, F., Garcia, P., Godo, L., Mántaras, R.L.d., Prade, H.: Case-based reasoning: A fuzzy approach. In: IJCAI 1997: Selected and Invited Papers from the Workshop on Fuzzy Logic in Artificial Intelligence, pp. 79–90. Springer, Heidelberg (1999)
17. Tawfik, A.Y., Kasrin, N.: Integrating causal knowledge in case-based retrieval: Causal decomposition of cases. In: Petridis, M., Wiratunga, N. (eds.) Proceedings of the Thirteenth UK Workshop on Case Based Reasoning (UKCBR 2008). University of Cambridge, UK (2008)
18. Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers Inc., San Francisco (1988)
19. Xiang, Y., Poole, D., Beddoes, M.P.: Multiply sectioned bayesian networks and junction forests for large knowledge-based systems. Computational Intelligence 9, 171–220 (1993)
20. Racine, K., Yang, Q.: On the consistency management of large case bases: the case for validation. In: Proceedings of the AAAI 1996 Workshop on Knowledge Base Validation, American Association for Artificial Intelligence (AAAI), pp. 84–90 (1996)

21. Redmond, M.: Distributed cases for case-based reasoning: Facilitating use of multiple cases. In: AAAI, pp. 304–309 (1990)
22. Kolodner, J.L.: Retrieving events from a case memory: A parallel implementation. In: Proc. of a Workshop on Case-Based Reasoning, Holiday Inn, Clearwater Beach, FL, pp. 233–249 (1988)
23. Richter, M.M.: Introduction. In: Lenz, M., Bartsch-Spörl, B., Burkhard, H.-D., Wess, S. (eds.) Case-Based Reasoning Technology. LNCS (LNAI), vol. 1400, p. 1. Springer, Heidelberg (1998)
24. Weibelzahl, S.: Conception, implementation, and evaluation of a case based learning system for sales support in the internet. Master's thesis, Universität Trier (1999)
25. Weibelzahl, S., Weber, G.: Benutzermodellierung von Kundenwünschen durch Fallbasiertes Schliessen. In: Jörding, T. (ed.) Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen, ABIS 1999, Magdeburg, pp. 295–300 (1999)
26. Cunningham, P., Bonzano, A., Smyth, B.: An incremental case retrieval mechanism for diagnosis (1995)
27. Cunningham, P., Smyth, B., Bonzano, A.: An incremental retrieval mechanism for casebased electronic fault diagnosis (1998)
28. Jurisica, I., Glasgow, J., Mylopoulos, J.: Incremental iterative retrieval and browsingfor efficient conversational CBR systems. Applied Intelligence 12(3), 251–268 (2000)