An Analysis of Current Trends in CBR Research Using Multiview Clustering

Derek Greene, Jill Freyne, Barry Smyth, and Pádraig Cunningham

■ *The European Conference on Case-Based Reasoning (CBR) in 2008 marked 15 years of international and European CBR conferences where almost seven hundred research papers were published. In this report we review the research themes covered in these papers and identify the topics that are active at the moment. The main mechanism for this analysis is a clustering of the research papers based on both cocitation links and text similarity. It is interesting to note that the core set of papers has attracted citations from almost three thousand papers outside the conference collection so it is clear that the CBR conferences are a subpart of a much larger whole. It is remarkable that the research themes revealed by this analysis do not map directly to the subtopics of CBR that might appear in a textbook. Instead they reflect the applications-oriented focus of CBR research, and cover the promising application areas and research challenges that are faced.*

In 2008 the international series of conferences on case-based reasoning (CBR) celebrated their fifteenth anniversary. Each year since 1993 there has been an international or European conference on CBR. Up to 2007, this conference series produced 672 papers in all. In this report we examine the research themes evident in these papers and identify the most active research topics in CBR.

At the 2008 conference we presented an analysis of the research themes in CBR, based on an analysis of the cocitation links in the research literature (Greene et al. 2008). That analysis was based on the core set of 672 papers from the CBR conferences with cocitation data coming from a set of 3461 papers that cite these papers (details on how cocitation links are determined are given later in the article). While cocitation analysis has been proven to be very effective at uncovering relational structure in the research literature (White and Griffith 1981), it has the shortcoming that recent papers will have few cocitation links as papers citing pairs of papers in the core set (that is, the source of cocitation links) have not yet appeared. This issue is evident in the plot of citation counts shown in figure 1 and ultimately makes it impossible to recognize the influence of more recent papers.

We improve on this analysis here by integrating a new source of relational data based on text similarity with the existing coc-

 ISSN 0738-4602

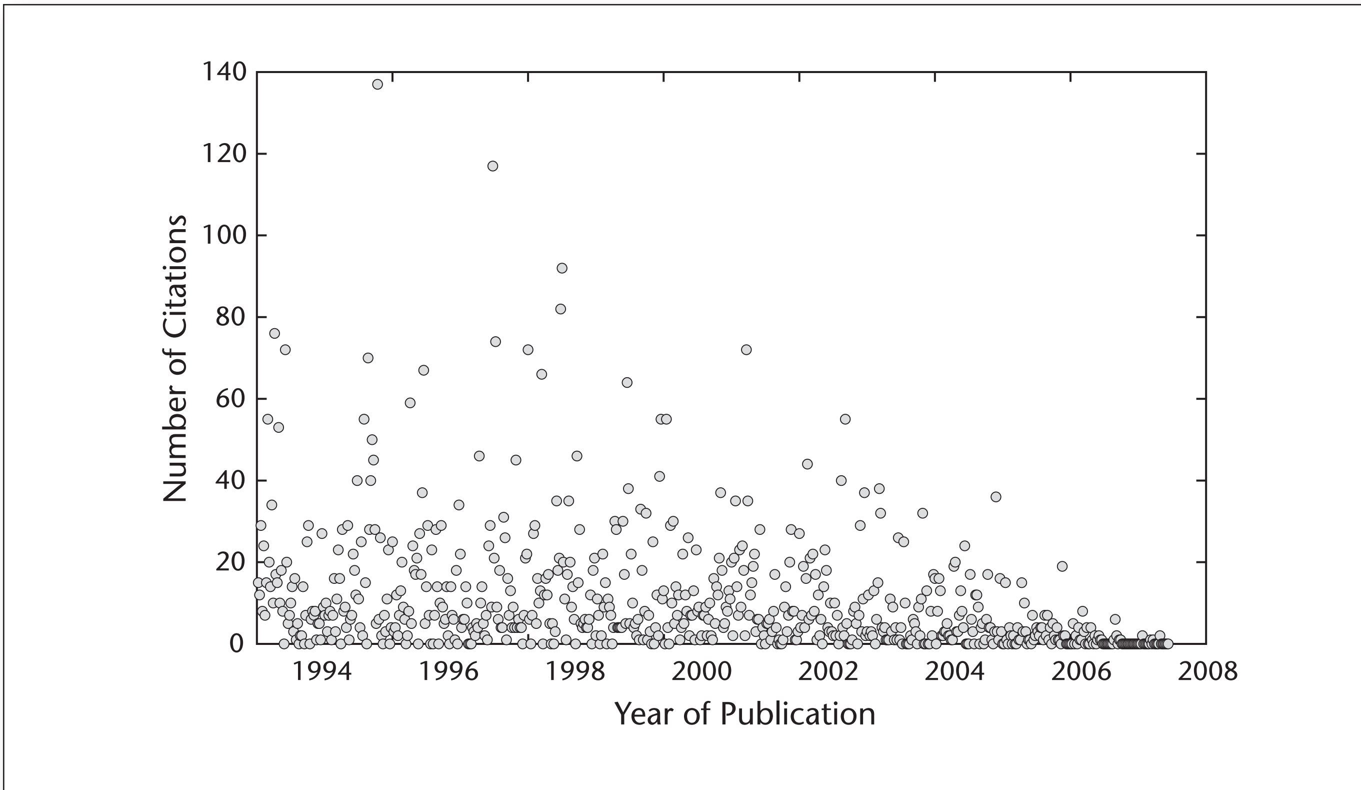

Figure 1. A Plot of the Citation Counts for Papers.

It is clear from this plot that citation and cocitation data contains little information about recent papers.

itation data in order to provide a more comprehensive picture of the research themes in the CBR literature. The evaluation in the research themes section shows that incorporating the text similarity view meets this objective of bringing very recent papers into the clustering process. The text view also allows older papers that did not attract citations (and thus do not have significant cocitation links) into the clustering. Whether this is always desirable is debatable, and it raises interesting questions about the significance of the research themes that have been identified. In the analysis based on cocitation links only, we can be confident that research themes that did emerge were based on a significant citation structure. It might be argued that a set of papers on an identifiable research theme that is not supported by a network of citations does not have the same status. On the other hand such themes may lie dormant for some time, becoming relevant at some future time when conditions are right—this is the case for the theme on explanation discussed in the new themes revealed by PICA section.

The data on which this analysis is based is described in the next section. The results of our initial analysis based on cocitation analysis are then summarized. The alternative views on the data that are used for multiview clustering are described in the data views section. Then the challenges of multiview clustering and the approach that we use are described in the multiview clustering section, and the research themes that have been identified are discussed in the research themes section.

The Data

Since the conception of the CBR conference series (ECCBR/ICCBR/EWCBR) in 1993, a total of 672 papers have been published by 828 individual authors. Data on these papers was gathered from the Springer online bibliographies[1] for each of the annual conference proceedings. These bibliographies are available in the form of RIS files, a tagged file format for expressing citation information, including details such as the issue title, paper titles, author lists, and abstracts for each publication in the conference series.

To determine the connections within the network of CBR publications, we submitted queries to Google Scholar[2] to retrieve the list of papers refer-

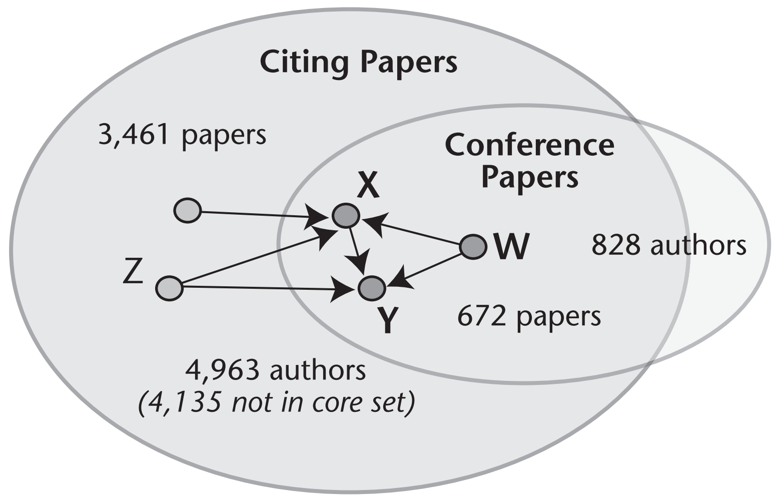

Figure 2. The Analysis Centers on the Core Set of 672 Papers Published in the CBR Conferences.

The cocitation links are revealed in a set of 3461 papers that cite these papers.

Conference	No. Papers	Maximum	Mean	Median
ECCBR	305	92	11.01	6
ICCBR	367	137	10.14	5

Table 1. Citation Statistics for the 672 Conference Papers.

encing each of the 672 "seed" papers. Each list contains all of the Google-verified citations that a given paper had received at query submission time (December 2007). In total 7078 relevant citation links were recorded. Note that, while citation information from the supplementary (that is, non-seed) set of papers was used to provide additional information regarding cocitations, only the 672 seed papers and their associated authors were considered as data objects in our analysis.

It is interesting to observe that of the 7078 citation links, only 1216 were internal cites with 5862 coming from papers outside the conference papers. While there are 828 authors represented in the core set of papers, there are 4135 authors in the wider set of papers making 4963 authors in all (see figure 2). This shows that the CBR conference papers are a small part of a very large research activity—while there are 672 papers in the core set there are 3461 "citing" papers.

Some citation statistics for the conference papers are shown in table 1. In all, 549 papers received citations and the total number of citations found for the collection is 7077. The most cited paper is titled "Weighting Features" by Wettschereck and Aha (1995), which at the time the data was collected had 137 citations. The overall mean number of citations is 10.5 and the overall median is 5. This is a very respectable number for a conference series. In another analysis comparing impact across a number of artificial intelligence and machine learning conferences, this was found to compare favorably with conferences such as European Conference of Artificial Intelligence and European Conference of Machine Learning (Coyle et al. 2008).

A Review of the Results of the Citation-Only Analysis

The analysis of case-based reasoning research themes presented in Greene et al. (2008) was based on cocitation analysis only. A cocitation link exists between two papers if they are both cited by a third paper (see the later discussion of the cocitation

No.	Paper Title	Authors	Year	Citations	Cocites
1	Weighting Features	Wettschereck and Aha	1995	137	522
2	Modeling the Competence of Case Bases	Smyth and McKenna	1998	92	525
3	Refining Conversational Case Libraries	Aha and Breslow	1997	117	518
4	Maintaining Unstructured Case Bases	Racine and Yang	1997	72	469
5	Using Introspective Learning to Improve Retrieval in CBR: A Case Study in Air Traffic Control	Bonzano, Cunningham, and Smith	1997	74	473
6	Similarity Versus Diversity	Smyth and McClave	2001	72	452
7	Building Compact Competent Case Bases	Smyth and McKenna	1999	64	399
8	Categorizing Case-Based Maintenance: Dimensions and Directions	Leake and Wilson	1998	82	322
9	Diversity-Conscious Retrieval	McSherry	2002	44	362
10	Similarity Measures for Object-Oriented Case Representations	Bergmann and Stahl	1998	66	403

Table 2. A Ranked List of the Top 10 Papers in the Overall Collection Based on Eigenvector Centrality.

The total number of citations and cocitations for these papers is also shown.

view for details). The results of this early analysis confirmed how contemporary CBR research has evolved from the early years of the field. Strong clusters of activity in contemporary research include the likes of recommender systems and diversity, textual CBR, case-base maintenance, and conversational CBR, which are characteristic of modern CBR research. It was interesting that many of the more traditional research themes did not feature prominently in the clusters of research that emerged from our analysis. For example, the traditional themes of *representation and indexing, analogy, architectures*, and *design and planning* are conspicuous by their absence, and even critical areas of research such as *adaptation* or *similarity and retrieval* have either become less active or have fundamentally changed their emphasis. It is also encouraging to note that new themes can and do emerge (for example, recommender systems and diversity), and that research activity in an area can wind down (for example, case-base maintenance), as it matures to deliver effective solutions to the community. We concluded then that this could be considered a sign of a healthy research area.

Prominent Papers: Centrality and Citation Count

Given that the main findings in the initial analysis entail a clustering of the papers based on cocitation links, it was interesting to see which papers are most "central" to the overall collection based on these cocitation links. Following the literature on centrality in social network analysis, we selected eigenvector centrality and degree centrality as appropriate measures for this exercise (Wasserman and Faust 1994). Table 2 shows the top 10 papers ranked by eigenvector centrality. This table also shows a count of cocitations for these papers—this corresponds to degree centrality and correlates well with eigenvector centrality. A further ranked list with papers ranked by raw citation count is shown in table 3. The evidence from these tables is that the most important paper in the collection is "Weighting Features" (Wettschereck and Aha 1995). These two lists of prominent papers are useful in that they do appear to encapsulate the main themes in CBR research over the last 15 years.

Prominent Research Themes

The main result of the initial analysis was the identification of 14 research themes that were evident in the cocitation structure—see table 4. For the most part these themes are still evident in the clustering based on both views (text and cocitation). For instance, figure 4 shows a cluster of papers relating to recommender systems that corresponds closely to one uncovered in the original analysis—the balance of green and blue bars in the panel on the right of the screenshot indicates that this cluster is supported in both the text and cocitation views. However, the themes of "CBR on Temporal Problems" and "Scheduling and Agents" that were previously considered minor are now more prominent as they have good support in the text view.

This previous study provides a starting point for the work presented in this paper. The cocitation-based analysis offers a single view of the mainstream CBR research literature, and it is not without its shortcomings. At the very least, relying on

No.	Paper Title	Authors	Year	Citations
1	Weighting Features	Wettschereck and Aha	1995	137
2	Refining Conversational Case Libraries	Aha and Breslow	1997	117
3	Modeling the Competence of Case Bases	Smyth and McKenna	1998	92
4	Categorizing Case-Base Maintenance: Dimensions and Directions	Leake and Wilson	1998	82
5	Using k-d Trees to Improve the Retrieval Step in Case-Based Reasoning	Wess, Althoff, and Derwand	1993	76
6	Using Introspective Learning to Improve Retrieval in CBR: A Case Study in Air Traffic Control	Bonzano, Cunningham, and Smith	1997	74
7	Explanation-Driven Case-Based Reasoning	Aamodt	1993	72
8	Maintaining Unstructured Case Bases	Racine and Yang	1997	72
9	Similarity Versus Diversity	Smyth and McClave	2001	72
10	Cases as Terms: A Feature Term Approach to the Structured Representation of Cases	Plaza	1995	70

Table 3. A Ranked List of the Top 10 Papers in the Overall Collection Based on Total Citation Count.

No.	Major Themes	Prominent Papers
1	Recommender Systems and Diversity	Bridge and Ferguson 2002; Doyle and Cunningham, 2000; Goker and Thompson 2000; McGinty and Smyth 2003; McSherry 2003; Mougouie, Richter, and Bergmann 2003; Smyth and McClave 2001
2	Case-Base Maintenance	Aha and Breslow 1997; Heister and Wilke 1998; Muñoz-Avila 1999; Portinale, Torasso, and Tavano 1999; Racine and Yang 1997; Reinartz, Iglezakis, and Roth-Berghofer 2000; Smyth 2000; Smyth and McKenna 1998; Surma and Tyburcy 1998
3	Case Retrieval	Cunningham, Doyle, and Loughrey 2003; Doyle et al. 2004; Gabel and Stahl 2004; Lenz, Burkhard, and Bruckner 1996; McSherry 2004; Osborne and Bridge 1996, 1997; Schaaf 1996; Smyth and McKenna 1999
4	Learning Similarity Measures	Bradley and Smyth 2004; Gabel and Stahl 2004; Gomes and Bento 2000; Hayes et al. 2005; Stahl 2005; Stahl and Gabel 2003
5	Adaptation	Bandini and Manzoni 2001; McSherry 1998; Neagu and Faltings 2001, 2003; Tonidandel and Rillo 2005
6	Image Analysis	Grimnes and Aamodt 1996; Macura and Macura 1995; Perner 1999
7	Textual CBR	Bruninghaus and Ashley 1997, 2001; Gu and Aamodt 2005; Gupta, Aha, and Sandhu 2002; Lamontagne and Lapalme 2004; Wiratunga, Koychev, and Massie 2004
8	Conversational CBR	Aha, Maney, and Breslow 1998; Doyle and Cunningham 2000; Goker and Thompson 2000
9	Feature Weighting and Similarity	Bonzano, Cunningham, and Smyth 1997; Faltings 1997; Jarmulak, Kerckhoffs, and Van't Veen 1997; Netten and Vingerhoeds 1995; Stahl 2001, 2005, 2006; Trott and Leng 1997; Vollrath 2000; Wettschereck and Aha 1995
10	Creativity and Knowledge-Intensive CBR	Armengol and Plaza 1994; Bunke and Messmer 1993; Kolodner 1993; Lluís Arcos and Plaza 1993; Nakatani and Israel 1993; Richards 1994; Sebag and Schoenauer 1993; Smyth and Keane 1993
	Minor Themes	
11	CBR on Temporal Problems	Jære, Aamodt, and Skalle 2002; Nakhaeizadeh 1993
12	Games and Chess	Flinter and Keane 1995
13	Scheduling and Agents	Macedo and Cardoso 2004
14	Structural Cases	Borner et al. 1996

Table 4. Research Themes Identified in the CBR Conference Literature Based on the Cocitation View Only.

cocitations as the basic unit of structure necessarily limits our analysis to those research works that have been successful at attracting citations, obscuring from view those research efforts that have yet to amass a critical citation history. Recognizing these clusters can help to reveal dormant, latent, and emergent research, and it is for this reason that we seek to extend this previous analysis by allowing a text-based approach to complement the cocitation approach to provide a more comprehensive view of modern CBR research.

In analyzing a body of research literature in order to identify research themes, there are a number of perspectives that can be taken on the data. The most fundamental decision to be made is whether to search for an informative organization of authors or research papers. In the initial analysis of the CBR corpus (Greene et al. 2008), we found that clustering papers was more informative than clustering authors, presumably because it is a reasonably compact research field with some authors participating in a number of research themes. The initial analysis was based on a cocitation perspective on the papers; this is extended here by also considering a view based on text similarity.

Cocitation View

The most fundamental representation used to model scientific literature in bibliometrics is the unweighted directed citation graph, where an edge exists between the paper P_i and the paper P_j if P_i cites P_j. This graph can be represented by its asymmetric adjacency matrix **A**. However, it has been established in bibliometrics research that cocitation information can be more effective in revealing the true associations between papers than citations alone (White and Griffith 1981).

The concept of cocitation analysis is illustrated in figure 2 where an arrow from paper *X* to paper *Y* indicates that paper *X* cites paper *Y*. A direct analysis of citation shows for instance that *X* is related to *Y*. However, the fact that *X* and *Y* are both cited by *Z* and *W* indicates a strong relationship between these papers. Cocitation has the potential to reveal indirect associations that are not always explicit in the citation graph. In addition, it can bring information from outside the collection (*Z* is not one of the core papers) to bear on the analysis.

Consequently, a network of publications is often represented by its weighted undirected cocitation graph. This graph has a symmetric adjacency matrix defined by $C = A^T A$, where the off-diagonal entry C_{ij} indicates the number of papers jointly citing both P_i and P_j. Note that the entry C_{ii} on the main diagonal corresponds to the total number of citations for the paper P_i.

Rather than using raw cocitation values in **C** as a basis for measuring the similarity between papers, a variety of normalization strategies have been proposed in the area of bibliometrics (He and Cheung Hui 2002). The *CoCit-Score*, proposed by Gmür (2003), has been shown to be a particularly effective choice for clustering cocitation data. This measure computes the association between a pair of papers (P_i, P_j) by normalizing their cocitation frequency with respect to the minimum and mean of the pair's respective citation counts as follows:

$$S_{ij} = \frac{C_{ij}^{2}}{\min\left(C_{ii}, C_{jj}\right) \times mean\left(C_{ii}, C_{jj}\right)}$$

Each entry S_{ij} is in the range [0,1], where a larger value is indicative of a stronger association between a pair of papers. At the time the data set was constructed, 518 of the core CBR papers had accrued at least one citation according to Google Scholar, thus yielding an "incomplete" cocitation view.

Text Similarity View

In addition to the information provided by citation links, the availability of paper titles and abstracts in the RIS format allowed us to construct an alternative view of the seed papers in the form of a "bag-of-words" text representation. This text representation was available for all 672 seed papers, although the resulting vector space model is highly sparse, with only 1949 unique terms occurring in more than one document after standard stemming and stop-word removal techniques were applied. Similarity values between the term vectors were computed by finding the cosine of the angle between their respective term vectors. This provided the second view that was used in the multiview clustering process. A key goal of the process was to produce a superior model of the CBR research network from these two "deficient" views.

Multiview Clustering

The challenge of integrating multiple perspectives on a problem to offer a more complete picture arises in a variety of contexts. In the work described here there is a significant degree of discord between different views so we employ a system called PICA (Parallel Integration Clustering Algorithm) that can bring together multiple potentially discordant views in an unsupervised learning framework (Greene, Bryan, and Cunningham 2008). For instance, in bibliographic networks certain papers may share several common terms in their abstract text but may never have been cocited together in a single paper. This is further complicated by the fact that cocitation relationships generally do not begin to reveal themselves for sev-

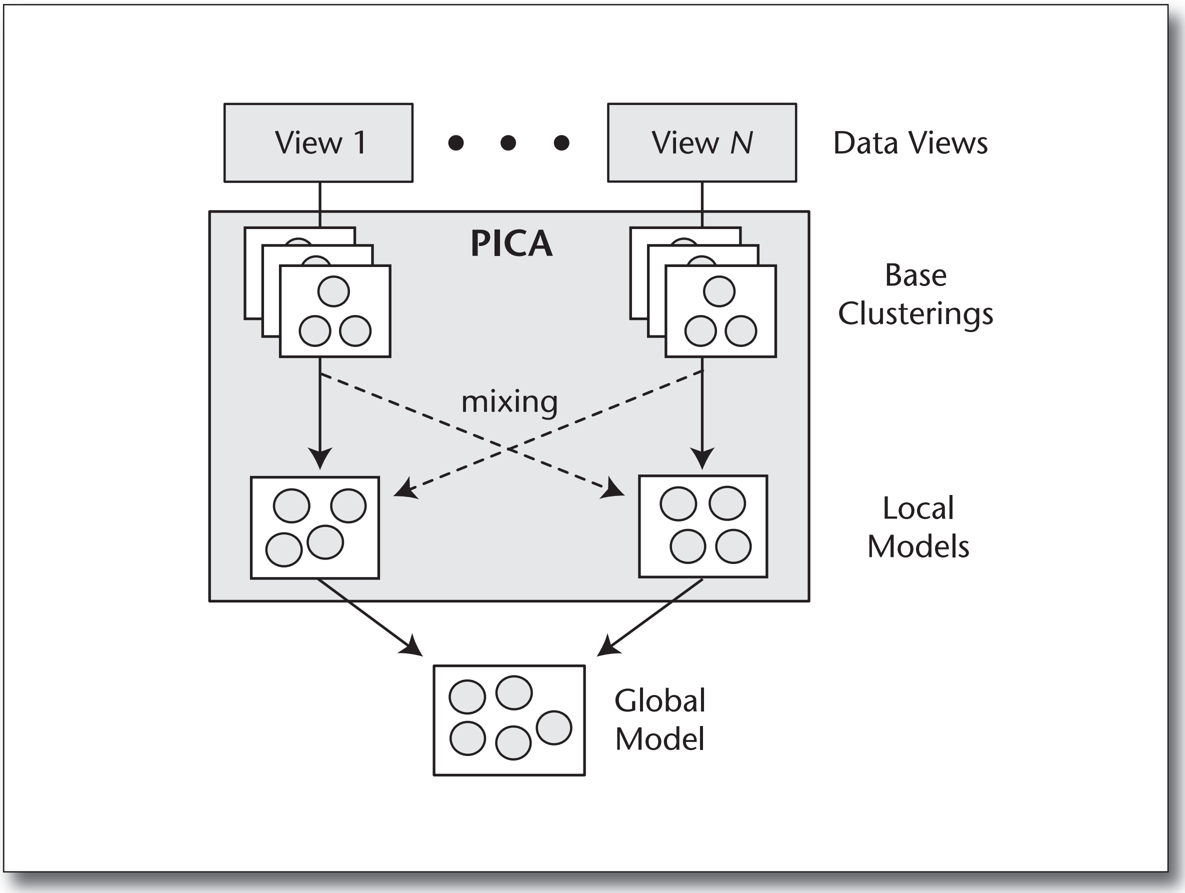

Figure 3. Overview of the Parallel Integration Clustering Algorithm (PICA).

eral years until papers begin to accrue citations. To deal with such cases, PICA has been developed based on the *parallel universe* (PU) framework for clustering presented by Wiswedel and Berthold (2007). The PU concept emphasizes the idea of sharing information between views in order to learn superior *local models* for the views, which can subsequently be combined to provide a comprehensive *global model* of the patterns present in the domain. For us, a key aspect of the PU framework is that structures can exist in some views but not in others. Another important aspect of many real-world data fusion tasks is that the available data sources will often be *incomplete* in nature (that is, each source may represent a different subset of the complete set of data objects in the problem domain). This is taken into account by PICA, as the input views do not necessarily need to group all possible objects in the domain. Some level of overlap between the objects present in the views is sufficient.

PICA

Rather than working on the original data, PICA takes as its input a collection of "base clusterings" constructed independently on each available view. These will typically be generated by applying a standard partitional clustering algorithm that will frequently converge to different local minima under different starting conditions. On the CBR network data, we employed the kernelized form of the *k*-means algorithm (Schölkopf, Smola, and Müller 1998). In the case of the text data, we clustered on a cosine kernel. For the cocitation data we used a kernel based on the *CoCit-Score* given in equation 1.

Given this input, PICA follows a two-stage process. Firstly, PICA constructs a local model on each available view in the form of a "soft" clustering (that is, a clustering with nonnegative real-valued membership weights that allows the representation of overlaps between clusters). Secondly, PICA combines the local models to produce a global model (in the form of a soft clustering of all data objects in the domain). This model merges the common aspects of the local models, while preserving those clusters that are unique to each local model. The complete PICA algorithm is illustrated in figure 3.

Local Model Construction: To initialize the local

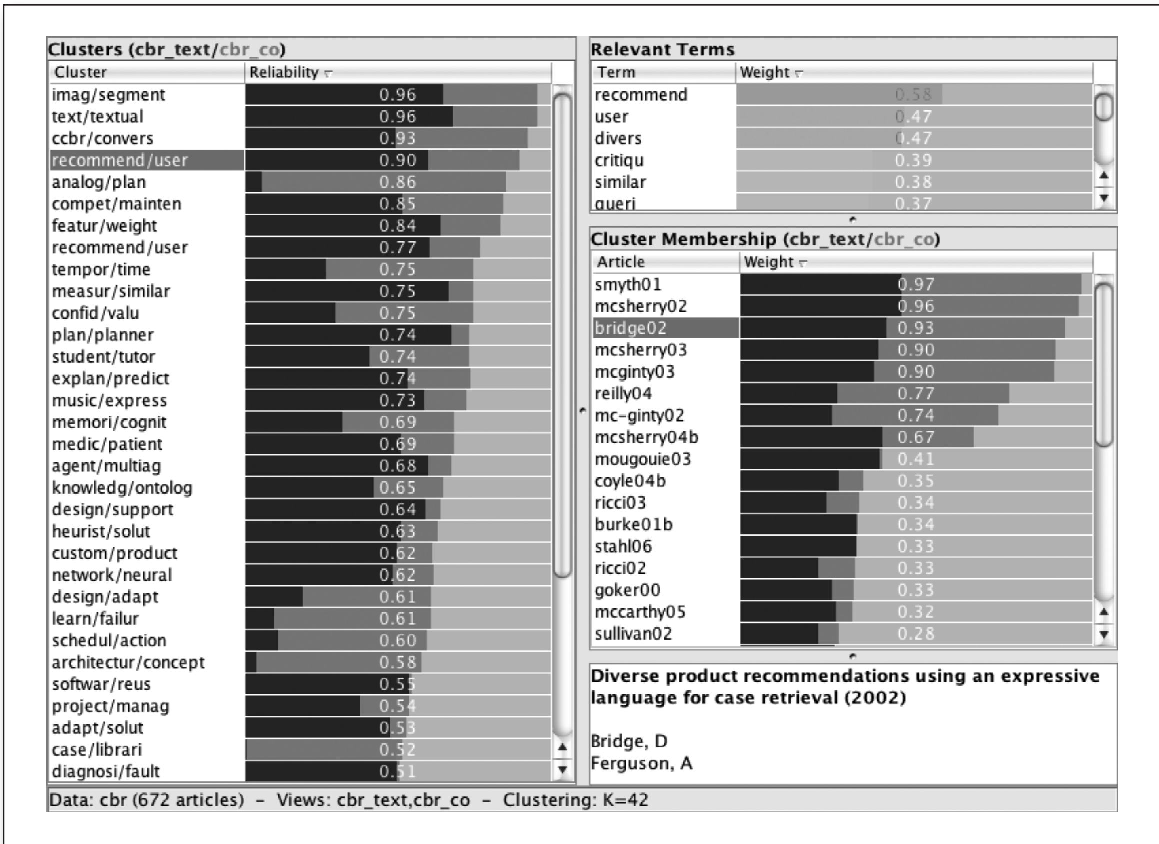

Figure 4. An Example of the Output of the PICA Browser Tool.

This shows a cluster of research on "Recommender Systems." The balance of green and blue in the panel on the bottom rights shows that this cluster is supported by both the text and cocitation view.

model for a given view, we select the most representative base clustering from the set of base clusterings generated on that view, using a measure of clustering "stability" based on pairwise average *normalized mutual information* (Strehl and Ghosh 2002). Next we attempt to improve our initial local model by adding information from the remaining base clusterings that were generated on *all* views. This has the effect of supporting "mixing" between the views, where information provided by a base clustering from one view can inform the model constructed for another view. In practice, the aggregation is performed by using a variation of the cumulative voting methods that have been previously proposed for efficiently combining an ensemble of clusterings (Dimitriadou, Weingessel, and Hornik 2002). We match the clusters in each base clustering with those in the current local model and merge these matched clusterings to update the local model. The optimal correspondence between clusters can be found by measuring the *binary overlap coefficient* similarity between pairs of clusters and solving the minimal weight bipartite matching problem. Note that "poorly matched" clusters (that is, pairs whose overlap similarity is below a user-defined threshold) are not included during mixing, reflecting the fact that structures in one view may not be present in another.

Global Model Construction: At this stage we have constructed a set of local models, one for each view. These may be of interest in their own right, but for ease of interpretation and evaluation, we would like to combine these partial models to produce a single global model providing a more complete picture of the domain. This is achieved by performing an additional matching procedure at this stage, where similar clusters from each local model are merged, so that redundant patterns are combined, while unique patterns are preserved. In practice, this can be done by performing complete-linkage agglomerative clustering on the local model clusters and choosing an appropriate cutoff level. The resulting global model is a soft clustering incorporating structures from all available views.

Model Visualization

To explore the models produced by PICA, includ-

Theme	Discriminating Terms
Confidence	confidence, value, solution, know, produce, expect, classify, estimate
Planning	plan, planner, route, analog, state, project, CBP, reformulate
Tutoring Systems	student, tutor, program, learn, skill, individual, concept, relevance
Explanation	explanation, predict, CBR, metric, outcome, explanation-based
CBR and Music	music, expression, perform, tempo, song, transform, phrase
CBR in Medicine	medicine, patient, care, health, expert, reason, therapy
Knowledge-Intensive CBR	knowledge, ontology, CBR, intensive, CBROnto, generation

Table 5. Research Themes Identified in the CBR Conference Literature Based on Both the Cocitation and Text Views.

ing the contributions made by each view to the models, we have developed the PICA Browser application.[3] An example of a cluster in a global model produced from the integration of two heterogeneous views is shown in figure 4. To highlight cluster provenance, the left side of the screenshot shows the list of clusters in the global model, with the blue/green bar showing the proportion of contribution coming from each view. Note that the clusters are arranged in descending order based on their *reliability* scores. These scores reflect the degree to which a cluster repeatedly appeared in the base clusterings across one or more views, and thus they quantify the prominence of a cluster in the research literature.

When one of the views under consideration is based on text data (such as the research abstracts available for the CBR conference series), we can use this data as a means of summarizing the content of the clusters generated by PICA for human inspection. As part of the PICA Browser interface, ordered lists of discriminating keywords are provided for each cluster (shown at the top right corner of figure 4). These keywords were automatically identified by ranking the terms for each cluster based on their Information Gain. Given a cluster of papers, the ranking of terms for the cluster is performed as follows: firstly the centroid vector of the cluster is computed on the text view; subsequently, we compute the Information Gain between the cluster centroid vector and the centroid vector for the entire set of papers. Terms that are more indicative of a cluster will receive a higher score, thereby achieving a higher ranking in the list of keywords for the cluster. Sample keywords for clusters generated by PICA are listed later in table 5.

Research Themes

An initial exploration of the thematic structure of the CBR conference literature has already been presented by Greene et al. in 2008. That analysis was based on cocitation links, an established technique for identifying relationships between research papers. Since cocitation data has the shortcoming that it cannot identify relationships between very recent papers or between those papers that are poorly cited, we extend that analysis by incorporating another view that is based on the similarity between the text of publication titles and abstracts.

The complete CBR conference literature network dataset[4] consists of 672 papers published by 828 individual authors. At the time the dataset was constructed (December 2007) 518 of these papers had accrued at least one citation according to Google Scholar, yielding an incomplete cocitation view. A text representation was available for all 672 papers, although the resulting vector space model was highly sparse, with only 1949 nonstopword terms occurring in more than one document. The goal of our evaluation was to take these two "deficient" views and use PICA to produce a superior model of the CBR research network.

New Themes Revealed by PICA

We now examine seven research themes revealed by the multiview analysis that were not evident in the original analysis performed on cocitation data only. These themes and the discriminating terms associated with them are shown in table 5. The interrelationships between these seven clusters can be seen in figure 5. Explanation is placed in the center by the graph-drawing algorithm because it is well connected to many of the other research themes, for example, recommender systems, planning, tutoring, and textual CBR. It is interesting to note the connections between tutoring and planning (after all tutoring has a significant planning component) between conversational CBR and recommender systems and between CBR in medicine and image analysis.

Confidence: This new cluster on confidence in CBR is a testament to the merits of including the text view in the clustering process (see figure 6). The most representative paper in this cluster is the paper by Cheetham and Price (2004), "Measures of

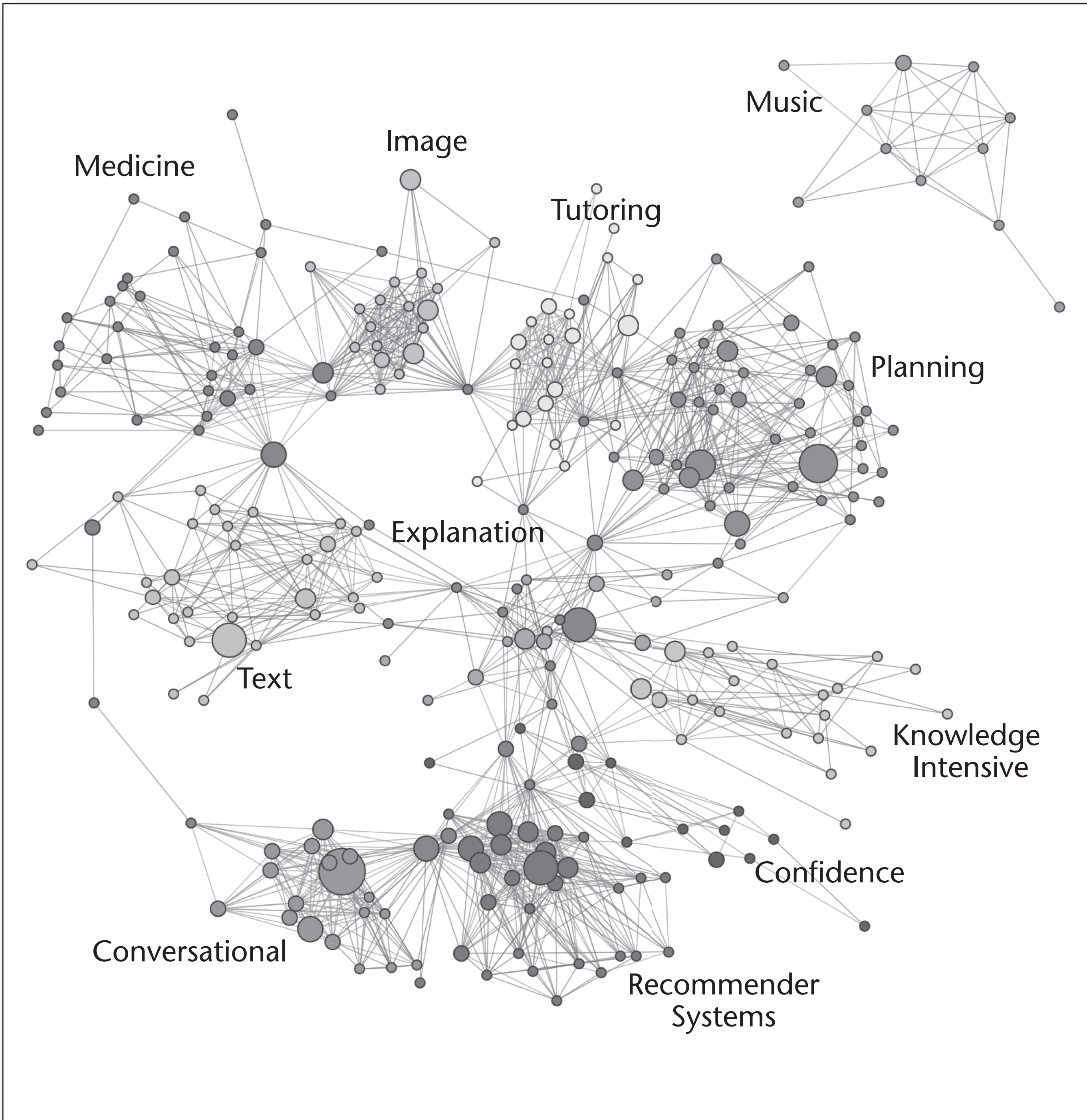

Figure 5. A Graphical Representation of Seven New Themes Revealed by the Multiview Analysis, Together with Four of the Original Applications-Oriented Themes.

Each node in the graph represents an individual paper. Red nodes denote papers that belong to more than one theme. The blue and green edges denote strong connections from the text similarity and cocitation views, respectively. Red edges indicate strong connections apparent in *both* views.

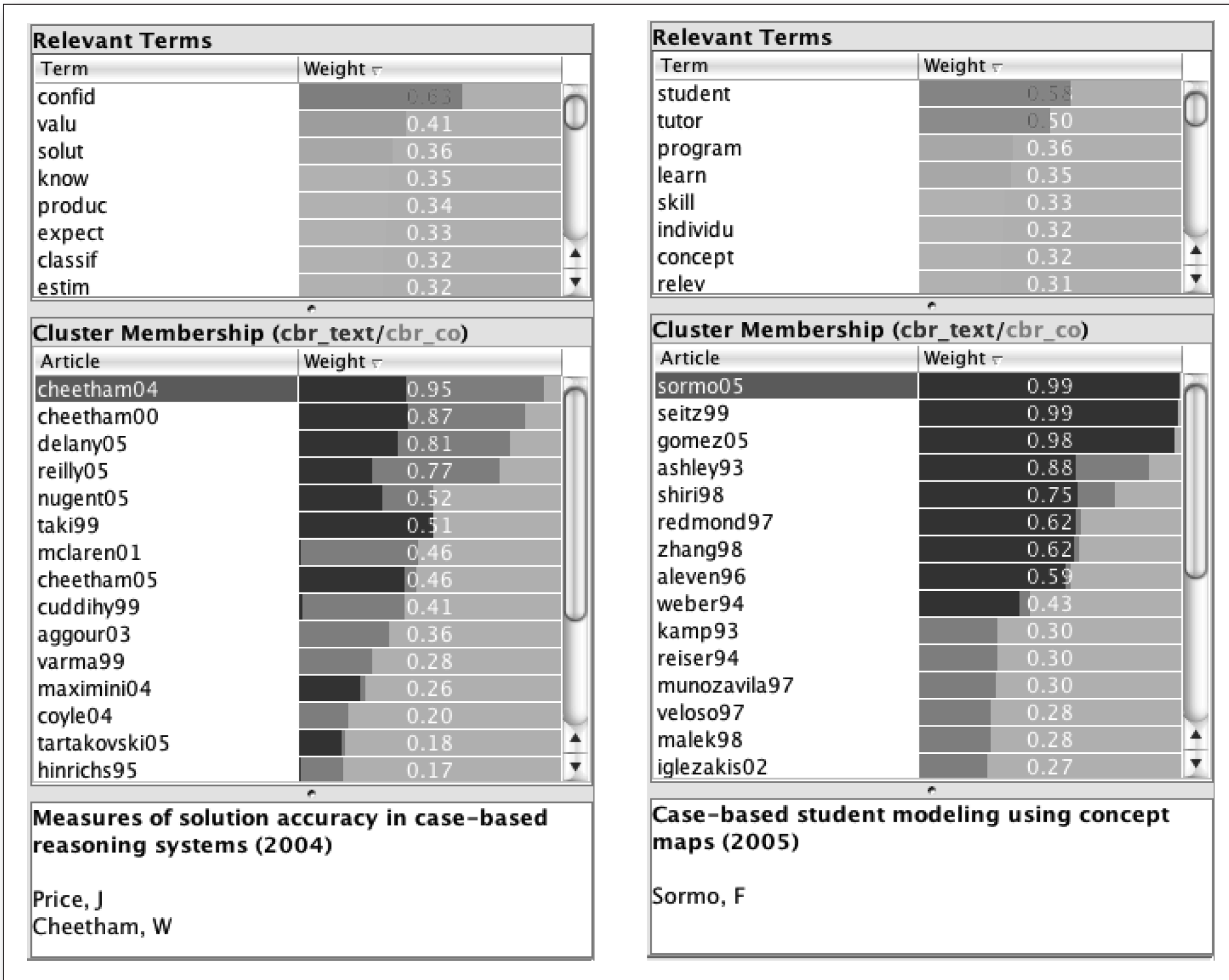

Figure 6. Two Further Examples of the Output of the PICA Browser Tool.

On the left is the cluster on "confidence," which has good support from the cocitation backed up by evidence from the text view. The cluster on the right covers research on "tutoring systems"—most of the evidence for this cluster comes from the text view.

Solution Accuracy in Case-Based Reasoning Systems." This paper is representative of a body of recent research activity on quantifying and predicting the reliability of solutions proposed by CBR systems. Many of the papers in the cluster are from 2005. These papers have picked up some citations already but not enough to form a clear cluster based on cocitation links only. However, the addition of the text view reveals a strong research theme with a lot of recent research activity.

Planning: The combined text and cocitation analysis reveals a cluster on planning that is supported almost exclusively from the text view. One paper in this cluster that is supported by some cocitation structure is the paper by Mukkamalla and Muñoz-Avila (2002), "Case Acquisition in a Project Planning Environment." Discriminating terms to describe this cluster are *plan, planner, route, analogy, state, project, CBP,* and *reformulate*. This indicates a cluster of papers on case-based planning (CBP) with a focus on applications in route planning and project planning. The absence of strong cocitation support for this cluster is probably explained by the fact that there are other clusters on related areas such as analogical reasoning and scheduling that contain papers on CBP. From a cocitation perspective CBP is strongly connected with analogical reasoning and scheduling and thus does not show up as an independent cluster when textual similarity is not considered.

Tutoring Systems: This cluster is evident in the combined view but not in the cocitation view (see figure 6) because the three most prominent papers in the cluster do not show up in the cocitation structure (Sørmo 2005, Seitz 1999, Gómez-Martín et al. 2005). It is not surprising that the papers by Gómez-Martín et al. and Sørmo do not show up in the cocitation structure, as they are recent papers. However, it is surprising that the paper by Seitz is absent as it is a frequently cited paper. The explanation appears to be that much of the work on CBR and tutoring is published outside the CBR conferences, and consequently the cocitation structure within the CBR conference literature is weak.

Explanation: This theme was already evident in the original analysis as a subtheme of case retrieval—retrieving cases to support retrieval is a recognized research issue in CBR. However, the text view brings in a few recent papers from 2005 to 2007, and this research theme is more evident in the multiview analysis. The most representative paper for this cluster is the invited talk from EWCBR'06 (Rissland 2006), titled "The Fun Begins with Retrieval: Explanation and CBR." The temporal distribution of papers in this cluster is bimodal with a number of papers appearing in the early days of the conference series in 1993 and 1994 and another concentration of activity in more recent years. The early papers (for example, Aamodt [1993], Bento et al. [1994]) report work on knowledge-intensive explanation while some of the more recent papers (Cunningham, Doyle, and Loughrey 2003; McSherry 2004) represent a knowledge-light approach.

CBR and Music: This research theme covers the use of CBR in music, with many of the papers having a creative or performance focus. There is some support for this theme in the cocitation view, but this support is not strong as many of the papers are from 2004 and later. This is a good example of the benefits of incorporating the text view as it reveals newer research themes that are not yet supported by cocitation. One of the top papers in this cluster (Baccigalupo and Plaza 2007) is slightly atypical because, while it is about CBR and music, it concerns song scheduling in music radio, whereas most of the papers in this research theme are concerned with performance (Tobudic and Widmer 2003; Grachten, Arcos, and de Mántaras 2004). It is interesting to note in the network diagram in figure 5 that the CBR and Music cluster is detached from the rest of the network. This supports the impression that this strand of CBR research is quite distinct.

CBR in Medicine: In the analysis based on cocitations only it was remarkable that applications of CBR in medicine did not emerge as a research theme, as this would be recognized as an application area for CBR where there is a significant amount of research activity. This theme is clearly evident in the multiview analysis, with contributions coming from both the text and cocitation views. It may be that the reason this did not show up in the original analysis is that much of this research is published outside the CBR conference series, and thus this theme does not have a strong signature in the available citation data. The most typical paper in this theme is that by Marling and Whitehouse (2001) on Alzheimer's care. Some papers with strong support from both the text and cocitation perspectives are also present (Opiyo 1995; Schmidt, Pollwein, and Gierl 1999; Montani et al. 2000).

Knowledge-Intensive CBR: The final cluster we choose to highlight is concerned with research on knowledge-intensive CBR, much of which is quite recent. The central papers in this cluster describe innovations around the *jColibri* CBR development environment, which is well suited for knowledge-intensive CBR (Díaz-Agudo and González-Calero 2001; Díaz-Agudo, Gervás, and González-Calero 2002). There are a number of other papers in this cluster that do not refer to jColibri but reflect independent CBR research with a knowledge-intensive focus (Kamp 1997; Bergmann and Mougouie 2006).

Other Themes: There are a number of other themes that can be identified in the PICA output. For instance, there is a theme on "Web Search" that contains a number of recent papers on CBR in Internet search. There are also identifiable clusters on the more established themes of "Software Reuse" and "Failure-Driven Learning." The papers in these clusters can be examined by downloading the PICA Browser tool and exploring the models generated on the CBR data.

Conclusion

Case-based reasoning research has its origins in the pioneering work of a number of researchers in the mid to late 1980s (Rissland, Valcarce, and Ashley 1984; Hammond 1986; Kolodner 1991; Schank and Leake 1989; Carbonell et al. 1991; Stanfill and Waltz 1986). These early researchers shared an interest in the role that experiences played in human problem solving and machine reasoning, and their early work represents the starting point for modern case-based reasoning research in which the capture and reuse of experiential problem solving plays a key role in intelligent systems design. This early research emphasized the foundations of case-based reasoning: case representation; similarity and case retrieval; solution adaptation, case learning, the CBR process model, and so on. Some 20 years on, case-based reasoning research continues to mature as ongoing basic research complements significant application success stories.

The 2008 European Conference on Case-Based Reasoning marked 15 years of international and European case-based reasoning conferences. These conference series alone have captured some 700 papers providing a comprehensive and coherent representative sample of evolving CBR research. This body of literature provides an excellent opportunity to review the development of CBR research and the evolution of this field's key research themes. Thus, the work presented by Greene et al. (2008) described an initial bibliometric analysis of CBR research themes, based on the cocitation structure that underpins a collection of more than 3000 CBR papers. The results confirmed that modern CBR research is characterized by a set of research themes that are significantly different from those present during the early years of the field. Classical themes such as case representation, similarity and retrieval, adaptation, and learning, while still evident, are overshadowed by stronger clusters of activities in areas such as recommender systems and diversity, textual CBR, case-base maintenance, and conversational CBR.

This original analysis is incomplete, however, and the focus on cocitation structure, while well motivated by the literature, means that it is unlikely to capture the influence of more recent papers, which have yet to attract a critical mass of citations. To this end, in this work we have extended this pure bibliometric approach by using multiview clustering techniques to integrate a new source of relational data, based on text similarity, as a way to provide a more comprehensive picture of contemporary CBR research. This new analysis has served a number of purposes. First of all, the results of the text-based clustering add support to our previous cocitation-based clusters, with prominent cocitation themes also featuring within the text-based view. More importantly perhaps, the text-based clustering has helped to uncover a number of new research themes that were not previously evident within the cocitation structure. These new themes are largely characterized by more recent research that has yet to attract a critical mass of citation links. However, the text view reveals a significant level of research activity that, in the future, may be expected to feature prominently within the broader field of CBR research.

There are many drivers that motivate a study such as this. From the standpoint of a research community such as case-based reasoning this type of study provides a useful type of literature review, one that focuses on macrolevel features of the research space (the evolution of trends and themes) instead of a more detailed analysis of particular research concepts. This can help a community to benchmark its own progress and recognize important trends that may be useful to guide future research efforts. At the same time it can also help researchers to recognise areas of research that are in decline and that are likely to prove less fruitful as a starting point for new research. In this context, we believe a review such as this can be especially helpful for new researchers entering a field as a tool to guide their early research efforts and to help point them in the direction of opportunities that may yet be hidden within the structure of recent research.

Methodologically speaking, we believe that the multiview clustering technique presented in this work serves as a useful template for this type of analysis. It is one that can be readily applied to other fields of research to good effect. For example, we are already considering this in the context of other reasonably well-defined communities such as machine learning, semantic web, and user modeling research. The combination of cocitation and text-based relational analysis provides alternative viewpoints with which to understand the evolution of mature and emerging research themes in a way that is readily reproducible given a core set of research papers and given the online citation resources that are readily available today.

Acknowledgements

This work was part supported by Science Foundation Ireland (SFI) Grant Nos. 05/IN.1/I24 and 08/SRC/I1407.

Notes

1. Downloaded from www.springer.com.
2. See scholar.google.com.
3. The PICA Browser tool and a Java implementation of PICA are available at mlg.ucd.ie/pica.
4. The dataset is available at mlg.ucd.ie/cbr.

References

Aamodt, A. 1993. Explanation-Driven Case-Based Reasoning. In *Topics in Case-Based Reasoning: Proceedings of the First European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837, 274–288. Berlin: Springer.

Aha, D., and Breslow, L. 1997. Refining Conversational Case Libraries. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 267–278. Berlin: Springer.

Aha, D. W.; Maney, T.; and Breslow, L. A. 1998. Supporting Dialogue Inferencing in Conversational Case-Based Reasoning. In *Advances in Case-Based Reasoning: Proceedings of the Fourth European Workshop on Case-Based Reasoning,* 262–273. Berlin: Springer.

Armengol, E., and Plaza, E. 1994. Integrating Induction in a Case-Based Reasoner. In *Advances in Case-Based Reasoning: Proceedings of the Second European Workshop on Case-Based Reasoning,* 2–17. Berlin: Springer.

Baccigalupo, C., and Plaza, E. 2007. A Case-Based Song Scheduler for Group Customised Radio. In *Proceedings of the Seventh International Conference on Case-Based Reasoning: Case-Based Reasoning Research and Development.* Lecture Notes in Artificial Intelligence Volume 4626, 433–448. Berlin: Springer.

Bandini, S., and Manzoni, S. 2001. CBR Adaptation for Chemical Formulation. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 634–647. Berlin: Springer.

Bento, C.; Exposto, J. L. P.; Francisco, V.; and Costa, E. 1994. Experimental Study of an Evaluation Function for Cases Imperfectly Explained. In *Proceedings of the Second European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 984, 33–44. Berlin: Springer.

Bergmann, R., and Stahl, A. 1998. Similarity Measures for Object-Oriented Case Representations. In *Proceedings of the Fourth European Workshop on Advances in Case-Based Reasoning.* Lecture Notes in Computer Science, Volume 1488, 25–36. Berlin: Springer.

Bergmann, R., and Mougouie, B. 2006. Finding Similar Deductive Consequences—A New Search-Based Framework for Unified Reasoning from Cases and General Knowledge. In *Proceedings of the 8th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 4106, 271. Berlin: Springer.

Bonzano, A.; Cunningham, P.; and Smyth, B. 1997. Using Introspective Learning to Improve Retrieval in CBR: A Case Study in Air Traffic Control. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 291–302. Berlin: Springer.

Borner, K.; Pippig, E.; Tammer, E.; and Coulon, C. 1996. Structural Similarity and Adaptation. In *Proceedings of the Third European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1168, 58–75. Berlin: Springer.

Bradley, K., and Smyth, B. 2004. An Architecture for Case-Based Personalised Search. In A*dvances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 518–532 Berlin: Springer.

Bridge, D., and Ferguson, A. 2002. Diverse Product Recommendations Using An Expressive Language for Case Retrieval. In *Proceedings of the 6th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2416, 291–298. Berlin: Springer.

Bruninghaus, S., and Ashley, K. 1997. Using Machine Learning for Assigning Indices to Textual Cases. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 303–314. Berlin: Springer.

Bruninghaus, S., and Ashley, K. D. 2001. The Role of Information Extraction for Textual CBR. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 74–89. Berlin: Springer.

Bunke, H., and Messmer, B. 1993. Similarity Measures for Structured Representations. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837. Berlin: Springer Verlag.

Carbonell, J. G.; Etzioni, O.; Gil, Y.; Joseph, R.; Knoblock, C. A.; Minton, S.; and Veloso, M. M. 1991. Prodigy: An Integrated Architecture for Planning and Learning. *SIGART Bulletin* 2(4): 51–55.

Cheetham, W., and Price, J. 2004. Measures of Solution Accuracy in Case-Based Reasoning Systems. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 106–118. Berlin: Springer.

Coyle, L.; Freyne, J.; Smyth, B.; and Cunningham, P. 2008. A Quantitative Evaluation of the Relative Status of Journal and Conference Publications in Computer Science. Technical Report UCD-CSI-2008-8, Computer Science and Information Department, University College Dublin.

Cunningham, P.; Doyle, D.; and Loughrey, J. 2003. An Evaluation of the Usefulness of Case-Based Explanation. In *Proceedings of the 5th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2689, 122–130. Berlin: Springer.

Díaz-Agudo, B., and González-Calero, P. A. 2001. A Declarative Similarity Framework for Knowledge Intensive CBR. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 158–172. Berlin: Springer.

Díaz-Agudo, B.; Gervás, P.; and González-Calero, P. A. 2002. Poetry Generation in Colibri. In *Proceedings of the 6th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2416, 73–102. Berlin: Springer.

Dimitriadou, E.; Weingessel, A.; and Hornik, K. 2002. A Combination Scheme for Fuzzy Clustering. *International Journal of Pattern Recognition and Artificial Intelligence* 16(7): 901–912.

Doyle, M., and Cunningham, P. 2000. A Dynamic Approach to Reducing Dialog in Online Decision Guides. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 323–350. Berlin: Springer.

Doyle, D.; Cunningham, P.; Bridge, D.; and Rahman, Y. 2004. Explanation Oriented Retrieval. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 157–168. Berlin: Springer.

Faltings, B. 1997. Probabilistic Indexing for Case-Based Prediction. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 611–622. Berlin: Springer.

Flinter, S., and Keane, M. 1995. On the Automatic Generation of Case Libraries by Chunking Chess Games. In *Case-Based Reasoning Research and Development: Proceedings of the 1st International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1010, 421–430. Berlin: Springer.

Gabel, T., and Stahl, A. 2004. Exploiting Background Knowledge When Learning Similarity Measures. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 169–183. Berlin: Springer.

Gmür, M. 2003. Co-Citation Analysis and the Search for Invisible Colleges: A Methodological Evaluation. *Scientometrics* 57(1): 27–57.

Goker, M., and Thompson, C. 2000. Personalized Conversational Case-Based Recommendation. In *Proceedings of the*

5th European Workshop on Case-Based Reasoning, Lecture Notes in Computer Science, Volume 1898, 29–82. Berlin: Springer.

Gomes, P., and Bento, C. 2000. Learning User Preferences in Case-Based Software Reuse. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 112–123. Berlin: Springer.

Gómez-Martín, P. P.; Gómez-Martín, M. A.; Díaz-Agudo, B.; and González-Calero, P. A. 2005. Opportunities for CBR in Learning by Doing. In *Case-Based Reasoning, Research and Development: Proceedings of the 6th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science Volume 3620, 267–281. Berlin: Springer.

Grachten, M.; Arcos, J. L.; and De Mántaras, R. L. 2004. Tempoexpress, a CBR Approach to Musical Tempo Transformations. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science Volume 3155, 601–615. Berlin: Springer.

Greene, D.; Bryan, K.; and Cunningham, P. 2008. Parallel Integration of Heterogeneous Genome-Wide Data Sources. In *Proceedings 8th International Conference on Bioinformatics and Bioengineering (Bibe'08),* Athens, 8–10 October. Piscataway, NJ: Institute of Electrical and Electronics Engineers, Inc.

Greene, D.; Freyne, J.; Smyth, B.; and Cunningham, P. 2008. An Analysis of Research Themes in the CBR Conference Literature. In *Proceedings of the 9th European Conference on Advances in Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 5239, 18–43. Berlin: Springer.

Grimnes, M., and Aamodt, A. 1996. A Two Layer Case-Based Reasoning Architecture for Medical Image Understanding. In *Proceedings of the Third European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1168, 164–178. Berlin: Springer.

Gu, M., and Aamodt, A. 2005. A Knowledge-Intensive Method for Conversational CBR. In *Case-Based Reasoning Research and Development: Proceedings of the 6th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 3620, 296–311. Berlin: Springer.

Gupta, K. M.; Aha, D. W.; and Sandhu, N. 2002. Exploiting Taxonomic and Causal Relations in Conversational Case Retrieval. In *Proceedings of the 6th European Conference on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2416, 175–182. Berlin: Springer.

Hammond, K. J. 1986. Chef: A Model of Case-Based Planning. In *Proceedings of the Fifth National Conference on Artificial Intelligence,* 267–271. Menlo Park, CA: AAAI Press.

Hayes, C.; Avesani, P.; Baldo, E.; and Cunningham, P. 2005. Re-Using Implicit Knowledge in Short-Term Information Profiles for Context-Sensitive Tasks. In *Case-Based Reasoning Research and Development: Proceedings of the 6th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 3620, 312–326. Berlin: Springer.

He, Y., and Cheung Hui, S. 2002. Mining a Web Citation Database for Author Cocitation Analysis. *Information Processing and Management* 38(4): 491–508.

Heister, F., and Wilke, W. 1998. An Architecture for Maintaining Case-Based Reasoning Systems. In *Proceedings of the 4th European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1488, 221–232. Berlin: Springer.

Jære, M.; Aamodt, A.; and Skalle, P. 2002. Representing Temporal Knowledge for Case-Based Prediction. In *Proceedings of the 6th European Conference on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2416, 225–234. Berlin: Springer.

Jarmulak, J.; Kerckhoffs, E.; and Van't Veen, P. 1997. Case-Based Reasoning in an Ultrasonic Rail-Inspection System. In *Proceedings of the Second International Conference on Case-Based Reasoning Research and Development,* Lecture Notes in Artificial Intelligence, Volume 1266, 43–52. Berlin: Springer.

Kamp, G. 1997. On the Admissibility of Concrete Domains for CBR Based on Description Logics. In *Proceedings of the 2nd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1266, 223–234. Berlin: Springer.

Kolodner, J. 1993. Understanding Creativity: A Case-Based Approach. In *Topics in Case-Based Reasoning: Proceedings of the First European Workshop,* Lecture Notes in Computer Science, Volume 837, 3–20. Berlin: Springer.

Kolodner, J. L. 1991. Improving Human Decision Making through Case-Based Decision Aiding. *AI Magazine* 12(2): 52–68.

Lamontagne, L., and Lapalme, G. 2004. Textual Reuse for Email Response. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 242–256. Berlin: Springer.

Leake, D., and Wilson, D. 1998. Categorizing Case-Base Maintenance: Dimensions and Directions. In *Proceedings of the 4th European Workshop on Advances in Case-Based Reasoning.* Lecture Notes in Computer Science, Volume 1488, 196. Berlin: Springer.

Lenz, M.; Burkhard, H.; and Bruckner, S. 1996. Applying Case Retrieval Nets to Diagnostic Tasks in Technical Domains. In *Proceedings of the Third European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1168. 219–233. Berlin: Springer.

Lluís Arcos, J., and Plaza, E. 1993. A Reflective Architecture for Integrated Memory-Based Learning and Reasoning. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837, 289–300. Berlin: Springer.

Macedo, L., and Cardoso, A. 2004. Using CBR in the Exploration of Unknown Environments with an Autonomous Agent. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 272–286. Berlin: Springer.

Macura, R., and Macura, K. 1995. Macrad: Radiology Image Resource with a Case-Based Retrieval System. In *Case-Based Reasoning Research and Development: Proceedings of the 1st International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1010, 43–54. Berlin: Springer.

Marling, C., and Whitehouse, P. 2001. Case-Based Reasoning in the Care of Alzheimer's Disease Patients. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 702–715. Berlin: Springer..

McGinty, L., and Smyth, B. 2003. On the Role of Diversity in Conversational Recommender Systems. In *Case-Based Reasoning Research and Development: Proceedings of the 5th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 2689, 1065–1065. Berlin: Springer.

McSherry, D. 1998. An Adaptation Heuristic for Case-Based Estimation. In *Proceedings of the 4th European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1488, 184–195. Berlin: Springer

McSherry, D. 2002. Diversity-Conscious Retrieval. In *Proceedings of the 6th European Workshop on Advances in Case-Based Reasoning.* Lecture Notes in Computer Science, Volume 2416, 27–53. Berlin: Springer.

McSherry, D. 2003. Similarity and Compromise. In *Proceedings of the 5th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2689, 1067–1067. Berlin: Springer.

McSherry, D. 2004. Explaining the Pros and Cons of Conclusions in CBR. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 317–330. Berlin: Springer.

Montani, S.; Bellazzi, R.; Portinale, L.; and Stefanelli, M. 2000. Evaluating a Multi-Modal Reasoning System in Diabetes Care. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 213–234. Berlin: Springer.

Mougouie, B.; Richter, M.; and Bergmann, R. 2003. Diversity-Conscious Retrieval from Generalized Cases: A Branch and Bound Algorithm. In *Case-Based Reasoning Research and Development: Proceedings of the 5th International Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2689, 1064–1064. Berlin: Springer.

Mukkamalla, S., and Muñoz-Avila, H. 2002. Case Acquisition in a Project Planning Environment. In *Proceedings of the 6th European Conference on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 2416, 264–277. Berlin: Springer.

Muñoz-Avila, H. 1999. A Case Retention Policy Based on Detrimental Retrieval. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 721. Berlin: Springer.

Nakatani, Y., and Israel, D. 1993. Tuning Rules by Cases. In *Topics in Case-Based Reasoning: Proceedings of the First European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837, 313—324. Berlin: Springer.

Nakhaeizadeh, G. 1993. Learning Prediction of Time Series—A Theoretical and Empirical Comparison of CBR with Some Other Approaches. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837. Berlin: Springer.

Neagu, N., and Faltings, B. 2001. Exploiting Interchangeabilities for Case Adaptation. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 74–89. Berlin: Springer.

Neagu, N., and Faltings, B. 2003. Soft Interchangeability for Case Adaptation. In *Case-Based Reasoning Research and Development: Proceedings of the 5th International Conference on Case-Based Reasoning,* Lecture Notes in Computer Science Volume 2689, 1066–1066. Berlin: Springer.

Netten, B., and Vingerhoeds, R. 1995. Large-Scale Fault Diagnosis for On-Board Train Systems. In *Case-Based Reasoning Research and Development: Proceedings of the 1st International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1010, 67–76. Berlin: Springer.

Opiyo, E. 1995. Case-Based Reasoning for Expertise Relocation in Support of Rural Health Workers in Developing Countries. In *Case-Based Reasoning Research and Development: Proceedings of the 1st International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1010, 77–87.Berlin: Springer.

Osborne, H., and Bridge, D. 1996. A Case Base Similarity Framework. In *Proceedings of the Third European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1168, 309–323. Berlin: Springer.

Osborne, H., and Bridge, D. 1997. Similarity Metrics: A Formal Unification of Cardinal and Non-Cardinal Similarity Measures. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 235–244. Berlin: Springer.

Perner, P. 1999. An Architecture for a CBR Image Segmentation System. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 724. Berlin: Springer.

Plaza, E. 1995. Cases as Terms: A Feature Term Approach to the Structured Representation of Cases. In *Proceedings of the First International Conference on Case-Based Reasoning: Case-Based Reasoning Research and Development*. Lecture Notes in Computer Science, Volume 1010. 265–276. Berlin: Springer.

Portinale, L.; Torasso, P.; and Tavano, P. 1999. Speed-Up, Quality and Competence in Multi-Modal Case-Based Reasoning. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 718. Berlin: Springer.

Racine, K., and Yang, Q. 1997. Maintaining Unstructured Case Bases. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 553–564. Berlin: Springer.

Reinartz, T.; Iglezakis, I.; and Roth-Berghofer, T. 2000. On Quality Measures for Case Base Maintenance. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 249–259. Berlin: Springer.

Richards, B. 1994. Qualitative Models as a Basis for Case Indices. In *Selected Papers from the Second European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 984, 126–135. Berlin: Springer.

Rissland, E. L. 2006. The Fun Begins with Retrieval: Explanation and CBR. In *Proceedings of the 8th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 4106, 1–8. Berlin: Springer.

Rissland, E. L.; Valcarce, E. M.; and Ashley, K. D. 1984. Explaining and Arguing with Examples. In *Proceedings of the Fourth National Conference on Artificial Intelligence,* 288–294. Menlo Park, CA: AAAI Press.

Schaaf, J. 1996. Fish and Shrink. A Next Step Towards Efficient Case Retrieval in Large Scaled Case Bases. *Proceedings of the Third European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1168, 362–376. Berlin: Springer.

Schank, R. C., and Leake, D. B. 1989. Creativity and Learning in a Case-Based Explainer. *Artificial Intelligence* 40(1–3): 353–385.

Schmidt, R.; Pollwein, B.; and Gierl, L. 1999. Case-Based Reasoning for Antibiotics Therapy Advice. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 723. Berlin: Springer.

Schölkopf, B.; Smola, A.; and Müller, K.-R. 1998. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation* 10(5): 1299–1319.

Sebag, M., and Schoenauer, M. 1993. A Rule-Based Similarity Measure. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837, 119–131. Berlin: Springer.

Seitz, A. 1999. A Case-Based Methodology for Planning Individualized Case Oriented Tutoring. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 318–328. Berlin: Springer.

Smyth, B., and Keane, M. 1993. Retrieving Adaptable Cases: The Role of Adaptation Knowledge in Case Retrieval. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 837, 209–220. Berlin: Springer.

Smyth, B., and McClave, P. 2001. Similarity Versus Diversity. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 347–361. Berlin: Springer.

Smyth, B., and McKenna, E. 1999. Footprint-Based Retrieval. In *Proceedings of the 3rd International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 1650, 343–357. Berlin: Springer.

Smyth, B., and McKenna, E. 1998. Modelling the Competence of Case-Bases. In *Proceedings of the 4th European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1488, 208–220. Berlin: Springer.

Smyth, B. 2000. Competence Models and Their Applications. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 1–2. Berlin: Springer.

Sørmo, F. 2005. Case-Based Student Modeling Using Concept Maps. In *Proceedings of the 6th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 3620, 492–506. Berlin: Springer.

Stahl, A. 2001. Learning Feature Weights from Case Order Feedback. In *Proceedings of the 4th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2080, 74–89. Berlin: Springer.

Stahl, A. 2005. Learning Similarity Measures: A Formal View Based on a Generalized CBR Model. In *Case-Based Reasoning Research and Development: Proceedings of the 6th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 3620, 507–521. Berlin: Springer.

Stahl, A. 2006. Combining Case-Based and Similarity-Based Product Recommendation. In *Proceedings of the 8th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 4106, 355–369. Berlin: Springer.

Stahl, A., and Gabel, T. 2003. Using Evolution Programs to Learn Local Similarity Measures. In *Case-Based Reasoning Research and Development: Proceedings of the 5th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 2689, 1064–1064.

Stanfill, C., and Waltz, D. L. 1986. Toward Memory-Based Reasoning. *Communications of the ACM* 29(12): 1213–1228.

Strehl, A., and Ghosh, J. 2002. Cluster Ensembles—A Knowledge Reuse Framework for Combining Multiple Partitions. *Journal of Machine Learning Research* 3: 583–617.

Surma, J., and Tyburcy, J. 1998. A Study on Competence-Preserving Case Replacing Strategies in Case-Based Reasoning. In *Proceedings of the 4th European Workshop on Advances in Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1488, 233–238. Berlin: Springer.

Tobudic, A., and Widmer, G. 2003. Playing Mozart Phrase by Phrase. In *Proceedings of the 5th International Conference on Case Based Reasoning,* Lecture Notes in Computer Science, Volume 2689, ed. K. D. Ashley and D. G. Bridge, 343–357. Berlin: Springer.

Tonidandel, F., and Rillo, M. 2005. Case Adaptation by Segment Replanning for Case-Based Planning Systems. In *Case-Based Reasoning Research and Development: Proceedings of the 6th International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 3620, 579–594.

Trott, J., and Leng, B. 1997. An Engineering Approach for Troubleshooting Case Bases. In *Case-Based Reasoning Research and Development: Proceedings of the 2nd International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1266, 178–189. Berlin: Springer.

Vollrath, I. 2000. Handling Vague and Qualitative Criteria in Case-Based Reasoning Applications. In *Proceedings of the 5th European Workshop on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 1898, 403–444. Berlin: Springer.

Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications.* New York: Cambridge University Press.

Wess, S.; Althoff, K.; and Derwand, G. 1993. Using K-D Trees to Improve the Retrieval Step in Case-Based Reasoning. In *Selected Papers from the First European Workshop on Topics in Case-Based Reasoning.* Lecture Notes in Computer Science, Volume 837, 181. Berlin: Springer.

Wettschereck, D., and Aha, D. 1995. Weighting Features. In *Case-Based Reasoning Research and Development: Proceedings of the 1st International Conference on Case-Based Reasoning,* Lecture Notes in Artificial Intelligence, Volume 1010, 347–358. Berlin: Springer.

White, H., and Griffith, C. 1981. Author Co-Citation: A Literature Measure of Intellectual Structure. *Journal of the American Society for Information Science* 32(3): 163–171.

Wiratunga, N.; Koychev, I.; and Massie, S. 2004. Feature Selection and Generalisation for Retrieval of Textual Cases. In *Advances in Case-Based Reasoning: Proceedings of the 7th European Conference on Case-Based Reasoning,* Lecture Notes in Computer Science, Volume 3155, 806–820. Berlin: Springer.

Wiswedel, B., and Berthold, M. 2007. Fuzzy Clustering in Parallel Universes. *International Journal of Approximate Reasoning* 45(3): 439–454.

Derek Greene is a senior postdoctoral researcher in the Clique Research Cluster, based at the School of Computer Science and Informatics, University College Dublin. He received his Ph.D. from Trinity College Dublin in 2006, and his current research interests include unsupervised learning, text mining, social network analysis, and the application of machine learning techniques in bioinformatics.

Jill Freyne is a research scientist at CSIRO's ICT Centre based in Hobart, Australia. Freyne was awarded her Ph.D. in computer science specializing in collaborative web search from University College Dublin, Ireland. Freyne's research interests are in human-computer interaction with a special interest in personalization, recommender technologies, information retrieval, and social media.

Barry Smyth holds the Digital Chair of Computer Science in the School of Computer Science and Informatics at University College Dublin. He has a Ph.D. from Trinity College Dublin. His research interests include personalization, recommender systems, case-based reasoning, machine learning, and information retrieval, and he has published widely on these topics. Smyth is a cofounder of ChangingWorlds Ltd., a leading provider of mobile content discovery solutions and now a division of Amdocs. Smyth is currently the director of Clarity: the Centre for Sensor Web Technologies.

Pádraig Cunningham is the Professor of Knowledge and Data Engineering in the School of Computer Science and Informatics at University College Dublin. His current research focus is on the analysis of graph and network data and on the use of machine learning techniques in processing high-dimension data. He has a B.E. and M.Eng.Sci. from NUI Galway and a Ph.D. from Dublin University, which he received in 1989.

www.aaai.org/iaai10

Nestor Rychtyckyj, Chair & Daniel Shapiro, Cochair

IAAI-10 Atlanta