Explaining Compound Critiquing *

James Reilly, Kevin McCarthy, Lorraine McGinty, and Barry Smyth

Adaptive Information Cluster, Smart Media Institute,
Department of Computer Science, University College Dublin (UCD), Ireland.
{james.d.reilly, kevin.mccarthy, lorraine.mcginty, barry.smyth}@ucd.ie

Abstract. Interactive recommender applications are an important technology for online retailers who want to increase their sales by providing potential buyers with suitable recommendations. Critique-based navigation, where a user applies a directional critique to a specific feature of a presented recommendation, is often used to guide users through such complex product spaces. Dynamic critiquing is a novel extension of this approach, where users are presented with compound critiques that reflect the product options that are available in a given cycle of a recommendation session. We believe that these compound critiques have considerable explanatory capability. In this paper we are interested in how these kind of explanations can be used to help the user to better understand, and effectively navigate, complex product spaces in interactive recommender systems by understanding the trade-offs/inter-dependencies that might exist between product features. We present sample screen-shots from a purpose-built prototype application to illustrate the potential benefits of our approach.

1 Introduction

The advent of Electronic Commerce has opened up a whole new world to both online shoppers and retailers. One of the many benefits to both parties is that the range of product opportunities that exist is no longer limited by the amount of available real-estate. However, strange as it may seem, the vastness of the product-space also poses an enormous problem to both; online shoppers often find it difficult to find what they want, and this results in missed sales opportunities for the retailer. Several approaches have been proposed to help users navigate through the complex product space of an electronic shop [9]. While *keyword search* and *category-based browsing* are widely used, both online shoppers and retailers have reported that they are substandard [7]. Instead many well-known online retailers, such as Amazon.com, use recommender technology to help the user to narrow-down the range of possible options [16].

Although recommender systems are a popular solution to the narrowing problem, one criticism is commonplace. Typically, recommendations are retrieved on the basis of their *match score*; that is, how closely they match the user's evolving

* This material is based on works supported by Science Foundation Ireland under Grant No. 03/IN.3/I361

query. Rarely is either the score or how it was computed shown to the user, and even when it is this provides little comprehension value to the user. It is this lack of transparency that prevents users from perceiving the recommendations as credible [6]. Instead users prefer the feature-based style of navigation, where the relationship between their stated requirements and the provided product descriptions is obvious. Critique-based navigation, proposed by Burke *et al.* [3], is an example of a navigation approach often used by recommender systems. The approach is interactive and incremental and does not require the user to have a specified need at the start. However it is susceptible to two distinct problems; first, this approach can lead to retrieval failures depending on the order in which the critiques are executed. Secondly, since only one feature critique can be executed in a given cycle, lengthly recommendation sessions tend to result.

In previously published work, we have described our *Dynamic Critiquing* approach, which concentrates on presenting the user with a selection of appropriate *compound* (i.e., multi-feature) critiques. Importantly, these compound critiques are representative of the product opportunities that exist at a given point in a recommendation session, as they are generated in real-time. We have already demonstrated the potential of our dynamic critiquing approach when it comes to improving recommendation efficiency. In this paper, we turn our attention to the explanatory benefits of the dynamic critiquing approach. Explanations have an important role to play in helping users to understand the suggestions made by recommender systems. Much of the research conducted to date has focused on ways to justify a particular recommendation to the end-user. Here, we take a different stance by proposing that compound critiques serve as rich explanations, as well as functioning as an intuitive and efficient navigation mechanism. Specifically, we suggest that they help the user to understand the recommendation opportunities that exist beyond the current suggestion, on the assumption that this current suggestion does not satisfy all of the user's implicit requirements.

We present a prototype recommendation application that demonstrates how these kind of explanations can be used to help the user to effectively navigate complex product spaces in interactive CBR systems. Indeed, the subject of Case-Based Explanation (CBE) is one that has recently become very topical at both conferences and workshops in the CBR arena. The following sections describe how our recent work, in the area of *Dynamic Critiquing* [10, 14], has potential explanatory benefits that better support the user's navigation task, as well as significant recommendation efficiency benefits. First, we discuss some of the related work in the broader area of CBE, and distinguish how our contribution differs from other CBE approaches relating to product-space navigation.

2 Related work in the Area of Case-Based Explanation

In this section we discuss related work in the area of case-based explanation, but first we distinguish between two apparent research focuses: (1) system-oriented explanations, and (2)user-oriented explanation. Early work in the area of CBE has focused on the former; the use of explanation-based techniques in order to drive the CBR process model. In this sense explanation structures are generated to fulfill various functions within a CBR system: the explanations are constructed **by** the system, **for** the system (see for example, [1, 2, 8]). More relevant here is the use of explanations for the benefit of the user by developing systems that are capable of explaining the reasoning steps and conclusions. The following sections discuss some of the most recent work that has been carried out in the areas of (1) diagnosis and classification tasks, and (2) product recommendation tasks.

2.1 CBE for Diagnosis and Classification

Diagnosis and classification systems may generate explanations in order to justify a predicted outcome and satisfy the user of its validity. Recently, there has been an increasing interest in the use of cases as a source of explanation in these type of situations. Indeed there is considerable optimism among the CBR community regarding the potential value of case-based or *precedent-based* approaches to explanation, when compared to their more traditional rule-based counterparts. The argument has been made that past cases provide a more natural and convincing form of explanation. For example, Cunningham *et al.* [4] report how real users find similar cases to be more convincing than rule-based explanations in a classification task.

However, using the most similar case in order to justify or explain a particular classification outcome to the user is perhaps the simplest form of case-based explanation and many issues remain a source of active research within the community. Recently, the work of Doyle *et al.* [5] demonstrates that in some classification tasks presenting the nearest-neighbour case may not be the best way to explain or justify a particular classification outcome. This occurs when the nearest-neighbour happens to be farther from the decision surface than the target case. The work demonstrates how superior explanation cases can be selected by using an explanation utility metric that formalises this idea.

Presenting the user with an *explanation* case is just the beginning of the explanation process and on its own may not be sufficient to convince the user that a particular diagnosis or classification outcome is justified. For instance, as McSherry points out, attempting to justify a predicted outcome by presenting the user with the particular explanation case (whether it is chosen because it was a nearest-neighbour to the target or because of some alternative strategy), ignores the possibility that some of the features of this explanation case may conflict with some of the target features [11, 12]. This may mislead the user if they mistakenly view these opposing features to be evidence in favour of the predicted outcome [11]. To combat this problem, McSherry proposes an evidential approach to precedent-based explanation [11, 12]. The user is presented with evidence that both supports and opposes a particular outcome in order to explain the pros and cons of case-based conclusions.

2.2 CBE for Product Recommendation

Generally a product recommender may use explanations to explain (1) the reasons WHY a particular product was recommended, or indeed why there is no product that can be recommended [13], and (2) WHAT opportunities remain; that is, *"where can I get to from here"*, when presented with an unsuitable recommendation. Since the vast majority of related work in the area of CBE and product recommendation has also concentrated on the idea of using cases as a source of explanation, many of the research efforts discussed above apply here also. Of course, product recommenders may not always be able to satisfy all of a user's requests so it is important that the system tries to explain the cause of the retrieval failure. A good example of this is presented in [13], which describes a mixed-initiative approach to recovery from retrieval failure by helping the user to eliminate certain constraints from their initial queries in a conversational product recommender. An important aspect of this work is that explanations are generated in order to explain retrieval failures by highlighting subsets of query features that cannot be satisfied (eg. *"there are no cameras with price less than 300 Euro and resolution greater than 4 mega-pixels"*). This is particularly relevant and complementary to the work described in this paper if one regards a recommendation cycle where the user has not found their target product to be a type of retrieval failure.

In our approach [10, 14], instead of trying to explain the retrieval failure we attempt to explain the retrieval opportunities that remain (eg. *"there are 10 cameras less than 300 Euro but their resolution is between 1 and 4 mega-pixels"* or *"there are 20 cameras with between 4 and 6 mega-pixels but their price is more than 300 Euro"*). To put this another way: instead of explaining what types of products do not exist, we explain what types of products do exist. We argue that this type of *positive* explanation is likely to be more useful as it is more intuitive for users to respond to explanations that tell them where they can go rather than where they can't go.

3 Explanation-Guided Navigation

In the following sections we describe how the standard approach to critique-based navigation operates, discuss how it is subject to retrieval failures, and suggest how the user may better understand remaining product opportunities through explanation.

3.1 The Standard Approach to Critique-Based Navigation

The standard critique-based navigation policy [3] is very simple. The key idea is that by critiquing presented examples (i.e., cases), the search can be re-directed to home-in on appropriate products for the interacting user. Very briefly, each recommendation session is initiated by a vague user query and this results in the retrieval of the most similar case (*the recommended case*), and a set of fixed directional feature critiques. The user has opportunity to accept this case, thereby ending the recommendation session, or to critique it in line with their requirements. For example, a digital camera recommender might present the user with a particular camera recommendation and the user may request to see something *"like this but cheaper"*; here *cheaper* is a critique over the price feature. Individual critiques act as a filter over the remaining cases, such that the case chosen for the next cycle is the one that is compatible with the critique and maximally similar to the previously recommended case. Figure 1 illustrates how

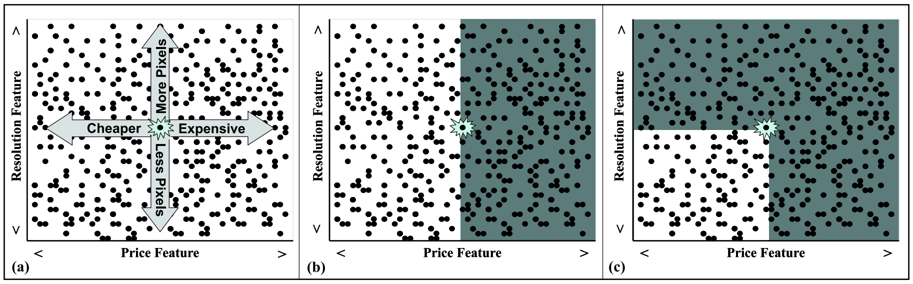

Fig. 1. Product space; (a) two features and their available critiques, (b) cases available after "*cheaper*" unit critique chosen, (c) cases available after "*cheaper & less pixels*" compound critique chosen (non grey area).

critiquing serves as a useful navigation mechanism. When the user is presented with a particular recommendation, he can immediately re-direct the search focus by applying a navigation action (i.e., feature-critique) in line with his requirements. Figure 1(a) shows two features, Resolution and Price, and their available critiques. The user can now choose to critique either of these features in two directions, respectively. Figure 1(b) shows the product space after the user has decided to look at cheaper products by choosing the *Cheaper* critique. This critique has re-directed the retrieval focus to a different part of the product space and closer to what the user is looking for.

3.2 The Problem: Limited Visibility

From a user's perspective it is often difficult to know how best to critique a proposed case in order to make useful progress through the product space. For example, consider a recommender system for Digital Cameras and suppose the user is presented with a *300 Euro* model with a *2 mega-pixel* resolution, *64MB* of card-memory and *3x* optical zoom. Let's suppose that the user is actually looking for a greater resolution and more memory, but at a lower price. Ordinarily he would start by critiquing one of the features, *price, resolution or memory.* There may be no cameras which satisfy all three of these constraints and by requesting a camera that is *cheaper* he may find that he is recommended a new camera that is cheaper but has lower resolution and less memory, which may not be satisfactory. The point is, that normally this type of recommender system provides the user with little visibility; that is, it does not help the user to understand what product options exist beyond the current case. This can mean that the user spends time following false leads and backtracking as he tries different sequences of critiques.

We believe that explanations can play a useful role in this regard by providing the user with an indication of the type of products that remain available. For instance, in the above example it might be useful for the system to highlight that there are 30 products with a lower price and greater resolution, but that these products also have less memory. Or that there are 15 products with greater resolution and more memory but they are also more expensive.

3.3 Compound Critiques as Explanations

Previously [10, 14], we have described how to generate a set of compound critiques during each recommendation cycle, and a method for presenting the promising compound critiques to the user for consideration. Each compound critique covers a number of features (typically 2 or 3) and they are generated in real time based on the cases that remain for a particular cycle. A data mining algorithm is used to extract common patterns of feature relationships from the remaining cases. These common patterns are then converted into compound critiques, the best of which are selected for presentation. Further details of how we dynamically generate and select compound critiques to present in each recommendation cycle can be found in [10, 14]. We have also demonstrated the potential of our dynamic critiquing approach when it comes to improving recommendation efficiency.

The core hypothesis is that compound critiques help the user to better understand the *recommendation opportunities* that exist beyond the current cycle by helping them to appreciate common interactions between features. We believe that in many recommender domains, where the user is likely to have incomplete knowledge about the finer details of the feature-space, that compound critiques will help to effectively map out this space. For this reason we believe that users will actually find it easier to work with compound critiques, and their associated explanations, than unit critiques and this may, for example, help the user to make fewer critiquing errors. For instance, with standard critiquing in the digital camera domain a user might naively select the $[Price <]$ unit critique in the mistaken belief that this may deliver a cheaper camera that satisfies all of their other requirements. However, reducing price in this way may lead to a reduction in resolution that the user might not find acceptable and, as a result, they will have to backtrack. This problem is less likely to occur if the compound critique $\{[Price <], [Resolution <]\}$ is presented because the user will come to understand the implications of a price-drop prior to selecting any critique. In 1(c) we can see that when a user picks a "*Cheaper & Less Pixels*" compound critique, the remaining product space is focused considerably.

4 A Demonstration Prototype

The importance of system transparency for fostering depth of user understanding has been shown by empirical studies [17, 18]. We propose that by presenting the user with a selection of compound critiques that best describe the product opportunities that remain, the user can gain a deeper understanding of the recommendation process. Thus, the compound critiques themselves serve as the explanatory mechanism that facilitates this understanding. To illustrate our idea, Figures 2-3 present a series of screenshots from a prototype application of our dynamic critiquing approach in the context of an online digital camera store. The screenshots present a sequence of recommendation cycles and in each we see the currently recommended case, its features and their unit critiques, plus a set of 3 compound critiques translated into natural language.

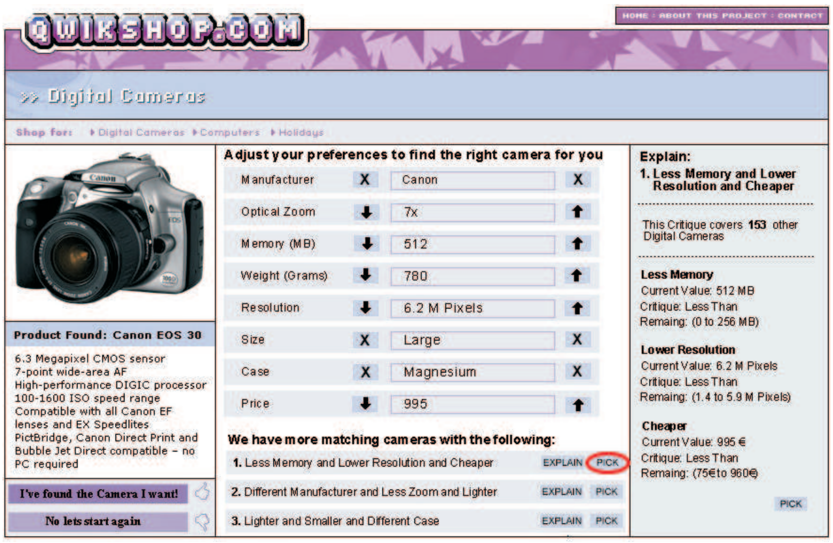

Fig. 2. Initially the user is presented with a high-end Canon camera but selects the first compound critique to indicate that they are looking for something cheaper and are willing to come down on memory and resolution.

After the user has provided some initial information they are presented with a high-end Canon camera for 995 Euro with 512MB of memory and 6.2 megapixels, as shown in Figure 2. The user can critique any of the individual features, such as *manufacturer, memory, resolution* etc., by selecting the appropriate critique icon on either side of the feature value fields that are displayed for the current camera. In addition, just below these features, three compound critiques are displayed. The compound critiques indicate to the user what other types of cameras are available and help the user appreciate the relationships that exist between digital camera features. For example, in Figure 2, the first compound critique, *"Less Memory, Lower Resolution and Cheaper"*, tells the user that if he wants a cheaper model than the currently recommended camera then he should also expect to compromise on the *resolution* and *memory* features.

Figure 3 shows the next recommendation cycle in which the user is presented with a different camera and a different set of compound critiques. In this cycle, the compound critiques are perhaps not as intuitive and understandable, but

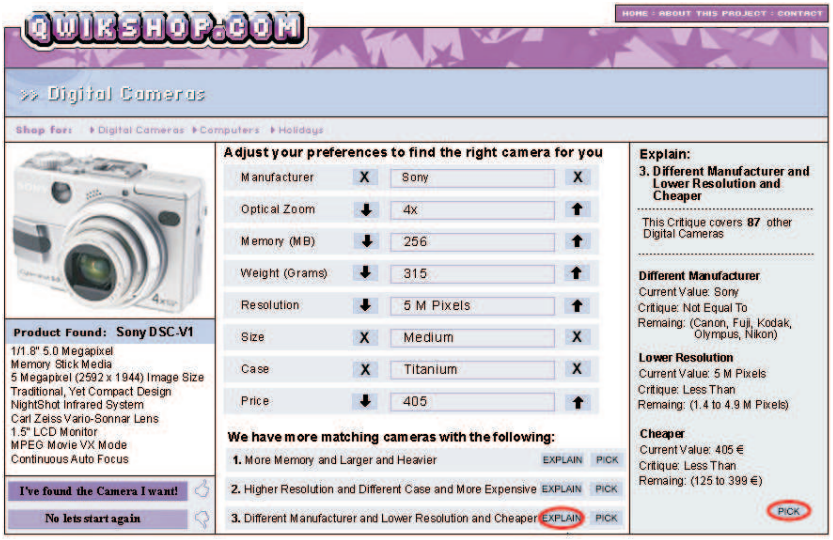

Fig. 3. This camera is cheaper with less memory and a lower resolution. It is still too expensive though. The compound critique explains to the user that there are reputable manufacturers remaining, and that the lower resolutions offered are still acceptable given the significant price drop that is available.

that is not to say they are any less relevant. By clicking on the 'explain' option for a compound critique the user can request a more detailed explanation. This explanation is presented in the pane to the right of the feature values, and defaults to be an explanation of the first compound critique. For example in Figure 3, the user asks for a further explanation of the third compound critique (*"Different Manufacturer, Lower Resolution and Cheaper"*). The resulting explanation tells the user that there are 87 remaining cameras that satisfy this critique—that is, there are 87 cameras that are cheaper, with a lower resolution, and made by a different manufacturer, than the currently recommended Sony camera. In addition, the explanation provides information about the ranges of values for these critiqued features. For instance, the user is told that these 87 cameras are made by manufacturers such as Canon, Fuji, Olympus, Kodak and Nikon, that they have resolutions from 1.4 to 4.8 mega-pixels, and that their price ranges from 125 to 399 Euro.

5 Discussion

At this point we are at the stage where the core technology has been developed and is being deployed in a realistic application setting. In addition the evaluation described in [14] provides evidence from an off-line study that if compound critiques are selected by users then their potential to improve recommender performance is significant.

However, feedback we received from various seminars and workshop events where we presented our ideas suggested that we should carry out a set of 'real-user' trails to further validate our claims before we should publish the results. Recently, we conducted tested our approach on a set of 20 live-users to determine if the benefits we expected benefits were realistic in practice. We found this study to be very useful; both in terms of validating our claims, and highlighting other areas for further research that would have been overlooked but for this study was carried out. We hope that the feedback in this instance will be just as fruitful!

The results of our live-user trial suggest that dynamic critiquing approach has the potential to help users find their desired products more efficiently that traditional unit critiques. In most cases the trialists found the compound critiques that they were presented with to be relevant, and indicated that "they could better understand the options that where available to them and why". However, while these results are promising, during the course of the trial a number of important issues were raised by the participants. These problems, as well as the steps we have taken to address them, are briefly discussed below.

5.1 Critique Diversity

Some users mentioned that the system sometimes presented them with compound critiques that lacked diversity among their feature constraints. Figure 4 shows a typical example in which a camera is presented to the user along with three compound critiques. The problem is that these three different critiques overlap considerably in terms of their individual critiques this limiting their applicability, and explanation depth. We are currently looking at ways of addressing this problem; that is, how can we increase the diversity of the compound critiques presented without compromising the other performance benefits of the approach (i.e., recommendation efficiency, user applicability, explanatory power etc).

5.2 Critique Continuity

Other users commented that sometimes the recommender system appeared to 'forget' about their earlier critiques. For example, one user indicated that they had asked for at least a 2x optical zoom during one cycle but that later, when the asked for an increased digital zoom, the recommender suggested a camera with a lower optical zoom. This problem stems from the fact that in the prototype application critiques are not tracked from cycle to cycle; each critique (compound or unit) is applied in isolation. This can lead to a number of *continuity* problems, especially when users are unsure of their requirements. To cater for this we have recently developed and evaluated an extended approach to dynamic critiquing called *incremental critiquing* which successfully solves this problem, and which has been shown to offer further potential efficiency improvements [15]. Very briefly, incremental critiquing maintains a history of a users critiques within a given session and new recommendations are selected not only because of their compatibility with the current critique but also on the basis that they match as many previous critiques as possible. This combination of factors means

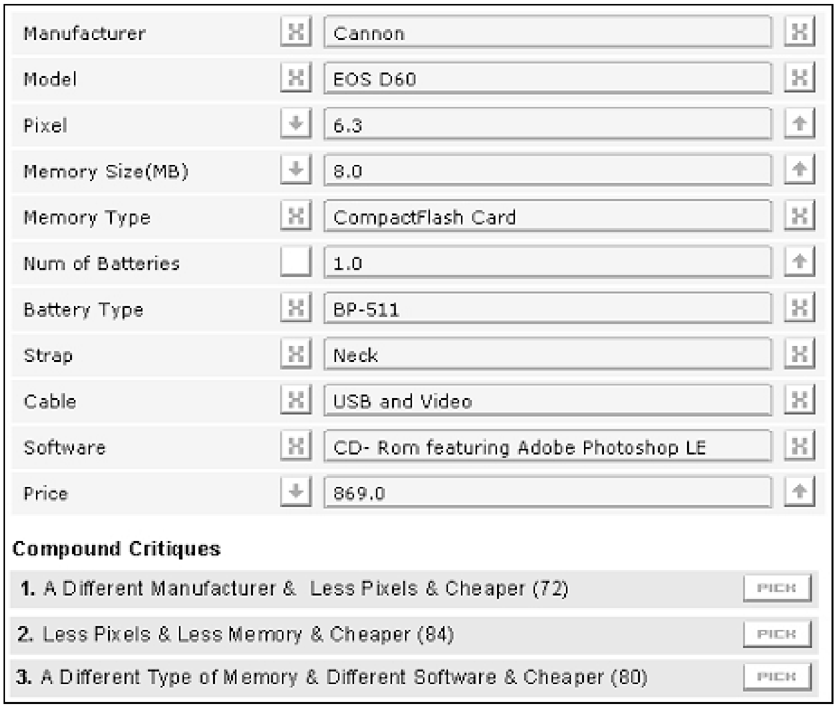

Fig. 4. Sometimes the compound critiques presented can lack diversity which, in turn, can compromise the depth of explanation they can offer to a user.

that suggestions that conflict with past critiques are less likely and helps the recommender to better focus its search through the product-space, as well prevent unnecessarily confusing the user.

6 Conclusions

Explanations have an important role to play in helping users to understand the suggestions made by recommender systems. Much of the research conducted to date has focused on ways to justify a particular recommendation to the end-user. In this paper we have taken a different stance by highlighting the importance of helping the user to understand the recommendation opportunities that exist beyond the current suggestion, on the assumption that this current suggestion does not satisfy all of the user's implicit requirements. Specifically, we discuss how compound critiques may serve as rich explanations that assist the user to better navigate the product space.

Our work has been implemented in a live demonstration system and off-line evaluations have demonstrated that compound critiquing has a potentially valuable role to play in explanation.

Finally it is worth pointing out, of course, that our approach to explanation is but one of many different approaches to explanation. We have focused on the need to help users to understand what options remain available, if the current recommendations should not meet their requirements. As we have seen in Section 2, other approaches to explanation have a different focus, such as justifying a particular recommendation or explaining its pros and cons. In the future, it is likely that we will come to see many of these explanation strategies playing their own particular roles in the next generation of interactive recommender systems.

References

1. A. Aamodt. Explanation-Driven Case-Based Reasoning. In S. Wess, K.D. Althoff, and M. Richter, editors, *Proceedings of the European Workshop on Case-Based Reasoning (EWCBR-94)*. Springer, 1994.
2. R. Barletta and W. Mark. Explanation-Based Indexing of Cases. In *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-88)*. AAAI, 1988. Minneapolis, MN, US.
3. R. Burke, K. Hammond, and B.C. Young. The FindMe Approach to Assisted Browsing. *Journal of IEEE Expert*, 12(4):32–40, 1997.
4. P. Cunningham, D. Doyle, and J. Loughrey. An Evaluation of the Usefulness of Case-Based Explanation. In K. Ashley and D. Bridge, editors, *Case-Based Reasoning Research and Development. LNAI, Vol. 2689.*, pages 191–199. Springer-Verlag, 2003. Berlin.
5. D. Doyle, P. Cunningham, D. Bridge, and Y. Rahman. Explanation Oriented Retrieval. In *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04)*. Springer, 2004. Spain.
6. B.J. Fogg. *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, 2002.
7. P. Hagen, H. Manning, and Y. Paul. Must Search Stink? *The Forester Report*, June 2000.
8. L. Ihrig and S. Kambhampati. An Explanation-Based Approach to Improve Retrieval in Case-Based Planning. In M. Ghallab and A. Milani, editors, *Proceedings of the European Workshop on Planning (EWSP-95)*, pages 395–406. IOS Press, 1995. Amsterdam.
9. A. Kobsa, J. Koenemann, and W. Pohl. Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships. *The Knowledge Enginnering Review*, 16(2):111–155, 2001.
10. K. McCarthy, J. Reilly, L. McGinty, and B. Smyth. On the Dynamic Generation of Compound Critiques in Conversational Recommender Systems. In Paul De Bra, editor, *Proceedings of the Third International Conference on Adaptive Hypermedia and Web-Based Systems (AH-04)*. Springer, 2004. Eindhoven, The Netherlands.
11. D. McSherry. Explanation in Case-Based Reasoning: An Evidential Approach. In B. Lees, editor, *Proceedings of the 8th UK Workshop on Case-Based Reasoning*, 2003. Cambridge, UK.
12. D. McSherry. Explaining the Pros and Cons of Conclusions in CBR. In P. A. Gonzalez Calero and P. Funk, editors, *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04)*. Springer, 2004. Madrid, Spain.
13. D. McSherry. Incremental Relaxation of Unsuccessful Queries. In P. A. Gonzalez Calero and P. Funk, editors, *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04)*. Springer, 2004. Madrid, Spain.
14. J. Reilly, K. McCarthy, L. McGinty, and B. Smyth. Dynamic Critiquing. In P. A. Gonzalez Calero and P. Funk, editors, *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04)*. Springer, 2004. Madrid, Spain.

15. J. Reilly, K. McCarthy, L. McGinty, and B. Smyth. Incremental Critiquing. In *Proceedings of The 24th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence. (AI-04)*, 2004. Cambridge, UK.
16. J.B. Schafer, J.A. Konstan, and E. Riedl. E-Commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, 5(1/2):115–153.
17. R. Sinha and K. Swearingen. On the Role of Transparency in Recommender Systems. In *Proceedings of the ACM CHI 02 Conference on Human Factors in Computing Systems.*, pages 830–831. Conference Companion, 2002.
18. J. Zimmerman and K. Kurapati. Exposing Profiles to Build Trust in a Recommender. In *Proceedings of the ACM CHI 02 Conference on Human Factors in Computing Systems.*, pages 608–609. Conference Companion, 2002.