Involving the User in Semantic Search

Axel-Cyrille Ngonga Ngomo and Frank Schumacher

University of Leipzig, Institute for Computer Sciences, Johannisgasse 26,
04103 Leipzig, Germany
{ngonga,schumacher}@informatik.uni-leipzig.de

Abstract. Retrieval systems have become one of the most used categories of computer tools. Yet the interfaces of modern information retrieval systems fail to address the correlation between the user's context and her information need. Furthermore, they usually do not integrate methods that allow whole communities or groups of users to profit of single retrieval instances. In this paper, we present our vision of an innovative, collaborative information retrieval and presentation approach based on the human factors balance model. Our approach combines automatic natural language processing results with handcrafted knowledge models and integrates implicit retrieval based on intelligent document segmentation and presentation, providing users with contextually relevant information.

Keywords: Information presentation, Intelligent Systems, Knowledge Management.

1 Introduction

The increasing tendency to utilize computer systems in various organization related tasks leads to a steadily growing amount of digitalized information that needs to be retrieved when completing certain tasks. Retrieval systems have therefore become one of the most used categories of computer tools. In the context of companies, IR tools are used to retrieve supplementary information linked to the current task of the user, the primary information needs of the user being modeled in the business process along which she works. Nevertheless, retrieving such information can be of crucial nature for the completion of a task, especially when operating in the information society in which we live. Yet the interfaces of current information retrieval systems present many drawbacks that lead to poor search results and thus to high investments in search with respect to time. Most of them are still very primitive and usually consist of a simple text field for the input and a link list as output. Even more elaborate interfaces fail to integrate the user's context in her retrieval process and do not allow to map the users' knowledge with the content of the searchable corpus, leading primarily to an explorative but uncontrolled query refinement: the user successively reformulate her query to shape the retrieval result without knowledge of the relations between their semantic concepts, the keywords and the documents. The output is usually presented by displaying a short excerpt from the document and a link. The user is thus obliged to evaluate the relevance of each document based on

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2007, LNCS 4557, pp. 507–516, 2007.

very little knowledge about its real content. Furthermore she is not given the possibility to transfer her insights concerning the relevance of a piece of information for a query to the other members of the organization.

We present our vision of an innovative, collaborative IR and presentation approach based on the human factors balance model applied to individuals using a search engine to retrieve information and on interfaces for IR systems. We first present some related work, then some ergonomic considerations related to user interfaces for retrieval systems. In the fourth and fifth section user-friendly semantic search and an interface modeling this search are presented. The sixth section is concerned with further discussing our results and pointing out some future work.

2 Related Work

User interfaces for IR systems have been studied during the past four decades. The information access process is usually assumed to look as follows: For each information need, a query is formulated and a set of documents retrieved. This information is then analyzed by the user and the results of the analysis are then used to successively alter the input query in such a way that it represents the user needs in a more accurate way [8]. Yet other models such as the Bates' "berrypicking" [1], which takes into consideration the shift in interest of the user consuming retrieved information, are also valid for some users. There are different theories concerning how these research strategies can be supported. According to Hearst [3, p. 265] most of them can be categorized as contrasting browsing, querying, navigating or scanning.

Some modern interfaces try to assist the user during her search process by displaying large information spaces, through which the user can navigate [7]. Yet such approaches are likely to present too much information to the user. Therefore, different viewing operations such as view changes [6], zooming and panning [2], etc. have been defined. The major drawback of these techniques lies in the fact that they demand supplementary interaction from the user during the retrieval process. To achieve similar goals the category-based approach uses taxonomies or similar structures to cluster information and guide the users during their search [4]. They present drawbacks similar to those of the information space approach. Furthermore they potentially lead to a labeling problem, since different categories of users might tend to use different labels for the same categories. A possible solution consists of learning the labels or the meaning of the labels for users or user groups.

Another category of information that could be used to guide users is their context, since the demands to retrieval systems differ depending on the user's current task. Especially in companies, IR is a support process, designed to find information supplementary that complement the input information necessary to efficiently complete a given task. It is a much more controlled process than web retrieval and operates on a small number of data sources. Some interfaces utilize this fact and integrate views on the data sources in which the search should occur. Yet the users who know the content of a data source usually know where to find the relevant data and do not need to search. The fact that they search can be assumed to imply that they do know where the information they need can be found.

None of the methods presented above consider the drawbacks linked to the standard two-fold search process of IR, which consists of searching for relevant documents and subsequently searching for relevant information. Hearst calls these phases search and analysis [3]. We tackle this problem by providing the user with fine granular content and integrating implicit feedback collection in the interaction with the content. The design of the proposed interface is based on considerations on interactions between the user and her working environment, especially the technology she uses and the user himself. As theoretical basis, we use the human factors balance model.

3 Ergonomic Considerations

The interactions between users and their working environment can be modeled using a wide range of theories. We will consider the factors on a user retrieving information using the human factors balance model (HFBM) introduced by Smith and Sainfort [9]. It is a holistic approach to the factors, which potentially influence a user during the completion of tasks in the work context. The HFBM subdivides the influences on a user completing a task in four main points (see Fig 1): environment, task, organization and technology.

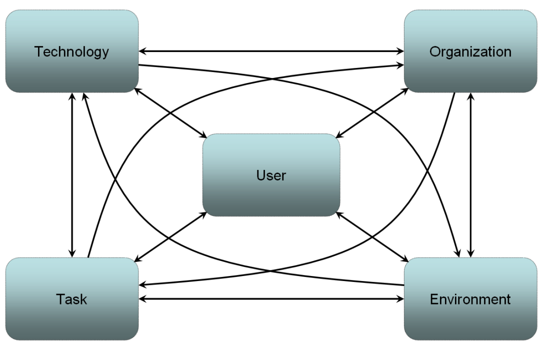

Fig. 1. The elements of a work system according to [9]

3.1 Environment

Meant here is the physical environment of the user. This aspect of the human factors can hardly be integrated in a user interface and will not be considered in our approach. Yet the "semantic environment", i.e. the set of cognitive concepts that are activated in the mind of the user when she retrieves information in a given context will be taken into consideration. Modeling the knowledge of a user is a difficult if not impossible task. Nevertheless, it is possible to model the knowledge contained in the data source(s) the user retrieves information from, making it easier for her to map the concepts activated in her mind and the concepts available in data. The semantic environment predefines the information need of the user. Hence it also alters the relevance she gives to the information she receives from a retrieval engine. It is thus of great importance to give the user the possibility to define the concepts she considers as relevant in a given task. Different knowledge modeling formalisms can be utilized when modeling concepts. In this work, we will focus on the use of terminological ontologies for modeling concepts.

3.2 Task

The user's task is of primordial importance when trying to model her semantic environment since it alters the relevance of information. This is due to the fact that the relevance of information is bound its potential usefulness when trying to achieve the goals set by the current task, on which it thus depends. Also relevant for the successful completion of a task are the input and output information that are related to it. The knowledge about these pieces of information is modeled in the business process model along which the user works. Since in- and output of a process are of static nature for a given process instance, the documents related to them should be aggregated by the process execution engine, which controls the information flow between tasks in processes. Thus it does not belong to the supplementary information that users might need when completing a task. Therefore we will not integrate this information in the user interface proposed.

3.3 Organization

The organization the user works for defines the global setting in which the search occurs. It is responsible for providing the user with accurate information sources and defines the duties and rights of the user. The duties (i.e. the tasks) of employees are condensed in the definition of their role(s) in the organization. Each role equally gives them certain rights. From the point of view of information retrieval, capturing the role of users is essential, since it is necessary for capturing the global scope in which the information retrieval takes place. Furthermore considerations such as data security are directly linked to the role definition. Thus the role of a user defines the view of the user on information, i.e. it defines the set of semantic concepts the user might be interested in at all. This implies that each role should be mapped to a concept model, i.e. an ontology. Considerations such as data sources and rights management are much more of administrative nature and will not be further considered. We will suppose that the user can only be presented with information she can access.

3.4 Technology

The technology utilized to complete a task can produce an information overload or underload [10]. From the point of view of information retrieval systems, both information underload and overload can be observed. The query input interface of modern information retrieval system is usually restricted to a text field, in which the user can give in boolean queries (this also underlines the insufficiency of the model in the background). A direct implication of this interface is that the user must formulate queries using a vocabulary that might differ completely from the one used in the corpus she is searching through. Thus there is a gap between the information available in the system and the knowledge the user has about this information. Another information underload occurs when the user is presented with the result of a retrieval process. The information she receives and out of which she must determine the relevance of a document for her query is usually reduced to the metadata of the document or a short excerpt from the document, generated without any consideration of context. The information overload comes about because the context of the user is not taken into consideration when processing her query. Thus she has to search through a large set of irrelevant information before getting access to the information that might of interest for him. Furthermore the relevance the user assigns to the documents once the retrieval results are presented is not registered by the system. Although explicit relevance feedback methods have been developed to tackle this drawback, they are seldom accepted by the users [5]. The need is thus for implicit methods able to collect the users' feedback.

In the following, we present an implementation of the aspects presented above in a user interface.

4 User Friendly Semantic Search

Required for our approach is a representative text corpus of adequate size containing domain specific documents. The preparation process consists of three steps. In the first step, we create a word similarity graph from the text corpus. The result is a graph, whose nodes represent words that are typical for the domain and whose edges depict the similarity of two adjacent words. We will refer to these words as terms. In the second step we extract domain specific ontologies from the calculated similarity graph. The computed ontologies are then manually revised. Alternatively, we can use existing ontologies and map their concepts with the terms in the similarity graph by computing the relevance value for each term-concept pair. For example, the relevance value between a concept and a term of the same word can be set to 1 by default. Last, we model typical, search intensive tasks executed frequently in this domain. We mark relevant concepts of the ontology for these tasks. With this data (i.e. an ontology with a set of sub ontologies and an underlying similarity graph connected to the concepts of the ontology), we can start our user friendly semantic search.

The three steps that lead the user to an improved IR experience combined with an improved search performance are

1. Configuration of the search environment
2. Query Generation
3. Presentation of IR results

The first step consists in configuring the search environment, i.e. in defining the context in which the user searches, by enabling the user to define her current context and task, which provides additional information that can be used in the next step: query generation. Current search engines do not assist the user in formulating their query, although this is an important point of any retrieval process. The best search algorithm cannot supply satisfying answers, when the user does not utilize relevant terms to describe her information need. In the query generation step, we use the role-dependent ontology and the similarity graph to offer the user an overview of relevant concepts and terms linked with these concepts, thus helping her to find the relevant keywords. The third step is the presentation of the IR results, which are subsequently analyzed by the user. Instead of only presenting links to possibly relevant documents, links which would lead to the user having to search through the document to find relevant pieces of information, user friendly semantic search displays only relevant parts of the document, called chunks. The user is now able to better and faster estimate the relevance of a document compared to only guessing the content from the title.

5 Graphical User Interface for Semantic Search

Our graphical user interface (short GUI, see Figure 2) consists of three frames according to the three steps mentioned above, of which each holds special functionality implementing each step of the search process: The upper right frame, called options frame, provides the user with the parameters necessary to configure her search and context, thus implementing the first step cited above. After setting these, the user can start exploring the ontology for concepts, which may be of interest in her search request. This functionality is realized in the navigation frame. For a quick start, the user can select the start node from within a list of the alphabetically ordered concepts. When the user has finished selecting concepts or words, she can start the search for chunks, i.e. document segments, which will be displayed in the lower frame, the chunk viewer. It is possible to interact with the shown chunks. The underlying framework monitors this interaction and uses it to adapt the ontology. Each of the frames will be presented in more detail in the following subsections.

5.1 The Options Frame

In the options frame, the user can customize the search process. The option frame offers a variety of roles the user can choose from. These roles mostly describe a job, like web designer, programmer, graphics designer and so on. The selection of a role is equivalent to the choice of a domain specific ontology, which is displayed in the navigation frame (see 5.2). The user then selects a specific task, which actually narrows down the ontology by solely activating a sub-ontology. With this method, it is possible to present potentially relevant concepts, from which the user can choose the most relevant for her search. Each ontology is modified during runtime by user behavior, which is collected using implicit feedback techniques. When the user activates a concept of the displayed ontology in the navigation frame, the framework checks the relevance values of terms in relation to the chosen concept. The user is given the possibility of setting upper and lower similarity thresholds. Terms with similarity values higher than the upper threshold value are automatically added to the search query. Likewise, terms are added, when they have a similarity value equal or higher the upper threshold in relation to the just added term. If the similarity value is between the upper and lower threshold value, the term is displayed as being relevant for the chosen concept but not added to the query.

Additionally, the user can influence the display of the frames with check boxes in the option frame. It is possible to deactivate the display of interface elements like the concept list or the search word line to achieve a maximum visibility for the remaining elements, like the ontology.

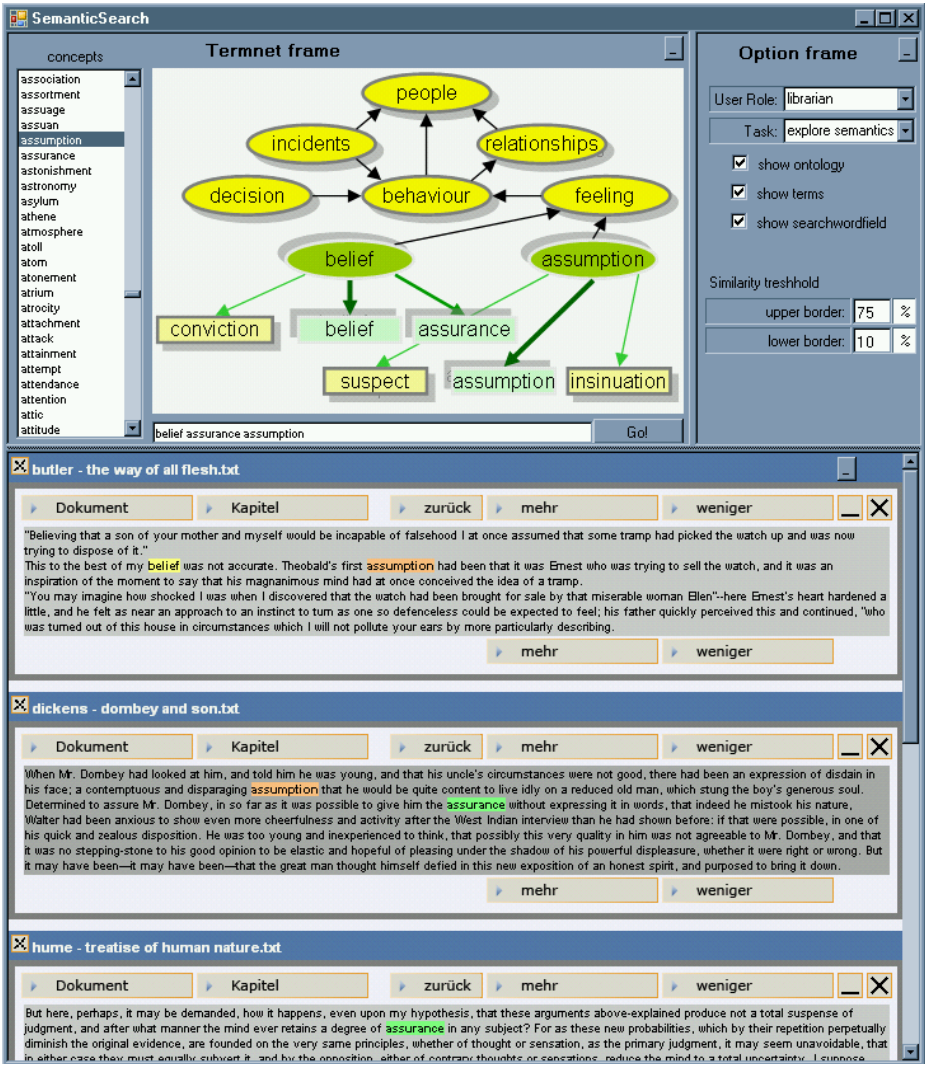

Fig. 2. The Semantic-Search GUI

5.2 The Navigation Frame

The navigation frame displays the part of the ontology that was activated with the selection of the task and user role in the options frame. Since this activated set can be rather large, the user has the possibility to select a first concept from a list box, which lists the concepts contained in the ontology. The window is then focused on this concept. The user can now examine the neighborhood of the chosen concept and can decide on the relevance of other concepts for her query. If the ontology does not fit into the navigation frame, the user can dynamically expand the frame, or drag the ontology inside the frame.

If the user considers a concept adequate, she can add the relevant terms with a double-click on the concept to the search query. The terms are extracted from the underlying term similarity graph, which forms the base of the ontology. In Fig. 2 concepts are displayed in oval boxes, terms from the similarity graph in square boxes. Terms that bear the same labels as the concepts are automatically selected and will be called direct terms from now on. The search query also automatically includes terms that have a similarity value to the direct term above the set threshold. These terms are called indirect terms. The added terms are indicated in the graphical view as child nodes of the chosen concepts. We use the color green to indicate the terms included in the search query. Terms with a similarity between the lower and upper threshold are also added as child nodes, but not to the search query. Therefore, they are displayed in yellow. To visualize the similarity of the terms with the given context, we show the connection arrows among concepts and terms in different brightness and thickness. The brighter the color and thinner the line, the lower is the similarity between the concept and the connected term. The user can now exclude or include words to the search query, simply by double-clicking the displayed terms. A direct input of the keywords is also possible.

The first instance of implicit user feedback recording occurs at this point: By deselecting a term from the similarity graph, the user produces negative feedback, which is learned by the system, leading to a decrease of the relevance of the given term for the concept *within the given context*. On the other hand, if the user adds a term to the search query, positive feedback is collected.

At the bottom of this frame is a text box, which holds the cumulated search query. It is possible for a user, to manipulate the query directly in this box, for instance, if she wants to insert words directly via copy and paste. If she has completed the exploration of the ontology and added all related search words, she can start the search using the Go! button.

5.3 The Chunk Viewer Frame

After the user has finished formulating her query, her focus changes to the third frame: the chunk viewer frame. Classical information retrieval systems show the search result as links to the retrieved document. In the majority of cases it is difficult for the user to estimate the relevance of the document. Consequently she needs to investigate the resulting documents to find out, whether they contain the information she is looking for. Often, this means to start another viewer, like the acrobat PDF-reader and to search through that document.

The chunk viewer frame uses a different approach. It shows the user document parts that most likely answer her question. To achieve this, we segment the documents into small coherent units we call chunks. Instead of showing only links to whole documents, we display sections of documents. This provides more information about the document than only the name of it. Ideally, it already contains the information, the user was looking for. At least she gets an idea about the content of the underlying document and whether it suits her need for information. This reduces the time she needs to investigate in analyzing search results, because she does not need to change her search environment and she does not need to search again through the results to encounter the exact location of the needed information.

The user can now decide to open the whole document (from which she already knows, that it is of high relevance), or she can use the functionality of the chunk viewer frame to get more information. She can e.g. expand the view of the chunk by the previous or next chunks or display the whole chapter in the chunk viewer. If chunks do not match with her search request, she can delete them from the result list. All those interactions are monitored by the system and stored in a database for further computation of the implicit feedback. The implicit feedback of a user's interaction with the concepts of the ontology and the terms of the similarity graph in the navigation frame, as well as the manipulation of the chunks in the chunk frame, result in a continuous re-adjustment of a term's relevance in the context of the user's role and task. Therefore, other users in the same role and task benefit from the re-adjustments generated due to the feedback of other users.

6 Discussion

We presented an approach that integrates human factors into the implementation of user interfaces for retrieval systems. We then elaborated on our vision of the process leading to user-friendly information retrieval, presenting an interface that implements the requirements generated from an analysis of existing systems and the HFBM. The proposed interface is constituted of three frames: the options frames, which allows the parameterization of the search process, the navigation frame, which enables the user to customize the search query and the chunk viewer frame, which display potentially relevant micro-content. The proposed implementation allows addressing some of the main issues discussed in section 2. It allows the user to browse through a representation of the available content, to query the content, to navigate through the retrieval results and eases the scan and analysis of retrieved document by displaying key micro-content directly in the browser. By these means, we implement all forms of interaction suggested by [3]. The complexity of the interface can be adapted to the user, since she can choose not to display certain windows. Our system uses a combination of information space and category-based browsing. A combination of these approaches might lead to an information overload. We address this potential drawback by learning through the user's feedback and thus ensuring that she is presented with a small number of contextually relevant concepts. Our system is based on the standard approach to information retrieval, which assumes an iterative reformulation of the user's query until the desired documents are found. Considering other approaches such as "berrypicking" might lead to some modifications of the interface.

References

1. Bates, M.J.: The design of browsing and berrypicking techniques for the online search interface. Online Review 13(5), 407–424 (1989)
2. Bederson, B.B., Hollan, J.D.: Pad: A zoomable graphical interface. In: Proceedings of ACM Human Factors in Computing Systems Conference Companion (CHI 95) (1995)
3. Hearst, M., Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. ACM Press/Addison-Wesley Longman Publishing Co., Harlow, England

4. Hearst, M.A., Karadi, C.: Cat-a-Cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. In: Proceedings of SIGIR-97, 20th ACM International Conference on Research and Development in Information Retrieval, Philadelphia, US, pp. 246–255 (1997)
5. Kelly, D., Teevan, J.: Implicit feedback for inferring user preference: a bibliography. SIGIR Forum 37(2), 18–28 (2003)
6. Mäkelä, E., Hyvönen, E., Sidoroff, T.: View-based user interfaces for information retrieval on the semantic web. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, Springer, Heidelberg (2005)
7. Robertson, G., Mackinlay, J., Card, S.: Information visualization using 3d interactive animation. In: CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 461–462. ACM Press, New York (1991)
8. Salton, G.: Automatic Text Processing – The Transformation, Analysis, and Retrieval of Information by Computer. Addison–Wesley, Reading, MA (1989)
9. Smith, M.J., Carayon-Sainfort, P.: A balance theory of job design for stress reduction. International Journal of Industrial Ergonomics (4), 67–79 (1989)
10. Smith, M.J., Cohen, B.G.F., Stammerjohn, L.W., Happ, A.: An investigation of health complaints and job stress in video display operations. Human Factors 4(23), 387–400 (1981)