An Aspectual Interface for Supporting Complex Search Tasks

Robert Villa, Iván Cantador, Hideo Joho, Joemon M. Jose
Department of Computing Science
University of Glasgow
Lilybank Gardens
Glasgow, G12 8QQ
{villar, cantador, hideo, jj}@dcs.gla.ac.uk

ABSTRACT

With the increasing importance of search systems on the web, there is a continuing push to design interfaces which are a better match with the kinds of real-world tasks in which users are engaged. In this paper, we consider how broad, complex search tasks may be supported via the search interface. In particular, we consider search tasks which may be composed of multiple aspects, or multiple related subtasks. For example, in decision making tasks the user may investigate multiple possible solutions before settling on a single, final solution, while other tasks, such as report writing, may involve searching on multiple interrelated topics.

A search interface is presented which is designed to support such broad search tasks, allowing a user to create search *aspects*, each of which models an independent subtask of some larger task. The interface is built on the intuition that users should be able to structure their searching environment when engaged on complex search tasks, where the act of structuring and organization may aid the user in understanding his or her task. A user study was carried out which compared our aspectual interface to a standard web-search interface. The results suggest that an aspectual interface can aid users when engaged in broad search tasks where the search aspects must be identified during searching; for a task where search aspects were pre-defined, no advantage over the baseline was found. Results for a decision making task were less clear cut, but show some evidence for improved task performance.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Search process; H.5.2 [**User Interfaces**]: Evaluation/methodology

General Terms

Experimentation, Human Factors

1. INTRODUCTION

With the continued integration of the World Wide Web into more and more work and social situations, web search systems are being increasingly used in many different scenarios. As reflected by models of information retrieval interaction, such as the cognitive model of Ingwersen [9], search tasks are part of the larger contexts and work tasks in which users are engaged – whether writing a school report, looking for a new job, or finding new questions for a pub quiz. Such tasks may sometimes require the user to search for many different aspects or perspectives about a single topic, or may involve a learning process, where the user must learn to make sense of the task as he or she searches. Such types of information seeking situations have been considered by the area of "exploratory search" [7, 14, 15, 20].

In this paper, we consider the exploratory search problem as one of aiding the user in the elicitation of their search requirements, and by extension, aiding their understanding of their possibly uncertain search task. One of the ways in which users make sense of the world is by classification, or as is stated by Lakoff [11]: "An understanding of how we categorize is central to any understanding of how we think and how we function". Our intuition is that by providing users with mechanisms which enable them to categorize while searching, we may be able to aid users in carrying out complex search tasks which may be difficult with existing systems.

There has been much work in Information Retrieval research which has taken advantage of classification in the search interface, most notably the work of clustering as an aid to the understanding and visualization of search results [5]. Other exploratory search techniques, such as faceted interfaces [7, 14, 15] also take advantage of categories by structuring the space users search via facets. In this paper, however, we take a different approach, where we wish to investigate whether allowing the users to explicitly structure their searching, i.e. allowing a searcher to classify his or her searches, to aid the user in both performing the task (discovering relevant documents), and with better understanding the search task that is being carried out.

In order to support this, we have developed an *aspectual* search interface, which allows a user to define multiple aspects, where an aspect allows a user to both search and mark relevant web pages in a bookmark style area. Aspects provide a way for users to classify and organize both their searching process, and the results of their searching process: i.e. each aspect provides a separate history of the searches which have occurred in that aspect plus provides an area where web page results can be stored and associated with the aspect. The interface as a whole is designed to facilitate the creation and organization of the aspects themselves, and the data stored within those aspects.

For example, consider an Arts student engaged in writing a report about the Viennese Secession. The student may want to investigate and write about some of the main people involved in the movement, such as Gustav Klimt and Josef Hoffmann; he or she may want to write about events which occurred, such as

SIGIR'09, July 19–23, 2009, Boston, Massachusetts, USA.

exhibitions held; or the student may want to write about particular paintings which were created by the members of the secession. All of these different elements we consider aspects of the same task – all aspects which together are related both to the overall subject matter, but which may also be interrelated together in unknown ways. Additionally, the student may learn about new people, events, or other facts while searching. From reading about Josef Hoffmann, the student may discover Koloman Moser, from which new facts may be discovered in turn.

1.1 Research questions

In this paper, we have three principle research questions:

RQ1: Does the aspectual interface allow the user to better explore and discover relevant material when compared to a purely sequential interface?

RQ2: Does the aspectual interface aid the user in better understanding the search task?

RQ3: What features of our aspectual interface are used by the users carrying out the search tasks?

Research questions 1 and 2 are the central questions addressed in this paper: can our aspectual interface, which allows the user to structure their searching, lead to the user discovering more relevant material when compared to a standard interface, and secondly, via this searching process, do users better understand the task they have carried out? Research question 3 is a more open ended, where we are interested in investigating how users take advantage of the different features of the aspectual interface. In order to test our ideas concerning the potential utility of aspectual search, and our aspectual interface, a baseline interface was created for comparison purposes, which restricted the user to carrying out one search at a time, as is common in current information retrieval interfaces.

The rest of the paper is structured as follows: one of the lessons learned from our previous work in [19] was the importance of the task type, and hence, in the next section, we provide a background section on some previous work into complex tasks. Following this is a description of the experimental design of our study, before our aspectual search interface and the baseline interface are described. The results are then presented, followed by a discussion. The paper finishes with a previous work section, followed by conclusions and future work.

2. COMPLEX SEARCH TASKS

There are many ways in which a user search or work task can be considered "complex", and there has been work in various different fields which has looked into this question. From the information science, Bystrom and Kalervo [3] consider the relationship between task complexity and information search and use. Five 'task categories' are defined: genuine decision task; known genuine decision task; normal decision task; normal information processing task; and automatic information processing task. These five categories are based on the uncertainty inherent in the tasks: in automatic information processing tasks, the task solutions and types of information required are all known in advance. At the other end of the scale, in genuine decision making tasks nothing is known about the types of information or solutions required.

The work of Bell and Ruthven [1] also uses the uncertainty inherent in a task as a measure of complexity, and presents a study which systematically alters search tasks in order to control task complexity. More complex topics were created by removing specifics from the topic statement, providing the user with less knowledge of the task to be undertaken. A similar method is used in the study reported in [3]. The work of Vakkari [18] is an attempt to synthesize a number of studies into a single model, and takes a similar approach, where task complexity is related to the lack of structure in a task, and the lack of knowledge a user has of the task.

An alternative perspective on task complexity is provided by Campbell [4], who reviews approaches to task complexity in various research fields. Based on this, he outlines four basic task characteristics which define what he calls "objective complexity", i.e. complexity which is based solely on the task, and not on the degree of the user's knowledge of the task. These are: (a) the presence of multiple paths to a solution; (b) the presence of multiple desired outcomes; (c) the presence of conflicting interdependences among the paths to the solution(s); and (d) the presence of uncertain links among the paths and outcomes. Task complexity is therefore defined as more than uncertainty, but also involves the presence of other factors, notably the possibility that complexity may also be due to the task having multiple solutions, or multiple paths to a single solution. This model of task complexity has provided an impetus to the aspectual interface described here, and has motivated the tasks created for the user study, described in Section 6.

3. EXPERIMENTAL DESIGN

In order to investigate the research questions outlined in Section 1.1, a between subjects user study was conducted. Based on the model by Campbell [4], we constructed three tasks, each of which were designed to represent three different possible types of complex task. Two separate groups of users then carried out the same three tasks on two different interfaces – the aspectual interface (described in Section 3.1.1) and a baseline interface (Section 3.1.2). The same underlying web search engine, Yahoo BOSS[1], was used for both interfaces.

Based on research questions 1 and 2, we created two hypotheses: first, that the aspectual interface would allow a user to investigate the tasks to a greater extent, measured by the documents marked as relevant, the results viewed, the searches carried out, and the query vocabulary size. Secondly, we hypothesized that the user perception of task complexity and difficulty would fall significantly for the aspectual interface, but not the baseline, when comparing a user's perceptions of the task before and after carrying out the search.

3.1 SEARCH INTERFACES

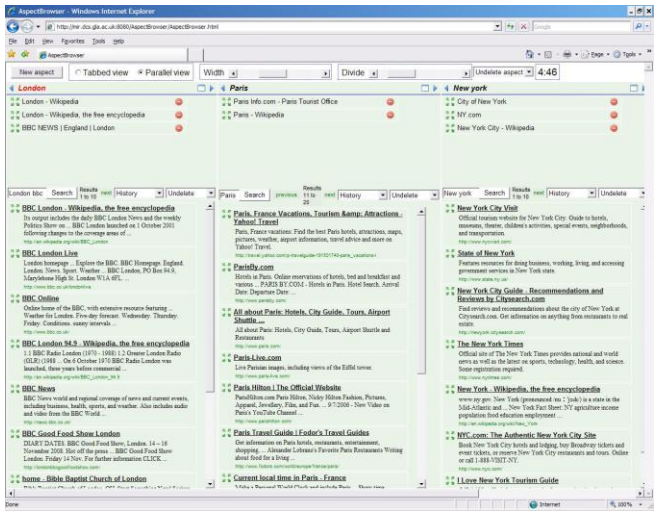
3.1.1 The aspectual interface

The aspectual interface is a development of the system described in [19]. Screenshots of the interface are shown in Figure 1 (a) and (b), showing the two different views supported. The interface is built around the concept of search aspects, where each aspect contains the following elements: (1) a name, which is by default set to the last query executed, but which can be explicitly set by the user when desired; (2) a list of selected documents, i.e. the web pages which the user judges as being relevant to the aspect; (3) the current search query entered by the user; (4) the list of search results for the current query, where clicking on the document's title will display the corresponding web page in a

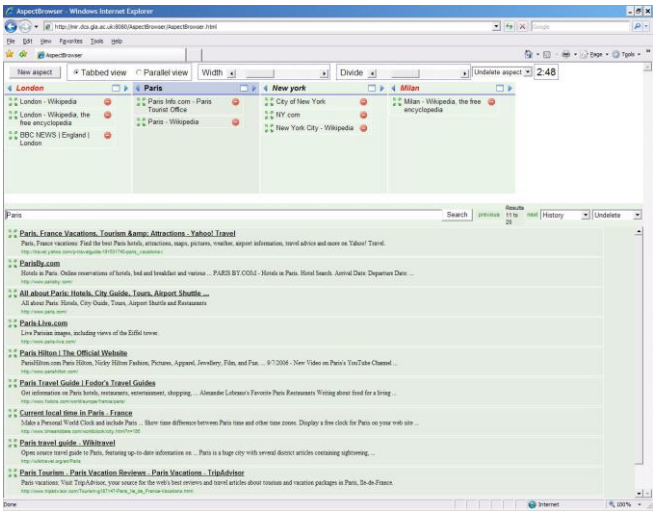
[1] http://developer.yahoo.com/search/boss/

pop-up window; (5) a history of the searches carried out in the aspect, and a list of any deleted documents; and finally, (6) the position of the aspect within the overall sequence of aspects.

It should be noted that an aspect is a self-contained entity containing all of the above states: each aspect has its own query history, undo history, current query, search results, etc. The "new aspect" button at the top left of the interface creates new aspects, adding them to the far right of the display. At startup, a default (empty) aspect is created and shown.



(a) Parallel view



(b) Tabbed view

Figure 1: The aspectual search interface showing the same information in two different views

The interface can be operated in two different visualizations: parallel view (Figure 1a) where each facet is displayed as a vertical slice of the screen, from left to right; and tabbed view (Figure 1b) where the name and examples of each aspect are shown on clickable "tabs" at the top of the interface. By clicking on an aspect, the rest of the screen fills to show the associated search query and results for that aspect.

The parallel view is similar to the interface in [19]. While it was liked by many users, the lack of space given over for search results caused problems for others, which resulted in the modified tabbed view shown in Figure 1b. The buttons at the top left of the interface allow the user to toggle between these two visualization. It should be noted that the same information is displayed in both modes, with only the presentation of the information changing.

The interface makes extensive use of drag and drop. Using the crosshair icon to the left of each search result, documents can be added to an aspect's marked list by dragging them onto the corresponding screen area. There is no restriction on which aspect a result can be dragged onto, therefore it is possible to drag a search result from one aspect directly onto the relevance list of a different aspect. Documents can also be dragged and dropped between different aspects, allowing the reorganization of material across the aspects, and can also be re-ranked within each aspect.

Relevant shots can be removed from the example lists using a delete button, which will add the document to the "undelete" drop down list for the aspect. Searches always return 10 documents, "next" and "previous" buttons allow the user to see more when necessary. In addition to these standard features, a pull down menu on each aspect allows a user to rename the aspect, duplicate the aspect, delete the entire aspect, and add a description. Finally, arrow buttons on each aspect allow the user to re-order the aspects, by moving aspects left or right by one.

In both this interface and the baseline interface described in Section 3.1.2, only web pages which are listed as a search result can dragged and dropped, and therefore marked as relevant. While the user can potentially browse to other pages, these browsed-to pages cannot be marked, thus restricting search results to only those found via the search engine.

3.1.2 Baseline interface

The baseline interface was a simplified version of the aspectual interface, where only a single panel was used to compile relevant results obtained during a searching session, independently of being related to different aspects or topics. The purpose of this baseline version of the system was to allow us to conduct the experiment in such a way that the aspectual interface could be compared against a classic web search interface.

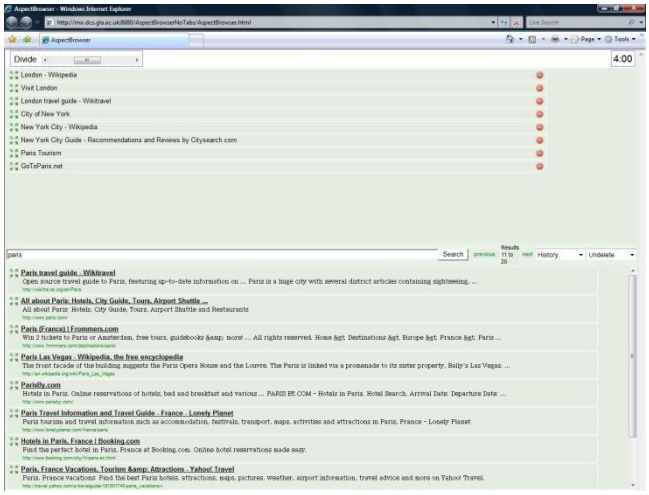


Figure 2: The baseline search interface

In this interface, no tabs are permitted so a user cannot organize the documents into different aspect-oriented panels. Figure 2

shows a screenshot of the interface. It consists of two main panels separated by a search bar. The bottom panel contains the list of documents retrieved by the query, and the top panel displaying the list of documents marked as relevant by the user. Functionality, the interface operates in exactly the same way as the aspectual interface, except there is only ever a single "aspect": e.g. the user can make use of the drag and drop functionality to mark documents as relevant, alter the order of the documents, etc.

3.2 TASKS

Based on the categories (a) to (c) of Campbell [4], summarized in Section 2, three different types of search task were defined:

A. Single solution, multiple potential paths to the solution

B. Multiple solution, where the aspects are largely independent, and are largely specified by the task

C. Multiple solution, where the aspects are implicit in the task and may be interdependent

The first task, A, was written as a decision making task, where the user was asked to find and decide on a new digital camera. The second task, B, was setup as a report writing task, where the user had to find information about the political leaders of four specified countries. The 'solution' was therefore plural, and the user was given the option to search for more than the specified minimum. The third task, C, was a summarization task, where the user was given some general background information for the task, but where specific aspects were not directly provided. The aim here was to force users to find and categorize the search task as they see fit. However, unlike Task A, the solution was again plural – users were asked to find multiple aspects. Each task was written as a simulated work task [2], which included a simulated situation, describing a context for the user, as shown in Table 1.

Our choice to carry out the evaluation on the web was a practical one – we wanted a document collection which was large enough, and rich enough to allow a user to explore material without hindrance.

3.3 PROCEDURE

In total, thirty-six users were recruited through email to take part in the study, split into two groups of eighteen users: one group of eighteen performed each of the three tasks with the aspectual interface, the other group of eighteen performing the three tasks with the baseline interface. The median age was 26 (with range 22 to 39) with the majority of users being either native or near-native English speakers. Most subjects were students or post-graduates at Glasgow University, all having university degrees or higher.

After arriving at the office where the study took place, users were welcomed before being presented with an information and consent form. After these preliminaries, an entry questionnaire was administered, before the experimenter then demonstrated the search interface (either the aspectual or baseline interface). This demonstration took approximately 10 minutes for each user, and was followed by a training task, where the user was allowed up to 15 minutes to interact and use the system.

After training, each of the three tasks was administered in an order which ensured counterbalancing. Before each task, the task description was presented to the user who was then able to read it, before a pre-task questionnaire was filled in. After this, the test interface would be started and the subject commenced searching. Both interfaces contained an automatic timer, always present to the user at the top right hand corner of the screen, showing the length of time which had elapsed. After 20 minutes, an "end task" dialog box would appear indicating the end of the task, although users were informed that they could end the task before the 20 minutes if they were satisfied with their search results. After each task, a post-task questionnaire was administered. This was repeated three times, one per task.

Table 1: Situated work tasks for the three tasks

Task	Simulated situation
A	You are looking for a new digital camera to replace your old film camera before going on a safari holiday in Africa. You would like to find out which models of camera are best for you and for your holiday trip. While liking photography, you would like something small and easy to use, which can also be used by your partner who dislikes technology.
B	Imagine you are a student working towards a media studies degree at the Open University. As part of your 2rd year "politics and the media" course, you have to write a report about some of the most influential political leaders currently in the news. This report is to provide a brief biography of each politician, of up to 2000 words, and should include a brief history of the most important issues which the leader has had to deal with over their term of office. You should emphasis the recent events which the leaders of each respective country had to deal with (i.e. those events which have occurred over the last year). Your lecturer has stated that, as part of the reports requirements, you MUST provide a biography of current presidents or prime ministers of the following countries: USA, Germany, Japan and the UK. In addition to the leaders from these four required countries, you may also include any other world leader you wish. The report must be written in English.
C	You work as a researcher for a well known UK newspaper, and as part of your remit, you work on a monthly news magazine, which summarizes and reviews the main sporting events of the previous month. It has now become necessary to put together the latest magazine for August 2008, covering the Olympics in Beijing. The magazine is 25 pages long, and typically covers between 4 to 8 main stories of a page or more, and a further unknown number of much shorter stories and news pieces, often only a paragraph long. You are responsible for researching the main stories from the Olympics, putting them in order of importance, and also finding as many other shorter stories as you can which are appropriate for the magazine.

At the end of the experiment, a final exit questionnaire was completed by all users. The procedure took up to 2 hours, and all users were paid 12 pounds on completion.

4. RESULTS

The results of the study are presented in this section, split to reflect the research questions outlined in Section 1.1. Unless otherwise stated, all results are described using medians and the interquartile range, due to the often asymmetric nature of the data; likewise the non-parametric Wilcoxon rank-sum test is used unless stated otherwise, with significance level $P \leq 0.05$.

4.1 Search performance

The first research question looks at the search performance of users using either the baseline or aspectual search interfaces. In particular, we are interested in measuring the degree of

exploration of the user, assuming that a user who is able to explore more of an information space is more likely to satisfy a task. We do not use precision and recall: the broad nature of the tasks makes defining a single set of relevance judgments difficult, and given the use of web search, likely impossible. Instead, performance is measured using the number of web pages marked as relevant, the number of search results viewed, the number of searches carried out, and the query vocabulary size. Table 2 shows the results of these four measures split by task and system.

Table 2: Number of web pages marked relevant, search results viewed and searches carried out. The median (and interquartile range) values are reported, bold indicating relevance at $P \leq 0.05$

	Task A	Task B	Task C
Marked relevant			
baseline	8 (6-9)	17 (11.3-20.5)	11 (8.5-25.3)
aspectual	**14 (9.3-20.8)**	22.5 (16.5-37.8)	**24 (18.5-32.5)**
Results viewed			
baseline	17 (11.3-21.0)	24.5 (22.3-29.5)	17.5 (14.3-24.5)
aspectual	22 (15.3-28.5)	23.5 (12.5-44.3)	**29.0 (17.0-43.3)**
Number of searches			
baseline	7 (5-9.8)	17 (16-20.8)	9.5 (7.3-15)
aspectual	**10 (7-13.8)**	18.5 (16.3-25.8)	**13 (12-16.8)**
Query vocabulary size			
baseline	11.5 (9-15)	21.5 (20.3-24.8)	15 (8.3-21.0)
aspectual	14.0 (11.0-16.0)	25.5 (21.3-30.5)	**21.0 (16.0-24.8)**

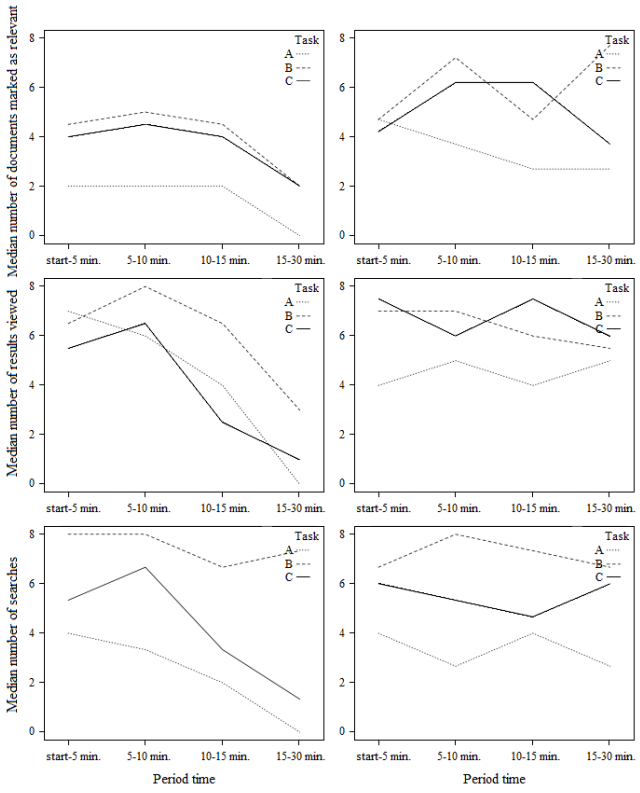
Considering each task individually, it can be seen that for Task C (multiple solutions, implicit aspects) there is a clear trend in all measures for an increase in performance for the aspectual interface when compared to the baseline interface: a statistically significant difference was found in all four cases at $P \leq 0.05$. Conversely, for Task B, no significant differences were found between the baseline and aspectual interfaces for all four measures. Lastly, for Task A, the picture is mixed: significant differences between the baseline and aspectual interfaces were found for two measures (number marked relevant, and number of searches), but not for the others.

One issue which is not considered in Table 2 is that of search time. All users in both conditions were informed before starting that they had up to 20 minutes to search for each task, but could finish early if they felt their task was finished. The instructions presented to both sets of users were the same. Table 3 shows the median task length for the two interfaces, and three tasks. As can be seen, there is a trend for users to spend less time searching with the baseline interface than with the aspectual interface. Out of the 54 sessions with the aspectual interface, only 7 finished before the maximum 20 minutes, with only two sessions finishing in less than 19 minutes (one session finished in 16 minutes, another 17 minutes).

Table 3: Median (interquartile range) for task time in minutes

	Task A	Task B	Task C
Baseline	13.8 (12.4-19.4)	17.9 (16.2-20)	17.3 (11.7-19.0)
Aspectual	20 (20-20)	20 (20-20)	20 (20-20)

All four of the measures reported in Table 2 consider each search session as a whole, irrespective of how long the user searched. It is also possible, as suggested by Kaki [10], to take time into account, i.e. to measure the number of web pages marked as relevant, viewed, or the number of searches carried out per minute. When this adjustment is made to the results in Table 2, no significant differences are found between the two systems.



(a) Baseline **(b) Aspectual**

Figure 3: Search performance over time for number marked relevant, number documents viewed, and number of searches. Baseline on left, aspectual on right

Figure 3 (a) and (b) present this same information, for number of documents marked as relevant, results viewed and searches carried out, but shown grouped into four 5 minute bins over time. On the left on Figure 3 (a) are the graphs for the baseline interface, on the right, the aspectual interface. It should be noted that these values are not cumulative – each bin represents the activity for only that 5 minute period of time.

For the baseline results, a clear fall in user activity can be seen in the final 5 minutes of time, reflecting the number of users who stopped before the full 20 minute period was up. An exception to this is the number of searches executed on Task A. The activity over time for the aspectual interface is somewhat more constant over the full 20 minutes, with activity roughly constant or at times increasing during the final 5 minutes. The graph for the change in query vocabulary size (not shown due to space) is similar to those shown in Figure 3, with the number of unique terms falling off in the final 5 minute period for the baseline interface, while remaining more stable for the aspectual interface.

4.2 User perceptions of search task difficulty

The second research question from Section 1.1 asks whether the aspectual interface aids the user in better understanding the search task. To attempt to measure this, as part of the pre- and post-task questionnaires, we asked each searcher to judge the complexity and difficulty of the task, before and after carrying out the search. If using the aspectual interface leads to a greater understanding of the task, we hypothesize that there will be a significant reduction in perceived task complexity and difficulty for the aspectual interface, and similarly a non-significant reduction for the baseline interface.

Table 4: Responses to the pre and post task question "The task [will be/was] complex], 1 = disagree, 5 = agree; median (interquartile range), bold indicates significance $P \leq 0.05$

		Task A	Task B	Task C
The task will be / was complex				
Baseline	Pre	2.5 (2-3.75)	3.0 (2.25-4)	3.5 (2-4)
	Post	2 (1.25-3.75)	2.5 (2-3)	3.0 (2-4)
Aspectual	Pre	2 (1-4)	3.0 (2.25-4)	**4.0 (3-5)**
	Post	2 (1-4)	2.0 (2-3)	**2.5 (2-3.75)**

Table 5: Responses to the pre and post task question "The task [will be/was] difficult], 1 = disagree, 5 = agree; median (interquartile range)

		Task A	Task B	Task C
The task will be / was difficult				
Baseline	Pre	3 (2-3)	3 (2-4)	3 (2-4)
	Post	2 (2-3)	3 (2-3.75)	3 (2-4)
Aspectual	Pre	2 (1-4)	3 (2.25-4)	4 (3-4)
	Post	2 (1-4)	3 (2.25-4)	3 (2-4)

In Tables 4 and 5, the median responses are presented for the pairs of questions "The task will be complex"/"The task was complex" and "The task will be difficult"/"The task was difficult". Both the responses in the pre-task and post-task questionnaires are given; Wilcoxon rank-sum tests were used to compare the pre and post task judgments, with one significant result, that of users judgment of complexity with Task C ($W = 242$, $P \leq 0.05$).

4.3 Interface features and user perceptions

For research question 3, we present some overview statistics concerning the usage of the extra facilities of the aspectual interface (Table 6). It can be seen that roughly similar numbers of aspects were created by all users for all tasks (medians 5 or 6). Very few aspects were deleted – in total 39 aspects were deleted by 9 users over the 54 sessions.

Few examples were also copied between aspects: over the 54 search sessions with the aspectual interface, 69 aspect to aspect copying events were recorded, across 26 sessions. Out of the 18 users, 4 did not move bookmarked documents between aspects at all. Others, such as users 9 and 12, used this facility more, with 15 and 13 recorded aspect to aspect moves for these two users.

Similarly, documents marked relevant were also rarely deleted. Over all users, 78 document deletions were recorded, and again, 4 users did not delete any documents once marked, while others deleted numerous documents (user 14 deleted ten documents, and user 3, eight). By comparison, only 7 undelete document events were recorded over all sessions.

Table 6: Median number of events with aspectual interface; median (interquartile range)

	Task A	Task B	Task C
Number of aspects	6 (5-7.75)	5 (5-7.75)	6 (5-7.75)
Copied between aspects	0.5 (0-2)	0 (0-2)	0.5 (0-2)
Delete relevant	0 (0-1)	1.0 (0-1)	0 (0-1)

The aspectual interface supports two different visualization views, as shown in Figure 1. Table 7 shows the mean length of time users spent in each of these views, showing a distinct preference for the tabbed view over the parallel view. Out of the 54 sessions, 39 were carried out for a majority of the time in the tabbed view, the other 15 for a majority of the time in the parallel view.

Table 7: Length of time spent in the tabbed and parallel views, mean (standard deviation), in minutes

	Task A	Task B	Task B
Tabbed	13.5 (9.0)	13.0 (9.2)	16.0 (7.5)
Parallel	6.5 (9.0)	7.0 (9.2)	3.7 (7.1)

Finally, we present the results of the exit questionnaire which asked a number of usability questions based on the USE questionnaire [12], shown in Table 8. A Wilcoxon rank-sum test found that the question "The system would help me be more effective" was significantly different at the 5% level ($W = 91$, $P \leq 0.05$), users agreeing with this statement significantly more for the aspectual interface than the baseline. A significant difference was also found for the question "The system can be used effectively without instruction" ($W=270.5$, $P \leq 0.05$), this time the baseline interface being judged as being easier to use without instruction. All other comparisons were found to be not significant.

5. DISCUSSION

Concerning the results for research question 1, in Section 4.1, there would appear to be strong evidence that the aspectual interface does aid the user, based on the four performance measures, for Task C (multiple solutions where the aspects are implicit). Conversely, for Task B (multiple solutions where aspects are explicitly provided), we found no significant difference between the baseline and aspectual interfaces. This is perhaps not so surprising given that in Task B the user does not have the added burden of determining aspects themselves, instead,

the organization is implicit within the task description. This is consistent with the thesis that the aspectual interface aids the user in performing a classification of the task (i.e. determining the appropriate aspects): if the aspectual interface does not aid the user in specifying aspects, we may expect both tasks B and C to produce similar results, which is not the case here.

Task A, the decision making task, is less clear cut than the other two, the results shown in Table 2 are mixed. Given that users do mark more relevant documents, and carry out more searches, we can say there is some tentative evidence that the aspectual interface does aid the user with this type of task.

Table 8: Results of the exit questionnaire, medians (interquartile range) where 5 = agree, 1 = disagree

Question	Baseline	Aspectual
The system is simple to use	4.5 (4-5)	5.0 (4-5)
The system is wonderful	3 (3-4)	4 (4-5)
The system is useful	4 (3.25-5)	5 (4 – 5)
The system is flexible	3.5 (3-4)	4 (4-5)
I learned to use the system quickly	5 (5-5)	5 (5-5)
The system would make things I want to accomplish easier to get done	4 (3 – 4.75)	4 (4-5)
The system would help me be more effective	3.5 (2.25-4)	**5 (4-5)**
Using the system is effortless	3 (3-5)	4 (3.25 – 5)
The system can be used effectively without instruction	**5 (4-5)**	4 (2.25 – 4)
The system was easy to learn to use	5 (5- 5)	5 (4.25 – 5)
The system works the way I want it to work	4 (3 – 4)	4 (4 – 4.75)
The search methods I used in this study were similar to those I use when I normally search the web	4 (3 – 5)	4 (3.25 – 5)
Overall the system is easy to use	5 (4 – 5)	5 (4 – 5)
If this system were available for use I would use it frequently	4 (3-4.75)	4 (4-5)
I am satisfied with the interface	3.5 (3-5)	4 (3.25-5)

One issue with the results presented in Table 2 is that of task time. As shown in Table 3, users tended to spend less time searching with the baseline when compared to the aspectual interface. This is also supported by the graphs of performance over time shown in Figure 3, where it can be seen that user activity with the aspectual interface is broadly similar to the start of the task. In contrast, there is a trend for user activity with the baseline interface to decrease over time for the three measures (although Task B shows evidence that users are still searching until the end of the task, as shown by the bottom graph in Figure 3a).

It is possible to interpret this in at least two ways: one could suggest that baseline users succeed at completing the task faster; alternatively, it can be suggested that the extra organizational facilities of the aspectual interface encourages users to search more, the shorter task times recorded by baseline users being indicative of boredom or an otherwise lack of impetuous to search more. We tend to err on the latter interpretation for a number of reasons. First, there is no performance difference between the interfaces on Task B suggesting that the baseline interface *can* perform as well as the aspectual interface, when the task explicitly specifies the aspects to search for. While not significant, it is also to be noted that the baseline task time in Table 3 for Task B is longer than for the other tasks, and the final five minute period of search activity graphed in Figure 3(a) for Task B shows more activity than the other tasks, especially for searches carried out. This extra activity at the end of the task is consistent with the extra aspect information allowing the users to search for longer with the baseline interface, therefore enabling a search performance more in line with the aspectual interface.

Turning to the second research question, and the results shown in Tables 4 and 5, no significant differences were found in user perceptions of task complexity or difficulty before and after the task, with the exception of Task C and complexity with the aspectual interface. This is consistent with the aspectual interface aiding the users understanding of the task, where in this case aspects must be identified by the user. I.e. after carrying out task C with the aspectual interface, users considered the task *less* complex, but *equally* difficult. This result backs up the previous results which suggest that users gain the most from the aspectual interface when performing Task C.

Lastly, individual features of the aspectual interface were analyzed, and the results of the exit questionnaire presented. These show mixed usage of the two different views, with a tendency for users to prefer the tabbed visualization (Figure 1a). Organization facilities were lightly used by most users (Table 6), most documents which are marked relevant not being moved between aspects. User feedback from the exit questionnaire suggests that the aspectual interface does require instruction when compared to the baseline interface, although user perceptions of ease of use and ease of learning do not vary significantly between systems. User perceptions of effectiveness, however, do vary significantly, with users suggesting that the aspectual interface would allow them to be more effective.

6. PREVIOUS WORK

As has already been stated, the work reported here is largely an outgrowth of [19]. In that work, a video retrieval interface, called 'FacetBrowser', was developed to allow users to create multiple 'facets' in the interface. These facets are similar to the aspects described in this paper, although the interface itself is different – the name change from facet to aspect was intended to reduce confusion with faceted search interfaces such as [7], [15], and [17]. The interface was inspired by the "storyboards" which are commonly used by artists and directors in the film industry.

Subjunctive interfaces [6, 13] are a type of interface which is similar in intention to the one described here, although more general in scope and design. Subjunctive interfaces aim to support mechanisms which allow users to view, manipulate and control multiple scenarios, to allow a user a greater degree of exploration than possible when working with interfaces with only a single result display. Lunzer and Hornbæk [13] describe three different subjunctive interfaces – a data browser, a simulator of ant behaviour, and a document editor, which allow users to create and compare multiple scenarios, and compare these scenarios on screen at the same time. In Fujima et al. [6], an interface called C3W is described which allows the user to create scenarios from existing web pages. These scenarios can be cloned, and the user is able to execute multiple queries in each scenario simultaneously.

Exploratory search [7, 14, 20] is an emerging area of Information Retrieval research which focuses on search tasks which are ill-defined, where the user must learn while searching, or which require a degree of browsing. Hearst [7] is an early example of a paper proposing exploratory search, and gives a definition of *facet* as an attribute which can be used by the system and user to split a database into disjoint partitions. Many modern e-commerce websites such as amazon.com and ebay.com use such techniques, allowing the user to browse products based on price, manufacturer, type, etc. Marchionini and Brunk [15] and Schraefel et al. [16] are two academic examples of this style of interface.

The definition of aspect as used in the TREC interactive track can be considered as related to the definition of aspect used in this work. The interactive track in TREC-5 [17] defines an aspect as "roughly one of many possible answers to a question which the topic in effect posed". Similar search topics were also used in TREC-7 and TREC-8. For example, topic 408i from TREC-8 [8] has description "What tropical storms (hurricanes and typhoons) have caused property damage and/or loss of life?", and its associated instances section asks the user to "find as many different storms of the sort described above as you can".

Finally, the tabbed functionality of modern web browsers such as Mozilla Firefox and Microsoft Internet Explorer must be mentioned. Tabs provide a simple mechanism which allows users to conduct multiple searches in different browser tabs, albeit this functionality is still basic – for example, it may be difficult to organize search results found in multiple browser tabs.

7. CONCLUSIONS AND FUTURE WORK

The results presented in this paper suggest that for tasks similar to Task C, where the user must identify aspects of a broad search task, an aspectual style interface similar to the one presented here has a clear advantage: users found more relevant web pages, viewed more results, and searched more. There is also evidence that by enabling the user to classify and structure their searching, the perceived complexity of the task decreased after the search session. This is in contrast to Task B, where task aspects were supplied predefined to the users: in this case there is no advantage to using the aspectual interface. For decision tasks such as Task A, the advantages of an aspectual interface are less clear cut: there is evidence that task performance improves, but it is not conclusive.

Future plans include investigating how "tab" functionality, as found in web browsers, is used and how this impacts searching. Additionally, it may be possible to alter or augment a browser's tab functionality, such as creating tab specific search and bookmark functions. This would allow users to carry out web searches and bookmark web pages within specific tabs, similar to the interface presented here.

8. ACKNOWLEDGMENTS

This research was supported by the European Commission, contract FP-027122-SALERO.

9. REFERENCES

[1] Bell, D. J., and Ruthven, I. 2004. Searchers' Assessments of Task Complexity for Web Searching. In ECIR 2004, Sunderland, UK, 57-71.

[2] Borlund, P. 2003. The IIR Evaluation Model: a Framework for Evaluation of Interactive Information Retrieval Systems. In Information Research, 8, 3, paper no. 152.

[3] Bystrom, K. and Kalervo, J. 1995. Task Complexity Affects Information Seeking and Use. In Information Processing and Management, 2, 191-213.

[4] Campbell, D. J. 1988 Task Complexity: A Review and Analysis. In Academy of Management Review, 13, 1, 40-52.

[5] Cutting, D., Karger, D., Pedersen, J., and Tukey, J. W. 1992. Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections. In SIGIR 1992. Copenhagen, Denmark.

[6] Fujima, J., Lunzer, A., Hornbæk, K., and Tanaka, Y. 2004. Clip, Connect, Clone: Combining Application Elements to Build Custom Interfaces for Information Access. In UIST 2004, New York, NY, USA.

[7] Hearst, M. 2000. Next Generation Web Search: Setting our Sites. In IEEE Data Engineering Bulletin, 23, 3, 38-48.

[8] Hersh, W., and Over, P. 2000. TREC-8 Interactive Track Report. In the 18th Text REtrieval Conference (TREC 8), Gaithersburg, MD, USA.

[9] Ingwersen, P. 1992 Information Retrieval Interaction. London: Taylor Graham.

[10] Kaki, M. 2004. Proportional Search Interface Usability Measures. In NordiCHI'04, Tampere, Finland.

[11] Lakoff, G. 1985. Women, Fire and Dangerous Things: What Categories Reveal about the Mind. The University of Chicago Press.

[12] Lund A. M. 2001. Measuring Usability with the USE Questionnaire. In Usability and User Experience, 8, 2.

[13] Lunzer, A., and Hornbæk, K. 2008. Subjunctive Interfaces: Extending Applications to Support Parallel Setup, Viewing and Control of Alternative Scenarios. In ACM Transactions on Computer Human Interaction 14, 4, 1-44.

[14] Marchionini, G. 2006. Exploratory Search: From Finding to Understanding. In Communications of the ACM, 49, 4, 41-46.

[15] Marchionini, G., and Brunk, B. 2003. Toward a General Relation Browser: A GUI for Information Architects. In Journal of Digital Information, 4, 1.

[16] Schraefel, M. C., Wilson, M., Russell, A., and Smith, D. A. 2006. mSpace: Improving Information Access to Multimedia Domains with Multimodal Exploratory Search. In Communications of the ACM, 49, 4, 47-49.

[17] Over, P. 1999. TREC-5 Interactive Track Report. In the 7th Text REtrieval Conference (TREC 7), Gaithersburg, MD, USA.

[18] Vakkari, P. 1999. Task Complexity, Problem Structure and Information Actions: Integrating Studies on Information Seeking and Retrieval Information In Processing and Management, 35, 819-837.

[19] Villa, R., Gildea, N. and Jose, J. M. 2008 FacetBrowser: a user interface for complex search tasks. In proceedings of ACM Multimedia 2008, 489-498.

[20] White, R. W., Kules, B., Drucker, S. M. and Schraefel, M. C. 2006. Supporting Exploratory Search. In Communications of the ACM, 49, 4, 36-39.