Focused Access to Wikipedia

Börkur Sigurbjörnsson[1] Jaap Kamps[1,2] Maarten de Rijke[1]

[1] ISLA, Faculty of Science, University of Amsterdam
[2] Archives and Information Studies, Faculty of Humanities, University of Amsterdam
{borkur,kamps,mdr}@science.uva.nl

ABSTRACT

Wikipedia is a "free" online encyclopedia. It contains millions of entries in many languages and is growing at a fast pace. Due to its volume, search engines play an important role in giving access to the information in Wikipedia. The "free" availability of the collection makes it an attractive corpus for information retrieval experiments. In this paper we describe the evaluation of a search engine that provides focused search access to Wikipedia, i.e., a search engine which gives direct access to individual sections of Wikipedia pages.

The main contributions of this paper are twofold. First, we introduce Wikipedia as a test corpus for information retrieval experiments in general and for semi-structured retrieval in particular. Second, we demonstrate that focused XML retrieval methods can be applied to a wider range of problems than searching scientific journals in XML format, including accessing reference works.

1. INTRODUCTION

Wikipedia [14] is a "free" online encyclopedia that can be edited by anyone. At the time of writing (February 2006), it contains a million articles in English as well as millions of articles in several dozens of other languages. Given the volume of the data, search engines provide an important tool for accessing the information contained in Wikipedia.

There are quite a few search facilities for Wikipedia [16]. The search engines differ both with respect to the indexing scheme and result presentation used. Some engines search over the full content while others only search over the title. Some engines display links to pages, with or without text snippets, while other engines cluster results by category.

In the area of semi-structured retrieval, focused information access has gained much attention, with direct access to relevant parts of documents being an important example. This is one of the major research issues addressed within the Initiative for the Evaluation of XML Retrieval (INEX) [5]. In a previous study, performed as part of the INEX interactive track [7], we evaluated focused access to scientific literature [6]. In the evaluation we used a home-grown XML retrieval interface developed in a student project [1]. The evaluation was also carried out in a student project.

In this paper we describe how we have adapted the XML retrieval interface to provide focused search access to Wikipedia [13]. We describe the system and its evaluation. The main goal of the experiment is to investigate the usefulness of focused access to information. More precisely, we explore the usefulness of giving users access the Wikipedia pages at individual section level, as opposed to page level only.

Our main findings are that users are positive toward focused information access. The main advantage of the section level access is that the users finish their search tasks in less time. Additionally, our experiment revealed that users access the Wikipedia pages equally via search result lists and via browsing within the encyclopedia itself.

The remainder of the paper is organized as follows. In Section 2 we survey Wikipedia and its use as a corpus for information retrieval experiments. In particular, we zoom in on how we use it in this paper. We introduce our Wikipedia search engine in Section 3. In Section 4 we describe the evaluation setup, and in Section 5 we present evaluation results. We conclude and describe future work in Section 6.

2. WIKIPEDIA AS A CORPUS

Wikipedia is *free* in the sense that its contents are written by web users and can be edited by any other web user. As a document collection, Wikipedia has many properties that make it attractive as a corpus for performing information retrieval experiments. For a start, the document collection is freely available, which makes it easy to distribute as part of a test collection. The multi-lingual aspects of the collection support a range multi-lingual retrieval efforts. Furthermore, the semi-structured format of the collection makes it usable for evaluation of semi-structured retrieval techniques, such as those developed for XML element retrieval. And last but not least, the dense link structure of the collection makes it interesting for investigating the interplay between searching and browsing when users seek information [3].

There are several information retrieval evaluation initiatives that plan to use Wikipedia as their corpus. At CLEF 2006 a pilot task will be running where Wikipedia is used as a corpus for question answering [17]. Within the INEX initiative there are ongoing efforts to convert the wiki markup language into a standard XML format, and use the corpus for the evaluation of ad-hoc XML retrieval [5]. In this study, we complement these initiatives by using Wikipedia as a corpus for an interactive experiment.

3. WIKIPEDIA SEARCH ENGINE

In this section we detail our Wikipedia search engine. We index and search Wikipedia using our XML retrieval system [10]. Although the content of the Wikipedia pages is

DIR 2006, March 13-14, 2006, Delft, The Netherlands.
.

not in XML format, it is semi-structured and can easily be interpreted as a hierarchy of text objects. In particular, the wiki syntax for nested section captions can be used to identify section boundaries and nesting levels.

3.1 Retrieval Engine

Our XML retrieval system is based on our home-grown extension of Lucene [8, 4]. The engine uses a simple multinomial language model to rank each indexing unit, in our case individual sections, with respect to relevance to the user's query. For now, no advanced XML specific retrieval methods are used. For example, we have found mixture models useful for ranking XML elements [11], but it remains as future work to make the implementation efficient enough for online usage.

3.2 Indexing Wikipedia

Since the content of Wikipedia pages is not marked up in XML, we created a simple parser for the Wikipedia syntax which allowed us to index the collection as if the pages were stored as XML. Our indexing units are either (sub)-sections (if present) or complete pages (in the absence of section structure). Our index is non-overlapping, where each text token is only indexed as part of its most deeply nested ancestor.

We also extracted and indexed two types of additional fields. Titles of pages and sections were indexed using the 'fields' mechanism of Lucene [8]. For each Wikipedia page we also extracted its categories and indexed as a separate field (of the page on which it occurs). These fields were not used in our current evaluation efforts.

We index the whole Wikipedia distribution package. This means that in addition to the "proper" encyclopedia pages we also index redirect pages and various log-pages. All included, we index 2,086,197 pages which are divided up into 4,095,103 indexing units.

3.3 Wikipedia Search Interface

We have created two interfaces to our Wikipedia search engine. One is a simple baseline interface which gives access to the start of Wikipedia pages only, while the other is a focused interface which gives access to individual sections of Wikipedia pages.

Our baseline search interface is a Google-like one where each result is presented as a pair: a link to the relevant page, and a short query dependent summary of the page in the form of a snippet. A screen-shot of the interface is shown in Figure 1.

Our focused Wikipedia search interface is based on our XML retrieval interface xmlfind [1, 6]. A screen-shot of the interface can be seen in Figure 2. We group the retrieved sections and subsections by the wiki page that they belong to. Hence, the main addition of the focused interface, compared to the baseline interface, is that the snippets are broken up by section boundaries, and hyper-links give access to individual sections.

The section-based linking and section-based snippets are the only difference between the two interfaces. They use the same underlying ranking scheme which means that documents are ranked precisely in the same order. The ranking of the documents is based on aggregated score of the relevant sections. The snippet used in the baseline system is created by concatenating the snippets of relevant sections. This means that both interfaces present precisely the same text to the user.

3.4 Logging User Interaction

Our system logs various interactions between the user and the system. This data can be used to better understand how users interact with our system.

- *Queries*: All queries posted by users are logged.
- *Visited Results*: The system stores information about which links on the result pages are clicked on by the user.
- *Site Navigation*: All internal navigation between Wikipedia pages is logged.

In Section 5 we describe how we use the collected data to answer the research questions that we will discuss in the next section.

4. EVALUATION

4.1 Research Questions

The goal of the experiments described in this paper is to gain a better understanding of the way in which users interact with a focused retrieval system. Our main research question, then, is:

> Do focused retrieval methods improve users' access to Wikipedia, compared with more traditional document retrieval methods?

Even before exploring this issue, we had reasons to believe that it would be difficult to come to a positive answer to this question if we look at the problem from a broad perspective. I.e., 'on average' the systems are likely to be rather similar. One of the main reasons for this 'pessimism' was that many of the Wikipedia entries are not very long. It is even the nature of Wikipedia as an encyclopedia to create a new separate entry for a 'sub-entry' that gets too long [15]. This means that for many entries there will be no, or little, difference between the two interfaces.

We believe, however, that there exist cases where focused access might be more useful. That is, if the information need is specific, and is satisfied by some text buried deep in a long entry. Hence, we reformulated our research question to a stronger question:

> Do there exist scenarios in which focused access improves users' access to Wikipedia, compared with a more traditional document access?

We believe that the search scenario plays a crucial role when the effect of this sort of focused retrieval is evaluated.

Although focused access is the main goal of this exercise, we will gather data of other interesting user behavior. One interesting aspect in the case of Wikipedia is the interplay between searching and browsing [3]. Wikipedia has very dense linking between entries. Since we keep an extensive log of user interaction with our system we are able to formulate a 'bonus' research question related to the link structure:

> What is the interplay between searching and browsing when users interact with densely hyper-linked sources such as Wikipedia?

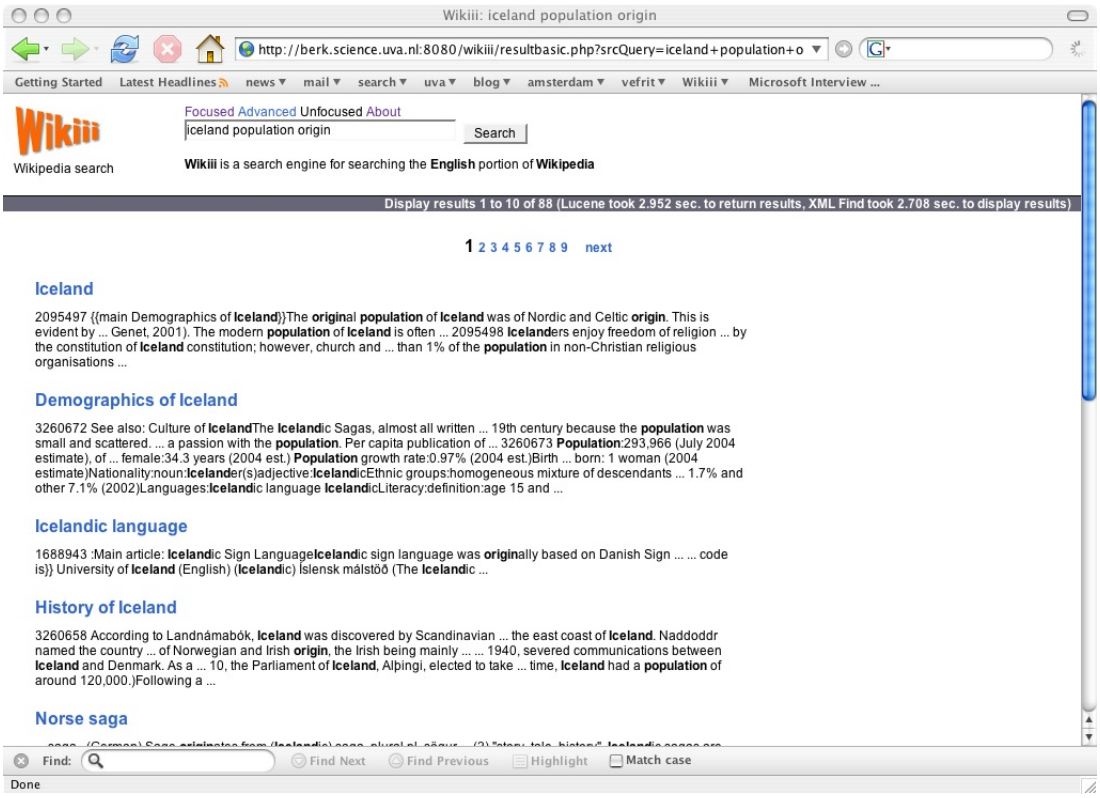

Figure 1: Baseline interface

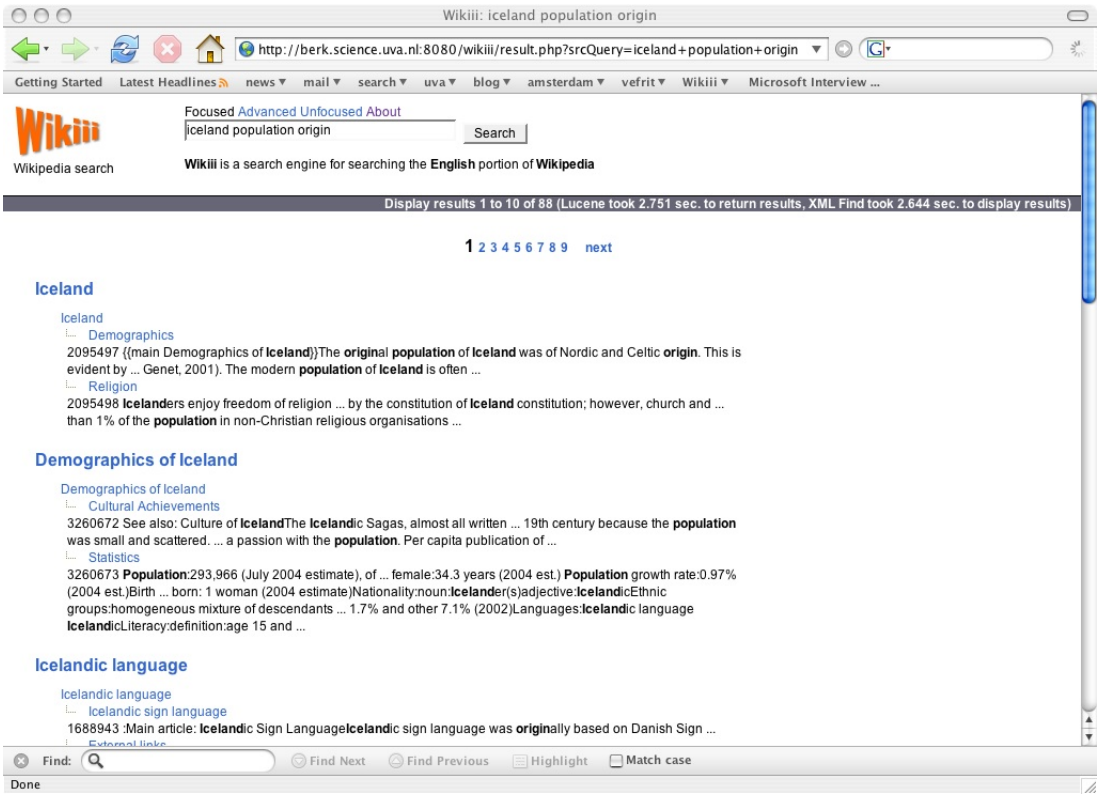

Figure 2: Focused interface

Motivation. Suppose you have just seen a report on the news about the recent earthquake in Pakistan. The report makes you want to get a better understanding of the Pakistan earthquake region.

Task. Please use the Wikipedia search engine to find the answer to the following questions:

- Where is Pakistan precisely?
- In which parts of Pakistan is there a great risk of earthquakes?
- What causes the earthquakes in Pakistan?
- Is there a difference between the cause of earthquakes in Pakistan, compared to other earthquake areas, such as California, Japan, or Iceland?

Figure 3: Example of a possible simulated work task.

Table 1: Experimental matrix for the interactive experiment.

Rotation	Task I	Task II
1	Baseline	Focused
2	Focused	Baseline

4.2 Experimental Setup

In order to answer our research questions we set up an interactive experiment where we asked people to perform simulated work tasks [2]. An example of a simulated work task can be seen in Figure 3. The actual work tasks that were used in the experiment can be found in Figures 7 and 8 in the appendix of this paper. Each of the actual work tasks consisted of three related search assignments. Each search assignment resembled a factoid question or a list question.

Each test subject performed two simulated work tasks, but using different system each time. The experiment matrix is shown in Table 1. Our analysis is based on 12 test persons, evenly distributed between the two rotations.[1]

The rotation removes the bias which is introduced by using one system before the other. The order of the simulated work tasks is always the same, leading to a potential interaction between the results for task I and task II.

In the beginning of the experiment the test person was asked to fill in a pre-experiment questionnaire on her background. After each task the user was asked to fill in a post-task questionnaire on her search experience during the task. Finally, the user was asked to fill in an post-experiment questionnaire after both task had been completed. The experiment, hence, involved the following steps:

1. Pre-experiment questionnaire
2. Simulated work task I
3. Post-task questionnaire
4. Simulated work task II
5. Post-task questionnaire
6. Post-experiment questionnaire

[1]In the original experiment there we 16 test cases, but from the system logs we found out that 4 of them did not fully follow the experiment guidelines.

Table 2: Responses on user *satisfaction*: Mean rating and standard deviation (in brackets). Answers were on a 5-point scale, ranging from 1 ("very dissatisfied") to 5 ("very satisfied").

	Task I		Task II		Overall	
Baseline	4.17	(0.75)	3.00	(1.26)	3.58	(1.16)
Focused	3.67	(1.41)	3.67	(0.52)	3.67	(0.65)

Table 3: Responses on user *effort*: Mean rating and standard deviation (in brackets). Answers were on a 5-point scale, ranging from 1 ("very difficult") to 5 ("very easy").

	Task I		Task II		Overall	
Baseline	3.17	(0.75)	2.83	(0.75)	3.00	(0.74)
Focused	2.67	(1.05)	3.50	(1.05)	3.08	(1.16)

5. RESULTS

We start by reporting on the user search experience while using our systems. These results are based on an analysis of the responses to the post-task questionnaires. We will then look at how users interacted with our system by mining the system interaction logs. Finally, we discuss the results in relation to the research questions stated in Section 4.

5.1 User Search Experience

In the post-task questionnaires there were two questions which addressed how the user experienced using the system for solving the task. One question asked about the user's satisfaction and the other about the user's effort.

Satisfaction: How satisfied are you with the answers given by this system?

The answers were given on a scale with range 1 to 5, where 1 stood for "very dissatisfied" and 5 for "very satisfied". The results for this question can be found in Table 2. The system satisfaction is mixed between the two tasks. Overall, there is little difference between the two systems.

Effort: The answers to the task-questions were in this system... [difficult/easy to find.]

The answers were given on a scale with range 1 to 5, where 1 stood for "very difficult to find" and 5 stood for "very easy to find". The results for this question are reported in Table 3. It is interesting to note that in solving the first task, the users rated the baseline system as easier to use. However, in solving the second task, the users rated the focused system as easier to use. Overall, there is very little difference between the two systems.

In the post-task questionnaire users were also asked how suitable they thought that the particular system was for answering respectively two types of questions, namely *specific questions* and *general questions*.

Specific questions: How well do you find this system suitable for specific questions?

Table 4: Responses on system suitability for answering *specific questions*: Mean rating and standard deviation (in brackets). Answers were on a 5-point scale, ranging from 1 ("very unsuitable") to 5 ("very suitable").

	Task I		Task II		Overall	
Baseline	2.50	(0.48)	2.50	(1.05)	2.50	(0.90)
Focused	3.00	(1.03)	3.17	(0.75)	3.08	(1.08)

Table 5: Responses on system suitability for answering *general questions*: Mean rating and standard deviation (in brackets). Answers we on a 5-point scale, ranging from 1 ("very unsuitable") to 5 ("very suitable")

	Task I		Task II		Overall	
Baseline	3.83	(0.75)	3.67	(1.03)	3.75	(0.87)
Focused	3.33	(1.21)	3.33	(1.03)	3.33	(0.98)

Table 4 shows how users rated the system's suitability for answering specific questions. The users find the focused system more suitable for specific tasks than the baseline system. Note, however, that the mean rating of the focused system is only slightly better than "neutral".

General questions: How well do you find this system suitable for general questions?

Table 5 shows how suitable the users rated the system's suitability for answering general questions. Now, both systems get a rating better than "neutral". The baseline system is rated above the focused system.

The notions of "specific questions" and "general questions" were not linked directly to the simulated work tasks performed, and may have been interpreted differently by each of the test persons. Still, the answers given do correspond to the expectation that focused search is particularly useful for specific information needs that could be answered with a relatively short amount of text [9].

5.2 User Interaction

We explore the user-system interaction by mining the interaction logs provided by the systems. Let us first look at the number of queries posted. Table 6 shows the mean number of queries issued in each search task. There is not much difference between users of the different systems. Next we look at the number of wiki pages viewed in each search task. Table 7 shows the mean number of pages viewed. This number includes all pages viewed, both via search results and via browsing within the Wikipedia site. Overall, users view more pages when using the focused system than when using the baseline system. The difference is not significant, however. If we look at the individual tasks, we see that we get different results. For Task I, more pages were visited by users of the baseline system. For Task II, the users of the focused system view more pages. In this case, the difference is significant (t-test: $p < 0.05$).

Table 7: Page views per search task: Mean number of page views and standard deviation (in brackets). Each search task was divided into three distinct search assignments.

	Task I		Task II		Overall	
Baseline	19.2	(12.4)	16.3	(4.6)	17.8	(9.0)
Focused	15.5	(8.1)	26.0	(8.0)	20.8	(9.4)

Table 8: Time spent per search task (minutes): mean time and standard deviation (in brackets). Each search task was divided into three distinct search assignments.

	Task I		Task II		Overall	
Baseline	31.2	(13.8)	27.0	(15.6)	29.1	(13.7)
Focused	23.3	(7.8)	22.5	(9.2)	22.9	(8.1)

Users of the focused system seem to spend more effort in terms of queries and page views, but what about time? Table 8 shows the average number of minutes needed to complete each search task. We see that despite all the page views, the users of the focused system finish their tasks quicker than the users of the baseline system. The difference is not significant, however.

Let us zoom in now on the interaction with the focused interface. Recall from Figure 2 that there are two types of links in the focused interface: *page-links* that bring you to the beginning of the page, and *focused-links* that take you to the relevant sections within a page. Let us look at whether users rather click on page-links or focused-links. Table 9 shows the average number of page-link and focused-link clicks for each search task. Overall, there is little difference between the popularity of the two access methods. If we look at each task separately, results are mixed. Users who used the focused system in in their first task preferred page-links over focused-links. Users who used the focused system for their second task had a slight preference for focused links. Figure 4 shows the ratio between page-link and focused-link clicks for each user. We see that the click-behavior is very user dependent.

Let us now take a closer look at the focused-links that were clicked. How deep into the documents do users dive? Table 10 shows both hierarchical and linear depth of user visits. The left part of the table shows where in the hierarchy the clicks are. No less than 70% of all clicks on focused links give access to sections or subsections, and the remaining 30% of the clicks are on the root element. The right part of the table shows a closer look at the section clicks.

Table 6: Queries per search task: Mean number of queries and standard deviation (in brackets). Each search task was divided into three distinct search assignments.

	Task I		Task II		Overall	
Baseline	11.33	(6.53)	8.67	(2.50)	10.00	(4.92)
Focused	12.50	(5.47)	9.50	(5.05)	11.00	(5.26)

Table 9: Page-link clicks vs. focused-link clicks in the focused interface: mean number of clicks and standard deviation (in brackets). Each search task contained three distinct search assignments.

	Task I		Task II		Overall	
Page-links	5.67	(5.85)	5.67	(4.59)	5.67	(5.02)
Focused-links	2.67	(1.03)	6.67	(4.63)	4.67	(3.82)

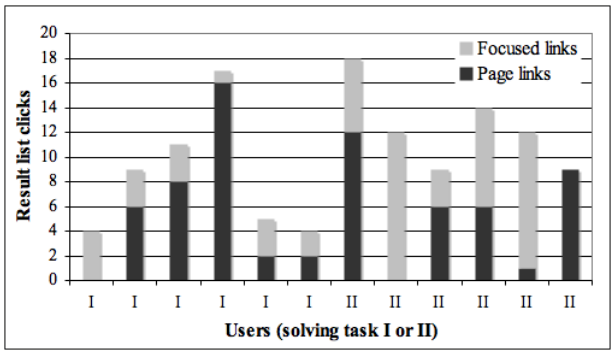

Figure 4: Number of result clicks per user in the focused interface. Dark: Clicks on page-links. Light: Clicks on focused-links.

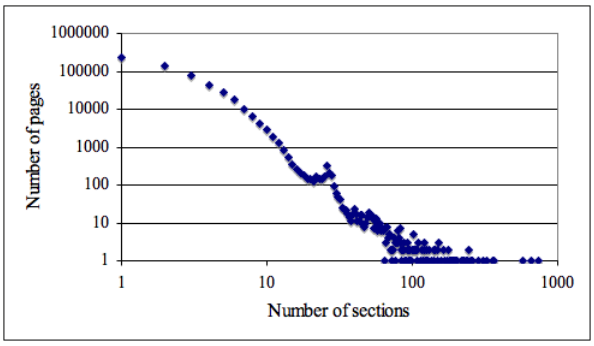

Figure 5: Linear depth of Wikipedia pages which have one or more sections. The distribution of pages over the number of sections is plotted on a log-log scale.

Table 10: Analysis of focused-clicks in the focused interface. *Left*: Type of element clicked (hierarchical depth). *Right*: Section number (in the Wikipedia source) of the of the sections clicked (linear depth).

Level	Clicks	
Root	17	30%
Section	31	55%
Subsection	8	15%

Section nr.	Clicks	
Section 1	16	52%
Section 2	5	16%
Section 3	5	16%
Section 4	4	13%
Section 9	1	3%

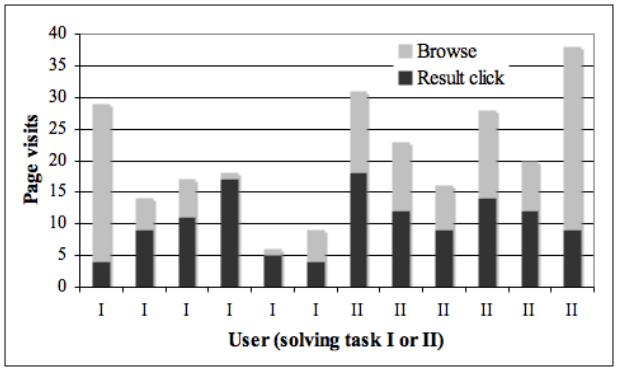

Figure 6: Number of page visits per user in the focused interface. Dark: Pages visited via the result list. Light: Pages visited via internal links.

Specifically, it shows how far into the document the section clicks go. About half of the links go to the first section of the Wikipedia article, while the other half goes deeper. This may seem a bit shallow access, but the collection itself is also rather shallow. About 560,000 pages are divided up into sections. Of these pages 224,000 have only one section, and 140,000 have two sections. Figure 5 shows the distribution of pages, based on the section count.

An important characteristic of Wikipedia is that the text is densely populated with hyper-links to other pages within the collection. Hence, it is important to see how users use these links as part of their information seeking behavior. In particular, it is of interest to see the ratio between pages visited via the search result list and pages visited via the internal link structure of Wikipedia. This ratio can be seen in Figure 6. Overall, 124 pages were reached via the search result list, while 125 were reached via internal links. The ratio is thus half-half. The ratio is slightly in favor of result visits for Task I and in favor of internal browsing for Task II.

5.3 Discussion

In the post-experiment questionnaire we asked the users which of the two systems they preferred. Most users chose the focused system. In their justification they argued that using the focused system the answers were found more quickly. They also complained that while using the baseline system too much text had to be read before the right answer was found. There were, however, several users that noted that there was little difference between the two systems.

Let's now recall our research questions as stated in Section 4. Our first research question read:

> Do focused retrieval methods improve users' access to Wikipedia, compared with more traditional document retrieval methods?

If time is an issue, the focused retrieval methods are promising. Users felt that they could find the right information quicker when using the focused system. This feeling is confirmed by the interaction log files.

It must be noted that most of the search assignments were about finding answers to factoid questions. That is, the assignments were aimed at satisfying "specific" information needs. Hence, our study provides evidence for the claim that focused retrieval methods are useful for "specific" information needs.

Next, let us look at our bonus question:

> What is the interplay between searching and browsing when users interact with densely hyper-linked sources such as Wikipedia?

In our experiment, page visits were evenly distributed between searching and browsing. The popularity of browsing was beyond our expectation. Earlier studies reported little interaction with the search results [12]. This issue deserves more attention. In future work, it might be interesting to go deeper into the role of browsing. Why do users browse?

Because they did not find the answer on the current page? Because they wanted to get broader support for their answer? Or even because they got distracted by an interesting hyper-link that was unrelated to their actual search assignment?

6. CONCLUSIONS AND FUTURE WORK

Wikipedia is an attractive corpus for performing information retrieval experiments. In this paper we described how it can be used to evaluate focused retrieval in an interactive experiment. One of our main findings is that focused access allows users to solve their search task quicker, at least when the information need is specific. Another main finding, derived as a by-product of our study, is that in a richly hyperlinked environment, users access pages equally via search result lists and via internal browsing. We believe that the interaction between searching and browsing deserves further study.

There are many options for extending the work in this paper. For the focused retrieval part, the outcome of the interactive experiment gives strong support to the effort to create a reusable system-oriented test collection based on Wikipedia. The first steps in this direction have already being taken within the INEX community. Focused information access to richly structured corpora also allows for retrieval using more expressive queries in which a user can combine content with structural constraints. With the creation of an XML version of Wikipedia this task becomes particularly interesting. Yet another form of focused information access is automatic question answering based on Wikipedia. Work on that task is already underway within the WiQA task at CLEF 2006. If we look beyond focused retrieval, Wikipedia is also a promising resource for evaluating multilingual retrieval, which will be (partly) addressed in the WiQA task.

Acknowledgments

Many thanks to the students of Project Information Retrieval who set up and carried out the interactive experiment: Monique Arlaud, Deborah Huijsman, Amelia Ibrahim, Natascha Ruimwijk, Deepak Sharma, Youri Sub-Laban, and Wendy van den Broek.

Jaap Kamps was supported by grants from the Netherlands Organization for Scientific Research (NWO) under project numbers 612.066.302 and 640.001.501. Maarten de Rijke was supported by grants from NWO under project numbers 017.001.190, 220-80-001, 264-70-050, 354-20-005, 612-13-001, 612.000.106, 612.000.207, 612.066.302, 612.069.-006, 640.001.501, and 640.002.501.

7. REFERENCES

[1] T. Bakker, M. Bedeker, S. van den Berg, P. van Blokland, J. de Lau, O. Kiszer, S. Reus, and J. Salomon. Evaluating XML retrieval interfaces: xmlfind. Technical report, University of Amsterdam, 2005.

[2] P. Borlund and P. Ingwersen. The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 53:225–250, 1997.

[3] Y. Chiaramella. Browsing and querying: Two complementary approaches for multimedia information retrieval. In *Hypertext – Information Retrieval – Multimedia (HIM'97)*, pages 9–26. Universitätsverlag Konstanz, 1997.

[4] ILPS. The ILPS extension of the Lucene search engine, 2006. http://ilps.science.uva.nl/Resources/.

[5] Initiative for the evaluation of XML retrieval (INEX), 2006. http://inex.is.informatik.uni-duisburg.de/.

[6] J. Kamps, M. de Rijke, and B. Sigurbjörnsson. What do users think of an XML element retrieval system? In *INEX 2005 Proceedings*, 2006. To appear.

[7] B. Larsen, S. Malik, and T. Tombros. The interactive track at INEX 2005. In *INEX 2005 Preproceedings*, pages 313–327, 2005.

[8] Lucene. Open-source search software, 2006. http://lucene.apache.org/.

[9] J. Reid, M. Lalmas, K. Finesilver, and M. Hertzum. Best entry points for structured document retrieval – part I: Characteristics. *Information Processing and Management*, 42:74–88, 2006.

[10] B. Sigurbjörnsson, J. Kamps, and M. de Rijke. An element-based approach to XML retrieval. In N. Fuhr, S. Malik, and M. Lalmas, editors, *INEX 2003 Workshop Proceedings*, pages 19–26, 2004.

[11] B. Sigurbjörnsson, J. Kamps, and M. de Rijke. Mixture models, overlap, and structural hints in XML element retrieval. In *Advances in XML Information Retrieval*, volume 3493 of *LNCS*, pages 196–210. Springer, 2005.

[12] A. Tombros, B. Larsen, and S. Malik. The interactive track at INEX 2004. In *Advances in XML Information Retrieval*, volume 3493 of *LNCS*, pages 410–423. Springer, 2005.

[13] Wikiii, 2006. http://berk.science.uva.nl:8080/wikiii.

[14] Wikipedia, the free encyclopedia, 2006. http://wikipedia.org/.

[15] Wikipedia:article_size, 2006. http://en.wikipedia.org/wiki/Article_size.

[16] Wikipedia:searching, 2006. http://en.wikipedia.org/wiki/Wikipedia:Searching.

[17] WiQA: Question answering using Wikipedia, 2006. http://ilps.science.uva.nl/WiQA/.

APPENDIX

The test persons in the Interactive experiment were all native Dutch speakers, and the simulated work tasks were formulated in Dutch. Figures 7 and 8 show the descriptions of the simulated work tasks I and II respectively.

Stel je bereidt je voor op het komende WK voetbal dat dit jaar in Duitsland wordt gehouden. Om in de juiste stemming te komen wil je wat meer weten over het volgende...

1. Wie heeft het eerste WK voetbal gewonnen en heeft dat land daarna ooit nog eens het kampioenschap gewonnen? Zo ja wanneer?

Stel je voor dat je naar een basketbalwedstrijd kijkt, die wordt gehouden tijdens de olympische spelen. Je vraagt je ineens af of basketbal altijd al een olympische sport is geweest. Dit blijkt wel het geval te zijn. Vervolgens stel je jezelf de volgende vraag...

2. Wie heeft de basketbal wedstrijd gewonnen tijdens de eerste olympische spelen?

Stel je voor dat je een vrouw bent en voetbal speelt. Je wilt wel eens weten wat er nou zo bijzonder is aan voetbal spelen op topniveau voor zowel mannen als vrouwen. Je stelt jezelf de volgende vraag...

3. Noem drie verschillen tussen het WK voetbal voor mannen en het WK voetbal voor vrouwen.

Figure 7: Task I: Simulated work task

Je probeert voor 't eerst mee te doen met de traditionele superbowl weddenschappen. Maar voor je je inzet kunt bepalen vraag je je af:

1. Welk football team heeft de eerste superbowl gewonnen, En heeft dit team daarna nog eens gewonnen? Zo ja, hoe vaak?

Je staat in de snowboard winkel, en vraagt je opeens af wanneer voor 't eerst snowboarden als olympische sport werd erkend... En je denkt:

2. Wie heeft de eerste olympische snowboard competitie gewonnen? [cat. Men's giant slalom]

Terwijl je op de bank zit te zappen, kom je bij eurosport opeens een sumo wedstrijd tegen. Waarop je je eigenlijk afvraagt hoe dat eigenlijk zit in de Verenigde Staten, bij football. Dus wil je weten:

3. Noem 3 verschillen tussen de Woman's professional football league [WPFL] en de [heren] football league [NFL].

Of

3. Noem 3 verschillen tussen [amateur, IFBB] body building competities tussen heren & dames.

Figure 8: Task II: Simulated work task

Like the tasks, the questionnaires were in Dutch. Below you can find the original Dutch version of the questions mentioned in Section 5.

Satisfaction: *In hoeverre heeft u in dit systeem een bevredigend antwoord gekregen op uw taakvragen?*

Effort: *De antwoorden op de taakvragen waren in dit systeem...* Here the answers vary from *erg makkelijk te vinden* to *erg moeilijk te vinden.*

Specific tasks: *In hoeverre is dit systeem volgens u geschikt voor specifieke taakvragen?*

General tasks: *In hoeverre is dit systeem volgens u geschikt voor algemene taakvragen?*