

Workshop Report

Seminar 09101 – Interactive Information Retrieval

Norbert Fuhr, Nicholas Belkin, Joemon M Jose, Cornelis J. van Rijsbergen

Abstract. There is need for more foundational research in the development of interactive information retrieval systems. The results of a week long discussion by a group of multi-disciplinary researchers have reported here. A brief description of main activities and major recommendations of the workshop are reported here.

1 Introduction

Interactive information retrieval (IIR) systems are a commodity nowadays; however, the scientific foundation of this type of system is rather limited. Information retrieval (IR) theory has widely ignored this area, and cognitive IR approaches have not yet led to detailed specifications for IIR systems. Within this context, we organized a week long seminar at Dagstuhl during March 2009 and the activities and recommendations are described in this report.

The general idea was to collect the state of the art in IIR research, and to define a research agenda for further work in this area. For this purpose, we brought together experts from the related areas such as information science, cognitive science, interactive IR, theoretical IR and human-computer interaction (HCI). We took a broad approach to the problem of IIR by highlighting latest results and naming crucial research issues. Based on these contributions, we identified open research problems and then point out steps towards resolving these problems.

2 Organization of the seminar

We had 32 participants from across the globe working on issues related to interactive information retrieval. The workshop started on a Monday and finished on Friday. The organization of the workshop included keynote talks (3), short talks from participants, demonstration sessions, special topic sessions, and break-out sessions. In addition, we had an afternoon visit to a nearby ancient city.

The technical activities focused around: evaluation methodology in information retrieval; adaptive and personalized retrieval; context and interfaces; and, semantic search. The three keynote talks were on: Cognitive & Context Modeling for Interactive IR; Evaluation of Interactive Retrieval systems; User Interfaces for Interactive Information Retrieval

Nick Belkin gave the first keynote talk with a historical overview of how cognitive (and other) models of users of interactive IR (IIR) systems have been elicited, constructed and used. Cognitive models in the domain of interactive information retrieval (IIR) are understood as models that a system (or a person) constructs of a(nother) person's "information need"; these are called "user models". Context models are models that a system (or a person) constructs of the conditions that led a(nother) person to engage in information seeking behavior, various characteristics of that person, and various aspects of that person's environment, broadly construed. Both types of models are used in personalization of IIR.

Belkin discussed the topic by reviewing Robert Taylor's 1968 article [Taylor 1968] in which he proposed four levels of "information need" or "query", and five "filters" according to which the librarian and the information seeker identify and clarify (i.e. model) various aspects of the user, the user's goals, the topic of interest and so on. He then presented various approaches to understanding of why people engage in information seeking behaviour, and of automatically constructing cognitive and contextual models. With respect to interactive developments, significant change points include the "cognitive turn" in the early to mid 1970s, modeling the human intermediary in the 1980s, and the "interactive turn" in the 1990s. He then outlined current state of cognitive and contextual modeling in IIR, which include cognitive modeling of "need", "intention" and recently incorporating inclusion of individual characteristics becoming more significant. In addition, contextual modeling of environmental factors being used and also contextual modeling of social factors becoming recognized as significant. There are also attempts to model long-term needs.

Maristella Agosti presented the second keynote and highlighted the need for modeling, organizing and managing scientific data produced in an evaluation campaign. In general, user studies and logs are used in a separate way, since they are adopted with different aims in mind. It seems more scientifically informative to combine logs together with observation in naturalistic settings. A systematic use of triangulation of different data collection techniques is needed as a general approach in order to get better knowledge of the Web information search process [Pharo & Järvelin 2004]. Taking inspiration from this general approach, a method of combining implicit and explicit user interaction data to gain information to be used for personalization purposes is outlined. The argument is that data log analysis can be combined with the results of data derived from user studies to evaluate information access services. Further, Agosti argued for using digital library systems as a tool to do this and presented a case study demonstrating this idea.

Harald Reiterer presented the third keynote talk giving an in-depth survey of interactive interfaces used for information access highlighting the lessons learned from these activities. Recent developments in interface technology are surveyed. Subsequently, the group discussed the role of interfaces in information seeking. Information is only useful when people interpret it in the context of their goals and activities. In order to design technologies that better support information work, it is necessary to better understand the details of user activity. In this context, the need for further studies on user information activity is needed.

Each of these keynotes was followed by short presentations from participants. Subsequently we formed shorter discussion groups, which are described briefly in the following.

3 Evaluation methodology

Starting from the Cranfield paradigm of evaluation methodology, we critically looked into the effect of searcher behaviour and the searcher's goals. The ultimate goal of information retrieval (IR) is to support humans to better access information in order to better carry out their tasks. Therefore IR research should provide methods and techniques to improve the retrieval/access process. In this regard Kalervo Järvelin led the debate and argued that IR evaluation methodology, in particular that based on Cranfield methodology, is not focusing its efforts properly to serve the user-centred goals. He argued further that the Cranfield paradigm of evaluation tends to lose its power as soon as one looks at human performance instead of system performance. Also, searchers using IR systems make use of rather unorthodox queries (from the Cranfield point-of-view) and sessions. Their search strategies have not been sufficiently described and cannot therefore be properly understood, supported or evaluated. Moreover, searchers engage in an information seeking process, which they have found effective enough based on their previous searching experiences. They try not to optimize the search result alone but the entire process (and effort) and its expected contribution to their primary task. Järvelin argued that this can be better understood in terms of the management science theory "incrementalism" than in terms of rationalism.

At the same time, the merits of current evaluation methodology in benchmarking various systems have been highlighted. We looked into the role of test collections in IR and emphasized their role as corner stones of IR evaluation. Sanderson led the debate and highlighted the papers that support test collection based evaluation [Sanderson 2009]. It is also argued that user experiments are slow to set up and expensive. Often they are not large enough to support any conclusive evidence in support or against the hypotheses. However, there are unsolved issues with respect to test collections. On the one side, there is strong evidence for relevance feedback and pseudo relevance feedback from test collection-based evaluations [Mitra et al. 1998]. Unfortunately such techniques have not been taken up for public utilization and we need to study why they haven't.

We also discussed simulated evaluation methodologies. One of the difficulties in interactive evaluation is the time needed to set-up experiments and the costs involved in terms of experimentation. These get exacerbated if we need to test multiple retrieval models. On the other hand classical IR evaluation methodology fails to consider interactive elements. An alternative is simulated evaluation in which the idea is to simulate all possible user interactions that might have happened in an actual usage of these systems. Using the ground truth given in test collections these strategies can be run and measurements can be taken. This allows one to benchmark various interactive retrieval models. However, this process will not consider the cognitive issues involved in user interaction. Hence it is important to conduct follow on user testing possibly with reduced number of interactive models. It is very clear that this methodology

needs further consideration in terms of simulation methods, measures etc. A possible alternative to this type of evaluation is to identify one or more specific types of users, limiting the user models to lead to prototypical interactive behaviors. The current difficulty with this approach is in developing credible and valid user models and associated behaviors.

4 Personalization, Adaptation & Context

Ann Blandford [Blandford 2009] pointed out that many of the current IIR systems are based on wrong assumptions about users and their behaviour. Thus, a better understanding of the human activity is needed when building new systems, aiming at covering the whole process of interacting with information: starting with the information need, followed by the information acquisition stage, then the found information is interpreted and finally used. Effective support of this process is only possible if the system takes the usage context into account and also allows for personalization and adaptation

We spent serious effort on discussing personalization research in the context of information retrieval. Adaptive IR may include adaptation of system features based on non-user factors, on the other hand personalization of IR is a subset of adaptive IR and is explicitly concerned with user-based factors. Personalization may be the more interesting, more difficult, and more fruitful approach

Belkin highlighted the facets of personalization: Relevance/usefulness/interest; Task; Problem state; Personal characteristics; Personal preferences; and Context/situation. There are lots of studies on investigating single facets in personalization however, not much study on integrating multiple facets and recommended further investigation is needed along these lines.

Context is an important factor in the information seeking process. There are many definitions of context and it is important to define this concept and highlight its role in information retrieval process. Context models are models that a system constructs of the conditions that led a user person to engage in information seeking behaviour. There are many facets of context and it can support understanding as well as retrieval. However, there are many technical challenges that need to be addressed. These include: what features of context can be used? How to elicit and represent those features? How to combine these features into a retrieval process? How to evaluate such a system?

5 Theoretical modeling of Interactive Information retrieval Systems

We also discussed the issues in modeling IIR systems. The classical Probability Ranking Principle (PRP) forms the theoretical basis for probabilistic Information Retrieval (IR) models, which have dominated IR theory for about 20 years. However, the assumptions underlying the PRP often do not hold, and its view is too narrow for interactive information retrieval (IIR).

Norbert Fuhr presented a new theoretical framework for interactive retrieval [Fuhr 2009]. The basic idea is that during IIR, a user moves between situations. In each situation, the system presents to the user a list of choices, about which s/he has to decide, and the first positive decision moves the user to a new situation. Each choice is associated with a number of cost and probability parameters. Based on these parameters, an optimum ordering of the choices can be derived - the PRP for IIR. Fuhr highlighted the relationship of this rule to the classical PRP and pointed out issues for further research. Massimo Melucci introduced a geometric model and its investigation for contextual information retrieval [Melucci 2009]. The geometric model leverages recent advances of vector space-based information retrieval.

The group observed that there is a lack of research activities in modeling of interactive IR systems and recommended this as one of the necessary action points.

6 Recommendations

In the closing session of the workshop, the group identified many research areas and highlighted the following recommendations:

1. Further effort is needed to define an evaluation methodology that can effectively evaluate context sensitive information retrieval systems. In this regard, the role of interactive test collections needs to be explored. In addition, the simulated evaluation methodology needs to be studied further.
2. Related to this the development of new evaluation measures, which evaluate system performance in terms of the entire information seeking interaction, rather than only in terms of the response to a single query.
3. There is an urgent need to define the concept of context and to study its exploitation in interactive information retrieval systems.
4. Theoretic models of interactive retrieval systems are very important. Serious efforts are needed to develop models that fits various interactive search scenarios
5. In order to reduce the effort for performing user-oriented evaluations, cooperation between research groups should be enforced and appropriate evaluation initiatives be launched.

References

1. Blandford, A. (2009). Interacting with information. In this volume.
2. Fuhr, N. (2009). A Probability Ranking Principle for Interactive IR. In this volume.
3. Melucci, M. (2009). Contextual Information Retrieval Using Vector Spaces. In this volume.
4. Mark Sanderson (2009). It's nice and warm in the cave. In this volume.
5. Mitra M, Singhal A, Buckley C. (1998). Improving Automatic Query Expansion. In: ACM SIGIR '98 Proceedings, pp 206-214
6. Taylor, R.S. (1968). Question-negotiation and information seeking in libraries. *College and Research Libraries*, 28:178-194.
7. Pharo, N. & Järvelin, K. (2004). The SST method: a tool for analyzing web information search process,. *Information Processing & Management*, 40: 633-654