Information Retrieval in Context – IRiX

Workshop at SIGIR 2004 - Sheffield

Peter Ingwersen, Department of Information Studies, Royal School of Library and Information Science, Copenhagen, Denmark;

Nick Belkin, School of Communication, Information and Library Science, Rutgers University, USA

Introduction.

There is a growing realisation that relevant information will be accessible increasingly across media and genres, across languages and across modalities. The retrieval of such information will depend on time, place, history of interaction, task in hand, and a range of other factors that are not given explicitly but are implicit in the interaction and ambient environment, namely the *context***.** IR research is now conducted in multi-media, multi-lingual, and multi-modal environments but largely out of context (Jarvelin & Ingwersen, 2004; Ingwersen & Jarvelin, forthcoming). However, such contextual data can be used effectively to constrain retrieval of information thereby reducing the complexity of the retrieval process. To achieve this, context models for different modalities will need to be developed so that they can be deployed effectively to enhance retrieval performance. Thus truly context-aware and -dependent retrieval will become feasible.

Information retrieval in context can mean one of two things. It can mean the design of methods and algorithms for IR that are context-aware, that is, able to elicit the implicit knowledge of the participatory searcher environment related to the context into an explicit representation. But it can also mean the specification of models for context and its features that can be used to focus an IR tasks according to such a context.

Commonly, context implies interactive IR and there may exist a stratification of contexts in association to IR engines and systems. For example, knowing where a searcher is focussing his or her attention during the interactive session of image retrieval can enhance the operation of relevance feedback to the system. The searcher's current task situation also acts as context as does his or her current information-seeking situation of which IIR forms part (Cool, 2001). However, traditionally IR research has dealt with contextual features of documents, e.g., other associated documents acting as context or document structures and portions functioning as context for their embedded NLP elements, such as, words nested in paragraphs. The underlying hypothesis (and belief) is that by taking account of context the next generation of retrieval engines dependent on models of context can be created, designed, developed and evaluated, delivering performance exceeding that of out-of-context engines.

Participation

The Organizing/Review Committee consisted of the following persons:

- Nick Belkin, Rutgers University, USA
- Pia Borlund, Royal School of LIS, Aalborg, Denmark
- Peter Ingwersen, Royal School of LIS, Copenhagen, Denmark

- Ian Ruthven, University of Strathclyde, Scotland
- Eero Sormunen, University of Tampere, Finland
- Amanda Spink, University of Pittsburgh, USA

Birger Larsen, Royal School of LIS, Denmark, and Ryen White, Glasgow University, organized the program and publication layout of presentations and the workshop website: http://ir.dcs.gla.ac.uk/context

The workshop attracted 15 position paper submissions, which were reduced via peer reviewing to 8 presentations with underlying papers, each maximum 2.500 words long, 2 demos and three panel presentations. We chose to allow for several presentations in order to obtain a diversified perspective on IR in context at this – the first SIGIR workshop on the topic. The structure of the program was such: first a panel presented three different approaches to IR in context: conceptual models of IRiX; context viewed from the perspective of IR systems design and evaluation; and the presentation of elements of a research framework for investigating IRiX.
This was followed by three sessions including discussion time: one session dealt with Models and General Approaches to IRiX; a second one presented IRiX Applications, including demos; and the third focussed on Mixing Contextual Evidence and Representations. The workshop ended with a wrapping up session and discussion of future tracks of IR in Context.

A total of 47 participants showed up at the workshop on July 29, whereby the IRiX workshop proved to be the largest one during SIGIR2004. Following the workshop a list of participants with e-mails has been distributed among the participants. Most of the presentations have also been made publicly available online at the workshop website, see above.

Workshop contents, recommendations and discussions

One might sum up the workshop into three major lines of action, preparing for follow-up events in coming years:

1. What are the elements of context, which are potentially significant to information retrieval?
    a. Work task and other task or interest features
    b. Searcher features
    c. System features
    d. Document features
    e. Environmental/physical features
    f. Temporal features
2. Which of these categories/elements are, or could be useful in improving information retrieval?
    a. This is dependent upon the situation
    b. This is also dependent upon the goal of the searcher
3. How can features of context be used to improve information retrieval?
    a. We need to develop methods of investigation, test beds, evaluation measures and techniques
    b. We need to explore potential operationalizations of aspects of context

Some efforts of capturing context features have been made in previous decades of interactive IR research, e.g., concerned with searcher goals and problem situations (Belkin, Oddy & Brooks, 1982). However, there is a lot of room for investigating

more intensively and specifically the nature of context associated with IR – and its usefulness to retrieval and seeking processes.
It was also evident to the participants that in future workshop events one should focus more on selected but central aspects of IRiX, for instance, contextual features and evidence that may support performance improvement and the suitable evaluation methodologies. It was suggested during workshop discussions that organizers should restrict the number of presentations in future, in order to make room for deeper discussion on such selected areas of research. This will be done at forthcoming events of IRiX.

Concluding remarks

We believe that everybody participating in the workshop agree that the topic is central to IR research and should be followed up at the ensuing SIGIR Conference in Brazil. It was especially interesting to note that a range of promising context-based IR models currently are under development and that demos of context-sensitive systems were feasible, mainly due to new (mobile) technologies, already at this primary event. Where the present workshop attempted to provide some sort of overview of issues and models of IRiX it seems now important to focus on specific facets of contextual IR. We hope to be allowed this possibility in 2005 in Salvador, Brazil.

References.

Belkin, N.J., Oddy, R.N. & Brooks, H. (1982). ASK for information retrieval. *Journal of Documentation, 38*(3): 145-164.
Cool, C. (2001) The concept of situation in information science. *Annual Review of Information Science and Technology*, v. 35: 5-42.
Ingwersen, P. & Jarvelin, K. (forthcoming). *The Turn: Integration of Information Seeking and Retrieval in Context.* Research monograph. Springer/Kluwer. 500 p.
Järvelin, K. & Ingwersen, P. (2004). "Information seeking research needs extension toward tasks and technology". Information Research, 10(1) paper 212 [Available at http://InformationR.net/ir/10-1/paper212.html]