Extracting Information Seeking Intentions for Web Search Sessions

Matthew Mitsui
Department of Computer Science
Rutgers University
New Brunswick, NJ, USA, 08901
mmitsui@cs.rutgers.edu

Chirag Shah, Nicholas J. Belkin
School of Communication and Information
Rutgers University
New Brunswick, NJ, USA, 08901
{chirags,belkin}@rutgers.edu

ABSTRACT

We present a method for extracting the self-reported intentions of users engaged in an information seeking episode. We recruited participants to conduct search sessions and subsequently asked them to self-report their intentions. A total of 27 users participated in a lab study, during which they worked on two search tasks. After each search session, participants indicated their intentions during that session while viewing a video replay. Results indicate that the set of search intentions provided to participants was sufficient to account for intentions in four journalism-related information seeking tasks: a copy editing task, interview preparation task, relationships task, and story pitch task. The results also suggest regular patterns in intentions that can be exploited for identification of task type as well as potential applications to personalization and recommendation during a search episode.

CCS Concepts

•Information systems → Query intent; Task models;

Keywords

search intentions; information seeking intentions; motivating task; information seeking episode; search session analysis

1. INTRODUCTION

It is becoming increasingly recognized that in general, information retrieval (IR) system performance cannot be evaluated in single query-single response format. Contrary to the Cranfield paradigm that has been presupposed in decades of IR system evaluation, performance must be measured over information seeking episodes - entire sessions of related queries, search results, clicks, and other interactions with IR systems. Moreover, these episodes are understood to be motivated by an external goal.

Therefore, it seems IR systems should directly support the accomplishment of user goals, whether it is the motivating

task/goal for the entire session or the various goals throughout the course of the session. We must hence be able to extract these goals - or an indicator of the goals - from patterns in interaction data, to gain a better understanding of the relationship between goals and behavioral patterns. Recent work has attempted to articulate and define searchers' *intentions* at various parts of a session [11], though little work has been done to extract such intentions for a fixed set of tasks. In this paper, we present a novel framework for extracting information seeking intentions from session data. The work reported here represents preliminary results of a project which aims to predict search intentions during a search session, on the basis of observable searcher behaviors.

2. BACKGROUND

Early proposals of why people engage in information seeking addressed the fundamental "why", discussing the motivating task to some extent (e.g. [1]; [10]). Later work has followed to characterize these motivating tasks. Much of this work focused on categorizing tasks into broad categories. Byström and Järvelin, for instance, characterize tasks by complexity [2]. Freund, Toms & Clarke relate task type to document genre and differences in usefulness and relevance [5]. Li classified motivating and search tasks into various facets, such as the product of the task, its complexity, and its degree of specificity [7].

Other work segmented information seeking into various levels of actions. Marchionini, for instance, differentiated between patterns, strategies, tactics, and moves, which are hierarchically related in increasing order of specificity [8]. In a later hierarchy proposed by Xie, the motivating search task eventually leads to more concrete "interactive search intentions", actual behaviors that are executed as steps in the search session [11]. Xie defined a concrete list of these intentions, identified through coding of empirical data, and found consistent patterns between the intentions and the information seeking strategies that were applied for these intentions. Literature exists in extracting intentions in a different sense, such as topical intent [6]. Teevan, Dumais & Horvitz also allowed users to give free form answers for intentions for search sessions as a whole, and related that to whether a query was worthwhile for personalization [9]. To our knowledge, though, little of the previous literature except Xie gives a fixed list of intentions that can serve as concrete, general steps in a query segment. In the work we present here, we choose a subset of the intentions of Xie to study. Rather than asking independent assessors to annotate query segments, we ask lab participants to annotate

their own segments. In additional contrast to Xie, rather than studying library users who bring their own library-related task, we study lab participants with a controlled set of topics and tasks, for the purpose of understanding the relationships between task, topic, and intentions.

Table 1: Search Tasks (Descriptions are on the coelacanths topic. Equivalent tasks were set for the topic of methane clathrates and global warming).

Assignment 1. Copy Editing
Your Assignment: You are a copy editor at a newspaper and you have only 20 minutes to check the accuracy of the six italicized statements in the excerpt of a piece of news story below.
Your Task: Please find and save an authoritative page that either confirms or disconfirms each statement.

Assignment 2. Story Pitch
Your Assignment: You are planning to pitch a science story to your editor and need to identify interesting facts about the coelacanth ("see-la-kanth"), a fish that dates from the time of dinosaurs and was thought to be extinct.
Your Task: Find and save web pages that contain the six most interesting facts about coelacanths and/or research about coelacanths and their preservation.

Assignment 3. Relationships
Your Assignment: You are writing an article about coelacanths and conservation efforts. You have found an interesting article about coelacanths but in order to develop your article you need to be able to explain the relationship between key facts you have learned.
Your Task: In the following there are five italicized passages, find an authoritative web page that explains the relationship between two of the italicized facts.

Assignment 4. Interview Preparation
Your Assignment: You are writing an article that profiles a scientist and their research work. You are preparing to interview Mark Erdmann, a marine biologist, about coelacanths and conservation programs.
Your Task: Identify and save authoritative web pages for the following: Identify two (living) people who likely can provide some personal stories about Dr. Erdmann and his work. Find the three most interesting facts about Dr. Erdmann's research. Find an interesting potential impact of Dr. Erdmann's work.

---

3. TASK AND DATASET

Our user data was collected in a lab setting. Participants were undergraduate students from one university, recruited from undergraduate journalism courses. To register, students were required to have completed at least one course in news writing. Each study session consisted of 2 search tasks, each followed by an annotation task, and several interspersed questionnaires, with a verbal exit interview at the end. All activity except for the exit interview was conducted at a desktop computer, with search activity recorded in Firefox by a browser plugin, eye-fixation behavior by GazePoint[1], and annotatable video of the search by Morae[2].

Participants began by answering a demographic questionnaire and watching a tutorial video on how to use our system

before beginning the search task. Then, participants read the task description and answered a short questionnaire on their familiarity with the topic and task as well as the anticipated difficulty. They then had 20 minutes to complete the search task; this was shown to be a sufficient amount of time in pilot tests, and the time limit needed to be constant among task types. They could finish before 20 minutes if they felt they completed their task early. Afterwards, participants answered a post questionnaire on the actual difficulty of the task. They then read a handout of the intention annotation task and watched a video demonstrating how to conduct the task. Users were also given a handout of a short description of each intention (see the Appendix for intention definitions). This was for further clarification and to also reduce variability in our data from differing interpretations of the intentions. They then completed the intention annotation task with no time limit. They then repeated the process with more questionnaires, another search task, and then another intention annotation task before the exit interview. The entire experimental session lasted about two hours.

For the intention annotation task, participants were asked to select which intentions applied to each *query segment* (all that occurred from one query to the next) in the search session. This was accomplished by playing the video of the search, segment by segment. They could select, from a displayed list, any number of intentions for a segment. For instance, if a participant knew nothing about coelacanths and issued the query "coelacanths" as the first query in a session, that person might mark "identify something to get started" and "learn domain knowledge". The participant was then asked to mark whether each of these checked intentions was satisfied. Our example participant may mark "yes" for "identify something to get started" but "no" for "learn domain knowledge". If a participant marks "no", she must then state why that intention was not satisfied. For example, while she found some new keywords to search, she may not have learned any knowledge that was required by the task description. If the participant had some other intention in addition to the 20 we listed, the participant may also check "other", give a short description of that additional intention, and also mark whether it was satisfied. They repeated this annotation process for each query segment separately. For the entire process, participants were incentivized with additional reward for being among the best performers. This provided incentive to issue good searches, instead of meeting the minimum requirements. Participants were told that "good performance" also included marking intentions well - i.e. marking all and only those that applied.

Table 2: Task characteristics (F=Factual, I=Intellectual, S=Specific, A=Amorphous).

Task	Product	Level	Goal	Named
CPE	F	Segment	S	True
STP	F	Segment	A	False
REL	I	Document	A	True
INT	I	Document	A	False

There were 4 possible tasks and 2 topics per task. Two of our task types are a copy editing task (CPE) and interview preparation task (INT), as specified in Cole et al [3]. Our other tasks - Relationships (REL) and Story Pitch (STP) task - were novel to this study. One topic was "coelacanths", and another was "methane clathrates and global warming".

We give descriptions of each task for the coelacanth topic in Table 1. The chosen topics were familiar enough to generate participant interest yet unfamiliar enough so participants would likely not know the requested information before arrival. We further give a faceted classification of each task in Table 2, according to a subset of facets in Li [7]. For definitions of each facet, see [3, 7]. Each user completed 2 search tasks and hence 2 annotation tasks. Task types were paired into 4 groups, based on differences in facet values. Each participant searched for 2 tasks in one of these 4 groups, each task on a different topic. Order of the 2 tasks and 2 topics in each group was flipped, yielding a total possible $4 \times 2 \times 2 = 16$ configurations. In our current dataset, we have 27 users. For analyzing intentions, we filtered out 3 users for misinterpreting instructions; they marked every intention for every query, thus skewing our data. This results in 24 searchers, conducting a total of 48 search sessions, and having assigned a total of 434 sets of intentions (i.e. 434 queries).

4. ANALYSIS

In defining our list of intentions, we took a subset of those from Xie [11], eliminating those that did not apply to our situation. While this may raise concerns that we may have missed some possibly valuable intentions, we found that the intentions were both necessary and sufficient. Intentions that were rarely selected in some task types were more frequently chosen in others. Good coverage was also demonstrated in our exit interviews. We asked participants if there were some missing intentions, to which no users gave a possible intention. We also found that users listed an "Other" intention in the task 19 out of 434 times. None of the "other" intentions that participants provided were repeated; they ranged from "I just wanted to see what the fish looked like" to "I found that Dr. Gerald Allen worked with Dr. Erdmann so I looked for Dr. Allen." With more data, we will attempt to categorize "other" intentions. Their answers suggest that while other intentions may be possible, this list of intentions gives a very good coverage of the possible intentions for these task types.

In Figure 1, we present the total counts of all selected intentions across all task types - i.e. the number of intentions that users marked as present in their query segments. We also show the number of users and queries per task. Because the number of search sessions is small, we provide only descriptive - and not inferential - statistics, but consistent patterns can still be drawn from the data. Two intentions consistently ranked among the 4 most common intentions in the query segments: "find specific information" and "obtain specific information". We can say this is due to the episodic nature of query segments. Since users break up information needs into succinct queries, each query almost always involves finding or obtaining specific information. More important is the obvious difference in relative occurrence of intentions in different task types. For instance, "evaluate correctness" is the third most frequently checked intention in CPE by a large margin, a pattern not exhibited by any other task type. "Identify something more to search" was common overall but also the most frequent intention in REL. In REL, users are more likely to "access items with common characteristics" (AC). Other differences suggest that some intentions can serve as distinguishing features for task facets. AC is much more common in Document tasks than

Segment tasks, while "evaluate duplication of an item" is lowest in the Document tasks. Lastly, the relative ordering of most to least frequent intention differs greatly between task types, suggesting that such differences may be useful for distinguishing task type.

5. CONCLUSION AND FUTURE WORK

Despite the relatively small data set, the results suggest the potential strength of our approach in identifying searchers' intentions during information seeking episodes. Our findings suggest that, at least in the four types of task we have studied, our chosen intentions provide sufficient coverage for characterizing search intentions. They also suggest that there may be regular patterns of intentions throughout a search session that differ from task to task. Similarly, Cole, et al. [4] identified a small number of clusters of sequences of eye-fixation behaviors, whose frequency of occurrence differs between tasks similar to those in our research. We therefore are continuing to collect search session data using these tasks and methods, in preparation for the next step in this line of research, which is to attempt to discover relationships between logged behaviors during search, and corresponding search intentions. The ultimate goal of the research of which this milestone is a part, is to be able to identify different search intentions during the course of an information seeking episode, in order to provide support specific to each different type of intention throughout.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] N. J. Belkin. Anomalous states of knowledge as a basis for information-retrieval. *Canadian Journal of Information Science-Revue Canadienne Des Sciences De L Information*, 5:133–143, 1980.

[2] K. Byström and K. Järvelin. Task complexity affects information seeking and use. *Inf. Process. Manage.*, 31(2):191–213, Mar. 1995.

[3] M. J. Cole, J. Gwizdka, C. Liu, R. Bierig, N. J. Belkin, and X. Zhang. Task and user effects on reading patterns in information search. *Interacting with Computers*, 23(4):346 – 362, 2011. Cognitive Ergonomics for Situated Human-Automation Collaboration.

[4] M. J. Cole, C. Hendahewa, N. J. Belkin, and C. Shah. User activity patterns during information search. *ACM Trans. Inf. Syst.*, 33(1):1:1–1:39, Mar. 2015.

[5] L. Freund, E. G. Toms, and C. L. Clarke. Modeling task-genre relationships for ir in the workplace. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '05, pages 441–448, New York, NY, USA, 2005. ACM.

[6] D. Kelly. *New Directions in Cognitive Information Retrieval*, chapter Implicit Feedback: Using Behavior to Infer Relevance, pages 169–186. Springer Netherlands, Dordrecht, 2005.

[7] Y. Li. *Relationships among work tasks, search tasks, and interactive information searching behavior*. ProQuest, 2008.

[8] G. Marchionini. *Information seeking in electronic environments*. Number 9. Cambridge university press, 1997.

[9] J. Teevan, S. T. Dumais, and E. Horvitz. Potential for personalization. *ACM Trans. Comput.-Hum. Interact.*, 17(1):4:1–4:31, Apr. 2010.
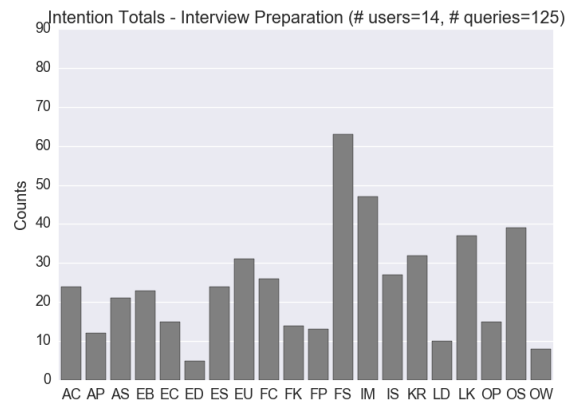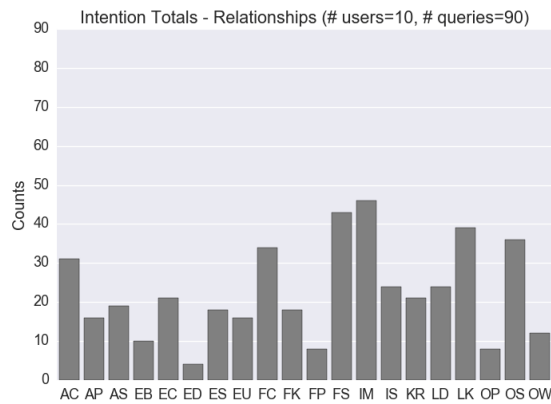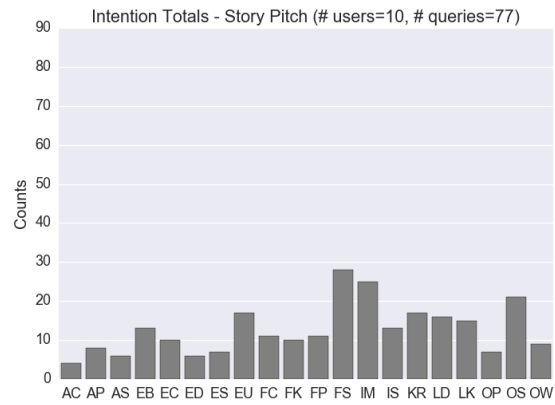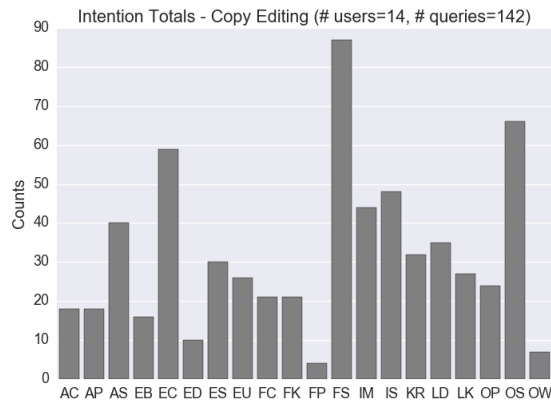
Figure 1: Total number of checked intentions for each task type.

[10] T. D. Wilson. On user studies and information needs. *Journal of documentation*, 31(1):3–15, 1981.

[11] H. I. Xie. Patterns between interactive intentions and information-seeking strategies. *Information Processing & Management*, 38(1):55 − 77, 2002.

APPENDIX

A. INTENTION DESCRIPTIONS

At walk-in, users were given two handouts: a description of the intention annotation task and a short bulleted list of each possible intention and a clarifying description. Here is the bulleted list they were given for the 20 intentions. Intentions were listed in this order on the handout, grouped by first word (e.g., "Identifying" and "finding").

- **Identify something to get started (IS)** - For instance, find good query terms.

- **Identify something more to search (IM)** - Explore a topic more broadly.

- **Learn domain knowledge (LK)** - Learn about the topic of a search.

- **Learn database content (LD)** - Learn the type of information/resources available at a particular website - e.g., a government database.

- **Find a known item (FK)** - Searching for an item that you were familiar with in advance.

- **Find specific information (FS)** - Finding a predetermined piece of information.

- **Find items sharing a named characteristic (FC)** - Finding items with something in common.

- **Find items without predefined criteria (FP)** - Finding items that will be useful for a task, but which haven't been specified in advance.

- **Keep record of a link (KR)** - Saving a good item or an item to look at later

- **Access a specific item (AS)** - Go to some item that you already know about.

- **Access items with common characteristics (AC)** - Go to some set of items with common characteristics.

- **Access a web site/home page or similar (AP)** - Relocating or going to a website.

- **Evaluate correctness of an item (EC)** - Determine whether an item is factually correct.

- **Evaluate usefulness of an item (EU)** - Determine whether an item is useful.

- **Pick best item(s) from all the useful ones (EB)** - Determine the best item among a set of items.

- **Evaluate specificity of an item (ES)** - Determine whether an item is specific or general enough.

- **Evaluate duplication of an item (ED)** - Determine whether the information in one item is the same as in another or others.

- **Obtain specific information (OS)** - Finding specific information to bookmark, highlight, or copy.

- **Obtain part of the item (OP)** - Finding part of an item to bookmark, highlight, or copy.

- **Obtain a whole item(s) (OW)** - Finding a whole item to bookmark, highlight, or copy.