Understanding the Brevity of Web Queries

Why do people submit brief, non-discriminating queries to web search engines?

Introduction

Internet search engines are the most common resource for people seeking answers to questions and access to web-based information, and in most search engines, keyword queries are the primary means of retrieving information. Based on transaction log analyses of web queries, we know that web queries are on average about two words in length, and have grown only slightly over time. This is much shorter than queries used in traditional information retrieval systems, and provides very limited input to support information retrieval. This study seeks to better understand why people use such brief and general queries when using search engines.

- What factors motivate the formulation of brief queries?
- How are brief queries being used in the web search process?
- What obstacles prevent the use of longer, more descriptive queries?

Methods

-FREC10 Interactive Track user study of web searching
  - 48 non-expert participants searched for 4 assigned tasks using a slightly modified Google interface.
  - Tasks were from four domains: Research, Consumer Health, Travel and Shopping.
  - Half the tasks were left partially open so that participants could personalize them.
  - We asked participants to use Keyword queries for 2 tasks, and Full Sentence queries for 2 tasks.
  - We used a transaction log to record queries and other aspects of search behaviour.
  - A semi-structured interview was conducted while re-playing the task using screen capture software.
  - The interview included questions on how and why queries were formulated and reformulated.

Results

I. Length

Keyword queries were on average 2.7 terms in length; those entered as sentences were 5.7 terms long. Of the 297 queries collected in the study, over 30% were 1 or 2 terms in length.

II. Reformulation

46% of searches were completed using a single query. Queries reformulated over the course of a search were more often lengthened (one or more terms added) than shortened. The most common strategy was to change one or more of the query terms in successive reformulations.

III. Use Context

Brief queries often occur in situations where the underlying information need is quite general, but also when it is well-defined and specific.

In some cases, users are constrained by preconceptions of how search engines work, by lack of knowledge of the available material on a topic, and by their own abilities (see right).

Short queries are used as opening gambits to provide easy entry into the webspace.

Establish a Universe

Some participants did not want or need a highly focused search.

These searches used general queries purposely as a coarse filter, to create a broader environment.

Some expressed this in terms of "contraining control" over the search.

I just prefer to find my own way around on the net, so I just start in a bigger range, and then I choose from there. [P07-34]

Well because actually... I want to decide, I want to sift through everything. Even though it's overwhelming [P00-22]

Open a Gateway

Brief queries are used as opening gambits to provide easy entry into the webspaces.

This allows the searcher to get to the heart of the "just sort of a general start to see what it would pull up with the two sort of areas I was interested in." [P14-13]

Refining the search by adding more terms to the query was a secondary tactic – only done if the initial results were unsatisfactory.

deciding what to query is not the challenging part for me. Figuring out how to refine the search is sometimes... Starting isn't the problem [P08-31]

if "second hand smoking" did not work, I would probably go back and do another search again but with either synonyms or added words, to make a more descriptive search [P13-12]

yeah, I'd go to the search box. If I was going to Lonely Planet and if I didn't know the where-its-are, the address, URL address or whatever it's called, I'd just go to the search box. [P80-31]

when I approach certain topics, I choose to go to a specific website and do a search there, if I think it's going to be more efficient... [P00-24]

Target a Site

Brief queries, often proper nouns, were submitted when participants were seeking websites of known organizations or other information portal sites.

"Futureshop", "Lonely Planet" and "British Museum" are examples of queries seeking organizational websites. "Titanic" and "global warming" are examples of attempts to reach information portal sites on these topics.

Participants preferred to search for detailed parameters once they reached the target website.

Yeah, I'd go to the search box. If I was going to Lonely Planet and if I didn't know the where-its-are, the address, URL address or whatever it's called, I'd just go to the search box. [P80-31]

when I approach certain topics, I choose to go to a specific website and do a search there, if I think it's going to be more efficient... [P00-24]

Participants seemed to have a strong sense that search engine queries should be brief.

Many participants expressed discomfort at being asked to submit queries in sentence form: "I don’t put sentences in these things [P10-21]

Some participants seemed to view query formulation as a subject classification problem, and restricted their queries to subject descriptors.

you just get into the point, you don't. I can’t imagine putting a sentence in there. [P80-42]

what is the base? What is the base topic?" [P92-41]

Conclusions

This study is part of an ongoing project to better understand the factors at play when people formulate web queries, in order to provide better support for that process.

Discussion and Future Work

This study provides further evidence that querying is a complex and interactive process.

Web-queries seem to play different roles in searching than traditional information retrieval queries, which is one reason for their brevity.

User perceptions of search engines and the Internet also play a role in shaping queries, which suggests a means of encouraging the use of longer queries, where this may improve search results.

This study is part of an ongoing project to better understand the factors at play when people formulate web queries, in order to provide better support for that process.

Acknowledgements

This work was partially funded by an IBM Centre for Advanced Studies Student Fellowship to the first author and a Natural Sciences and Engineering Research Council of Canada grant to the second author.

Keywords: search engine, query, reformulation, brevity, user studies, interview.