

HCSNet Workshop on Next Generation Search

Falk Scholer
RMIT University

Issue 1: Query Performance Prediction



The aim: estimate the *effectiveness* of a query in advance.

If it's likely that a query will return a poor set of answers, the search system can respond dynamically.

- Try a different retrieval model.
- Automatically reformulate query (query expansion).
- Ask user to reformulate query.

acsc - Google Search - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.google.com/search?hl=en&lr=&q=acsc&btnG=Search

Go

Google Web Images Video ^{New!} News Maps more »

acsc Search Advanced Search Preferences

Web Results 1 - 10 of about 501,000 for **acsc**. (0.28 seconds)

[Air Command and Staff College \(ACSC\)](#)
Overview, new student information, FAQ, and program details.
[www.acsc.af.mil/](#) - 10k - 18 Sep 2006 - [Cached](#) - [Similar pages](#)
[Downloads](#) - [www.acsc.af.mil/downloads.asp](#)
[Air Force Link](#) - [www.af.mil/](#)
[Welcome](#) - [www.au.af.mil/au/](#)
[Schools](#) - [www.maxwell.af.mil/au/schools.asp](#)
[More results from www.acsc.af.mil »](#)

[Australian Defence College](#)
[www.defence.gov.au/acsc/](#) - 2k - [Cached](#) - [Similar pages](#)

[Anderson Community School Corp.](#)
School information including employment, calendar, curriculum, staff and contacts.
[www.acsc.net/](#) - 2k - [Cached](#) - [Similar pages](#)

See results for: **[automobile club of southern california](#)**

[Automobile Club of Southern California](#)
The Auto **Club**: insurance quotes, travel planning and booking, auto purchasing and maintenance, and financial services for AAA members in **Southern ...**
[www.aaa-calif.com/](#)

[Javascript Enablement Required - AAA](#)
[www.aaa-calif.com/index.asp](#)

[AAA - American Automobile Association](#)
The American **Automobile** Association can help you obtain the best, most comprehensive coverage available ... Enter your zip Code to find your Local AAA **Club**: ...
[www.aaa.com/](#)

[ACSC2006 Workshop](#)
29th **Australasian Computer Science Conference-2006 (ACSC2006)** The University of Tasmania, Hobart, TAS, Australia January 16 - 19, 2006 ...
[www.cit.gu.edu.au/conferences/ACSC2006/](#) - 3k - [Cached](#) - [Similar pages](#)

Done

Evaluating Predictors



What makes a predictor good?

- Ideal: it can predict the utility of the result set for the user.
- Default approach: remove user from process, and use some standard IR metric as a proxy.
 - ▶ For example, calculate the Spearman rank correlation between the predicted difficulty for each query and the system's Average Precision for each query.



Current Approaches



Pre-retrieval: Use statistical information about the distribution of terms within documents and the collection

- IDF, query length, variance

Post-retrieval: Incorporate information from an initial result set obtained from running the query

- Query clarity, answer set cohesiveness

Learning predictors: use machine learning approaches to train on labelled data, then predict performance for new queries

- Neural networks, Support Vector Machines
-



Alternative Sources of Evidence



Natural language and semantic features:

- Synonymy
- Polysemy
- Use of proper nouns, pronouns
- Function words

User search behaviour:

- Clickthrough
- Past queries
- Other system interaction

Issue 2: Information Retrieval Evaluation



IR has a well-established tradition of experimental evaluation, involving:

- A set of queries.
- A static collection to search.
- Relevance judgements.

Many different metrics can be calculated.

- Aim to measure different aspects of retrieval.
- Give an objective view, allow comparisons to be made.

What About Users?



However:

- There is no correlation between the most widely-used IR performance metrics and actual user search performance.
- Users are abstracted out of the evaluation process.
- Some aspects of the retrieval process can't be evaluated at all using traditional approaches (e.g. document summarisation).

New metrics that reflect users and their search behaviour should therefore be developed.

Evaluation Assumptions



The underlying assumptions of current IR evaluation need to be re-visited.

- Direct observation and analysis of information-seeking behaviour in an online search environment
 - ▶ What do users actually do? How do they interact with the system?
- Identifying the different types of search tasks that users actually engage in
 - ▶ Informational, Navigational, Transactional... is that really how users plan their searches?

Evaluation Assumptions...



- Exploring how users learn while conducting searches
 - ▶ Different types of documents (structured, unstructured) and different media
 - ▶ Trust issues
- Accounting for factors such as interface design
 - ▶ How would users *like* to interact with a search system?

Once we understand these things, how can they be modelled, and turned into practical metrics for the evaluation of IR systems?

