Hierarchical Approach to Term Suggestion Device

Hideo Joho
Department of Information Studies
University of Sheffield
Western Bank, Sheffield, S10 2TN
+44 (0)114 222 2675
h.joho@sheffield.ac.uk

Mark Sanderson
Department of Information Studies
University of Sheffield
Western Bank, Sheffield, S10 2TN
+44 (0)114 222 2648
m.sanderson@sheffield.ac.uk

Micheline Beaulieu
Department of Information Studies
University of Sheffield
Western Bank, Sheffield, S10 2TN
+44 (0)114 222 2640
m.beaulieu@sheffield.ac.uk

ABSTRACT

Our demonstration shows the hierarchy system working on a locally run search engine. Hierarchies are dynamically generated from the retrieved documents, and visualised on the menus. When a user selects a term from the hierarchy, the documents linked to the term are listed, and the term is then added to the initial query to rerun a search. Through the demonstration we illustrate how hierarchical presentation of expansion terms is achieved, and how our approach supports users to articulate their information needs using the hierarchy.

Keywords

Concept hierarchies, interactive query expansion.

1. DESCRIPTION

Traditionally term suggestion devices in Interactive Query Expansion (IQE) have been designed to present candidate expansion terms in a form of list, regardless the underlying ranking techniques. While such a list has been the dominant form of presentation, it seems less effective to provide users with the contexts from which the terms are derived.

To explore an alternative approach to support IQE, we have been investigating methods of presenting expansion terms in a hierarchical form: automatically and dynamically derived from a set of retrieved documents; terms (concepts) are organised based on term specificity, with allowing links to contextually related terms (e.g. Yahoo Directory). It is anticipated that such a hierarchy will provide users with an overview of documents, and can be useful to allow users to articulate their information need.

In our demonstration, we use a technique called "subsumption hierarchy" [1] to construct a concept hierarchy. The subsumption hierarchy uses the co-occurrence information to identify a pair of terms that are related, and measures term specificity using document frequency. Although this is a simple technique, it is encouraging that the technique takes account into term specificity in constructing hierarchies. A user test has shown a positive reaction of the hierarchy in the expansion task [2].

Our live demonstration shows the hierarchy system working on a locally run search engine. Hierarchies are dynamically generated from the retrieved documents, and visualised on the menus (see Figure 1: the number in brackets beside terms indicates the number of documents linked to the terms). When a user selects a term from the hierarchy, the documents linked to the term are listed, and the term is then added to the initial query to rerun a search. Through the demonstration we illustrate how hierarchical presentation of expansion terms is achieved, and how our approach will support users to articulate their information needs using the hierarchy. We also briefly discuss possible ways to improve the current system.

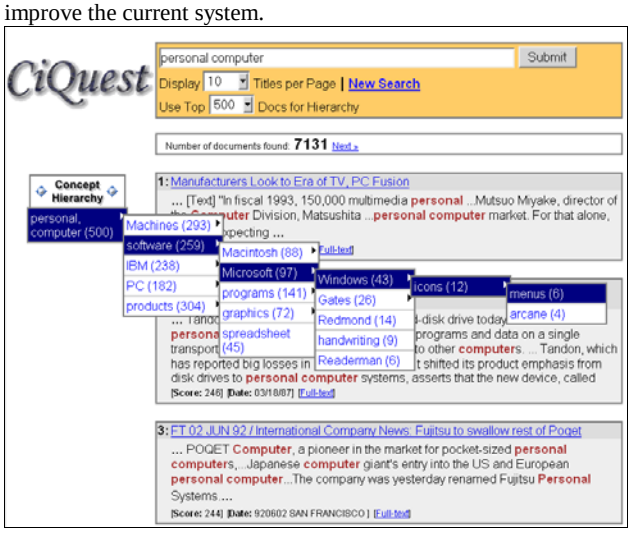

Figure 1. Screenshot of CiQuest System: a sample hierarchy generated from the top 500 documents retrieved in response to a query "personal computer". The maximum number of terms at each level was truncated to five for this screenshot.

2. ACKNOWLEGEMENT

This work was carried out as a part of the CiQuest project, which is funded by the Library and Information Commission (now *re*:source, the Council for Museums, Archives and Libraries). Any opinions, findings, and conclusions described here are the authors and do not necessarily reflect those of the sponsor.

3. REFERENCES

[1] Sanderson, M. and Croft, B. Deriving Concept Hierarchies from Text. In: *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 206-213, Berkeley, CA, 1999, ACM.

[2] Joho, H., Coverson, C., Sanderson, M. and Beaulieu, M. Hierarchical Presentation of Expansion Terms. In: *Proceedings of the 17th ACM Symposium on Applied Computing (SAC'2002)*, 645-649, Madrid, Spain, 2002, ACM.

SIGIR '02, August 11-15, 2002, Tampere, Finland.
ACM 1-58113-561-0/02/0008.