Using Syntagmatic Relationships based on a RAMEAU list as a Browsing Relevance Feedback Strategy in a WWW-OPAC

Abstract: This paper reports on an evaluation of the browsing behaviour of end users of a WWW-OPAC focusing on the browsing relevance feedback (BRF) strategy. Results of this study reveal that BRF is a popular strategy. We also find that the relationships involved in the BRF strategy are generally syntagmatic.

1. Hypertext and OPACs

According to Hildreth (1989), earlier browsing capabilities in OPACs (online public access catalogs) fall into two categories:
1: Pre-sequenced, linear, inflexible browsing
2: Non linear, multidirectional, flexible browsing.

The first category is related to vocabulary browsing. Most of the research over the past twenty years has been focused on enhancements in this category, e.g.:
• Enriching the entry vocabulary (Bates, 1986).
• Linking free text to vocabulary terms (Micco, 1993; Hildreth, 1993).
• Visualization of LCSH structure and relationships (Chryso, 1996).
• Developing more browsable displays of long lists of subdivisions (Allen, 1993; Markey and al, 1994).
• Defining a taxonomy of relationships in LCSH.

In France very few OPACs have incorporated the syndetic structure of RAMEAU (Répertoire des Autorités Matières Encyclopédique et Alphabetique Unifiée). In a recent survey on subject access to academic catalogs, we found that only 38% catalogs provided "see also" reference (see Table 1) (Ihadjedene, 1997).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>see</td>
<td>43 (62%)</td>
</tr>
<tr>
<td>see also</td>
<td>25 (38%)</td>
</tr>
<tr>
<td>BT, NT, RT</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1. Syndetic references in French academic OPACs

The problem of using hypertext to represent RAMEAU is presented elsewhere (Ihadjadene, 1997).

The second category is related to search mechanisms that support discovery through non-linear exploration of a database. This category of browsing was seldom used in the

---

1 These questions are under scrutiny by the American Library Association- ALCTS-CCS-SAC Subcommitee on Subject Relationships/Reference Structures.
second generation of OPACs: very few systems allow related searches to be carried out.

Although there are some experimental OPACs that offer non-linear browsing such as: Hypercatalog (Hjerppe, 1986), WHIZARD (Nelson, 1992) and SLS (Borgman, 1989), this functionality is to be expected of OPACS made available via a WWW browser, hence WWW-OPACs form the basis of this study.

According to Walker (Walker, 1989), experienced users often try to find one or a few relevant references and to look at the terms by which these records are indexed and then try to use some of these as new search terms. This is an example of relevance feedback. For Meadow (Meadow, 1992) the "use of a retrieved record for reformulation is much like relevance feedback, but the user does not rate each individual record".

Researchers in the past have found that users rarely exploited this strategy, because they often do not understand the collocative function of subject headings for identifying similar work (Hassoun, 1994). Most of the designers of traditional OPACs presuppose that the display of bibliographic records represents the end of a search.

This searching strategy has been named differently by different researchers:
Citation pearl growing (Markey and al, 1978), item instantiation strategy (Chen and al, 1991) and Query expansion (Walker, 1989). In a mediated online searching, Spink (Spink, 1995) called this strategy term relevance feedback.

All these strategies have two sub-processes in common:
- Examination of search results
- Extraction of information from retrieved items.

These strategies were also implemented in various expert systems and are equivalent to the "similarity selection" feature conceived by Swanson in 1964 (Su, 1994).

Some of the WWW-OPACs offer related record browsing (e.g.: records that share the same subject headings, or the same author). We refer to this strategy as Browsing Relevance Feedback (BRF). It is one of several inter-record browsing strategies.

The source that provides the terms for feedback is based on the search results. Browsing is the method used to select terms that will be used in reformulation. BRF is incremental and it is under the user's control (the determination of items is the responsibility of the users themselves rather than the system).

In this paper, we will evaluate the BRF strategy in two WWW-Opacs using RAMEAU. RAMEAU is the primary subject access and matching vocabulary used in French OPACs. It has two major kinds of relationships; paradigmatic and syntagmatic.

As in LCSH systems, the expression of syntagmatic relationships takes several forms in RAMEAU: Conjunctive phrase headings, prepositional phrase headings, qualified headings, adjectival phrases and subdivided headings. For more information about syntagmatics relationships and their role in enhancing precision in information retrieval, see (Bouché, 1989), (Green, 1994) and (Green and al, 1995).

The relationships involved in BRF strategy are generally syntagmatic.

Related Work
- In his doctoral thesis, Major (Major, 1981) studied the effect of these feedback mechanisms in an online catalog; display of adjacent terms, display of cross reference terms, display of all subject tracings (additional descriptors assigned to documents whose citations are retrieved). Results of this study showed the display of search terms related
through the subject authority structure (see also references) was significantly less effective than two other online treatments.

- Wolfram and colleagues (Wolfram and al, 1995) report on a study that examined the effect of inter record linkage (author, descriptor, term from title, term from abstract) in a hypertext-based bibliographic system. The authors conclude that increasing title and abstract keyword as a hypertext linking option does not improve retrieval performance.

- In a controlled experiment, results of the study of Hildreth (Hildreth, 1993) revealed that navigation on subject headings or title words resulted in higher recall but precision suffered in title word navigation.

- Spink et al (Spink, 1995) investigated the sources and the effectiveness of search terms used during 40 mediated online searching sessions. Term relevance feedback contributed 11% of search terms. Large proportions of the terms relevance feedback were selected from title (27%), descriptors (58%), identifiers (3%), abstracts (3%), journal names (3%) and only 3% were selected from subject headings.

- Beaulieu (Beaulieu, 1997) reports on a study that examined the effect of query expansion in OKAPI catalogs. Query expansion was used by about 33% of catalog users.

- Most of these studies have relevance for our study. However, we could not find any other study specifically addressing the browsing behavior of end-users in WWW-OPACs.

3. Methodology

Three types of data were collected:

- **Transaction log:** Online transaction logs data were collected during one year (1997) of the WWW-OPAC of ENSSIB and the online transaction logs data during two years (1996-1997) from a university webbed-catalog (Lyon 2). Both WWW-OPACs use frames.

- **Questionnaires:** Two questionnaires were used (pre and post-questionnaire) in order to capture demographic data of 44 humanities students from Lyon 2. We have also collected their interests, their previous experience and their search intentions.

- **Verbal data:** Searchers were asked to explain the reasons for their link selection. For the analysis we used think aloud protocol to examine the browsing behavior of the 44 users.

One limitation of this study is that we did not explore the effectiveness of search terms. That is, how the BRF influences search results.

4. Results

4.1. Usage data.

In a previous study (Ihadjadene, 1998) that examined the searching and browsing behaviour of ENSSIB users, we found that using syntagmatic relationships as a BRF strategy in a WWW-OPAC is popular for both local and remote users (14% vs 16%) (table1).

However, the users of ENSSIB being to a certain extent specialists in information science, we wanted to extend our study to a non-specialist public; we have done a similar study in another environment (public university catalogs). Results show that BRF is also a favourite strategy (12%) (table2).
Table 2: local vs remote access of WWW-OPACs (ENSSIB).

The enhancements of OPACs with hypertext interfaces such as those developed for the WWW leads users to develop search strategies.

Table 3. Access and browsing of two WWW-OPACs (ENSSIB vs Lyon2)

- The differences between the ENSSIB and Lyon2 usage may be explained by:
  - A large number of entries in the Lyon2 OPAC do not have subject headings.
  - Users of ENSSIB are more specialised users (in using the www and in documentation).
  - The Lyon2 WWW-OPAC offers publisher links.

Users browse generally by using subject headings (70% to ENSSIB vs 44% to Lyon2), followed by the author links (20% vs 31%) (see table 4). In information retrieval systems that incorporate feedback process, the assumption is that the subjects (keywords) only are of importance for reformulating the queries. Our results reveal the potential use of other sources of terms in reformulation such as author.
4.2 What motivates the users for browsing?

This study involved 44 students (30 of them are undergraduate and 14 are graduate students) from Lyon2. The majority of users (90%) had little experience in using WWW-OPACs.

Only 13 (30%) of users have used the BRF strategy which involves an average 2 links per session; they prefer browsing by subject headings rather than by other links (see table 5).

<table>
<thead>
<tr>
<th>Link Type</th>
<th>Percentage Lyon2 (N=4073)</th>
<th>Percentage ENSSIB (N=4727)</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>44%</td>
<td>70%</td>
</tr>
<tr>
<td>author</td>
<td>31%</td>
<td>20%</td>
</tr>
<tr>
<td>series</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>corporate author</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>publisher</td>
<td>11%</td>
<td>N/A</td>
</tr>
<tr>
<td>total</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 4. Links used for browsing the two WWW-OPACs

Users were asked to explain the reason of their selection of each link (subject heading, author, corporate author). Some reasons given by users for their subject heading link selection are:

- Six users have used the subject links to narrow a search. It takes many forms:
  - browsing assists users in identifying the correct form of a term
  - identifying terms that are narrower than their topic of interest
  - reducing search results (magnitude feedback)
- Four users have used the subject links to find similar works.
- Only one user has used a subject heading in order to broaden a search.
- Two users used 3 subject links as a browsing aid. They had some problems in linking back in a frame based WWW-OPAC, so they used the links to return to the list of references.

The main reasons given by users for not using browsing were:

- They were satisfied with the response because they had found the books(s) they were looking for (11 users)
- They choose one or two call numbers and they go on browsing in stacks (7 users).
- Users have some problems seeing if the retrieved record is relevant or not. This is due to the lack of information (subject headings or table of contents) (4 users)
- Users have some difficulty in understanding the meaning of subject headings. For example, the artificial forms of inverted headings do not invite them to browse. (3 users)
- They are not familiar with the WWW interface (3 users)
- Lack of time (3 users)
5. Conclusions:

The study highlights how the browsing relevance feedback strategy was used in two different environments (ENSSIB vs Lyon2). Results show that the browsing relevance feedback is quite often used. Basically, it appears that the strategy of the users is to narrow their search. One of the recurrent criticisms about subject headings is their linear structure which is unsuitable for online retrieval. Our results show that syntagmatic relationships (precoordination) can play an important role in refining a search and helping users to recognise what cannot be clearly specified in advance. However, the limited number of subject headings in records is perceived as a restriction. Since the number of subject headings per book was small (1.4 in the case of ENSSIB catalogue), the variety and choice offered to users to browse is also limited.

Enhancing the content of the bibliographic record will
- help an end user to determine the usefulness of a retrieved item (examination stage)
- present a variety of choices of subject links (extraction stage).

Until now, most of the research on the display of subject headings was done in order to help users in formulating a request (ALA,1992). We suggest that this research is reviewed to take into account the reformulation process, in particular to consider the following.
- link density: what is the average number of links to be displayed?
- horizontal overload: subject headings are sometimes quite long. To increase end-user understanding, effort must be made to shorten subject headings by reducing the number of subdivisions (Franz and al,1994).
- what is the impact of the display of headings at the top of records vs the bottom of records.
- develop more reformulation devices like the "summarize" function of the OASIS catalogs (Buckland and al,1992) which lists all subject headings in retrieved sets ranked by decreasing frequency of occurrences.

We plan to carry out a study of the browsing behavior of end-users in a non-frame environment. Moreover, we will do a usability study of the WWW-OPAC interfaces.

References

Advances in Knowledge Organization, Vol.6(1998)


Advances in Knowledge Organization, Vol.6(1998)