Distributed Agents for User-Friendly Access of Digital Libraries

DAFFODIL
Effective Support for Using Digital Libraries

Norbert Fuhr
University of Duisburg-Essen, Germany

*Supported by DFG as part of the Research Programme V$^3$D$^2*
Desktop system for accessing federated digital libraries
Daffodil concepts

1. Strategic information access support
2. Proactive system support
3. Digital library life cycle
4. Collaboration
1. Strategic information access support

Levels of search activities (Bates 1990):

1. **Move**: Low-level search function
   (e.g. type in search term, view retrieved document)

2. **Tactic**: several moves to further a search
   (e.g. broaden/narrow a query)

3. **Stratagem**: set of actions on a single domain
   (citation database, tables of contents of journals)

4. **Strategy**: complete plan for satisfying an information need
   (e.g. subject search, browse relevant journals, find referenced articles)
## Result list

**Search**

Query: Title=language model AND Free-Text=retrieval

### Found: 5

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Results

1. **Jerome R. Bellegarda**  
   Statistical *language model* adaptation: Review and perspectives.  

2. **P. C. Ching; Wai-Kit Lu; Helen Meng**  
   Cross-language spoken document retrieval using HMM-based *retrieval model* with multi-scale fusion.  
   (2003) from DBLP2; Achilles.

3. **W. Bruce Croft**  
   Language Models for Information Retrieval  

4. **Rohini K. Srihari; Munirathnam Srikanth**  
   Incorporating query term dependencies in *language models* for document retrieval.  

5. **Hirofumi Yamamoto; Shuntaro Isogai; Yoshinori Sagisaka**  
   Multi-class composite $N$-gram *language model*.  

---

**Moves** – Tactics – Stratagems – Strategies – GUI
Distributed Agents for User-Friendly Access of Digital Libraries

Abstract:

Cross-language spoken document retrieval (CL-SDR) is the technology that facilitates automatic retrieval of relevant information from a collection of spoken documents in a language that is different from that used in the queries. Information sources that are in different languages can then be retrieved automatically with CL-SDR, and the number of searchable information sources will increase significantly. The HMM-based retrieval model is a probabilistic formulation for the retrieval problem. Extensions to this retrieval model can be made by taking advantage of its probabilistic nature. Specifically, we have incorporated the translation component to make it possible to perform cross-language information retrieval (CLIR). In addition, this HMM-based CLIR retrieval model is also extended for retrieval at subword scales. In this work, the extended HMM-based retrieval model has been applied to an English-Mandarin CL-SDR task, which is to search the Mandarin spoken document collection with English queries at word and subword scales. Retrieval results obtained from these indexing scales are then fused for multi-scale retrieval. Experimental results demonstrate that improvement in CL-SDR retrieval performance can be achieved by fusion of word and subword scales.

Possible actions on this document:

- This document has the following external links:
  - Fulltext at portal.acm.org
  - Fulltext at doi.acm.org

- This document was found on the following datasources:
  - DBLP
  - Achilles
  - ACM

Theory - Search in Thesaurus
- Performance Search in Thesaurus

**Moves** – Tactics – Stratagems – Strategies – GUI
Distributed Agents for User-Friendly Access of Digital Libraries

Extract from Result List

Moves – Tactics – Stratagems – Strategies – GUI
Stratagems on the Desktop

- Subject Search
- Journal/Conference Run
- Author networks
- Citation Search
Distributed Agents for User-Friendly Access of Digital Libraries

Author Networks
Citation Search: Reference Tracking

- A relevant document is dragged to the tool for citation tracking
- Documents that cite or are cited are retrieved
- Browsing, Inspection and Navigation
- Drag and Drop
  - Can be performed iteratively
2. Proactive system support

Use form fields to specify query. The ; splits two query clauses. The default connection operator is AND. The query can use OR, AND, NOT. ( ) - (edit below or on next page).

Author = Thanos

Use context menu on item:
- animals
- coronary artery bypass
- database system
- digital library
- distributed database
- fragmented
- human
- laminin
- logical implication
- mobilization
- multimedia documents
- non classical
- object oriented
- optic nerve
- rotational
- sequent calculus
Proactivity during query formulation

Spell checking
Proactivity during query formulation (2)

Semantically inconsistent queries
3. Digital library life cycle

- Knowledge generation
- Resource selection
- Retrieving items
- Organizing items
- Evaluation of results
- Collate
- Re-present
- Discover
- Retrieve
- Interpret
Collate

Personal library:

Save all digital objects in a structured personal digital library

- Document Metadata
- Fulltexts
- Authors
- Journals, Conferences
- Hyperlinks, Bookmarks
- Query formulations

Discover – Retrieve – **Collate** – Interpret – Re-Present
Personal Library

- Personalized Keyword Search with Partial-Order Preferences
  - Preference Structures and Their Numerical Representations
  - Personalized Nonlinear Ranking Using Full-text Preferences
- Evaluation
- XML IR
- Preferences and IR
Distributed Agents for User-Friendly Access of Digital Libraries

Interpret: Annotations

Discover – Retrieve – Collate – **Interpret** – Re-Present
Distributed Agents for User-Friendly Access of Digital Libraries

Re-Present

(support generation of new information)

Export folders from personal library

Discover – Retrieve – Collate – Interpret – Re-Present
4. Collaboration

Discover – Retrieve – Collate – Interpret – Re-Present
Collation and awareness

- New objects
- Filed by other users
- Recommendation
- Notification mechanisms
- visual markers
- email

Discover – Retrieve – **Collate** – Interpret – Re-Present
Collaborate interpretation

Annotation threads

Discover – Retrieve – Collate – **Interpret** – Re-Present
Collaboration: Chat-Tool

- Help
- Discussion
- Object exchange
- Ask an expert (Librarian)
- Chatter-Bot
5. Evaluation

- Evaluation with 28 subjects
- Methods: questionnaire, logging, video
- Tasks of different complexity
- Major results:
  - Only for very simple tasks (known item retrieval), other methods may be faster
  - Complex tasks are supported very effectively
Usability studies

- 30 subjects
- Focus on query formulation

Results:
- Users feel rather uncertain
  - submit the same query several times,
  - stare at the keyboard
- Users ignore hints shown elsewhere on the desktop

Improvements through proactive functions:
- Fewer syntactical and semantical errors
- Better tactical decisions
- Users feel more confident
6. Summary

- Daffodil concepts
  - Strategic information access support
  - Proactive system support
  - Digital library life cycle
  - Collaboration
- Daffodil can be easily ported to other application domains
- The service-based architecture allows for easy integration into other systems
Representational Model of Common Elements

- Interaction
  - Understanding Queries
  - Massaging Results & User Annotation
  - Transformations
  - Discovery
  - Automatic Capture
  - Interpretive Capture
- User
  - Cognitive Completion
  - DWIM
  - Collaboration
  - Managing Personal Libraries
- Store

[Larsen & Wactlar: Summary of NSF WS on Future Directions for DLs, 2003]