Topic driven access to scientific handbooks

Caracciolo, C.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Contents

Acknowledgments xiii

1 Introduction 1
  1.1 Problem Statement .................................. 2
  1.2 Organization of this Thesis ........................... 5
  1.3 Scope of this Thesis .................................. 7
  1.4 Origins of the Material ............................... 8

2 A Vision on Access to Electronic Scientific Handbooks 9
  2.1 The Vision ........................................ 10
  2.2 Handbooks .......................................... 12
  2.3 Searching and Accessing Textual Documents .......... 13
  2.4 Semantic Structures .................................. 14
  2.5 A Modular Approach to Focused Access ................ 20
  2.6 Discussion ........................................ 21

3 A Browsable Map for Logic and Language 23
  3.1 User Requirements ................................... 23
  3.2 Design and Content of the Map ........................ 26
  3.3 Topics ............................................. 27
  3.4 Relations .......................................... 28
    3.4.1 Many Flavors of ISA ............................. 30
    3.4.2 Part-of .......................................... 33
    3.4.3 Instance ......................................... 34
    3.4.4 Domain Specific Relations ......................... 34
    3.4.5 Non-Hierarchical Relations ...................... 36
  3.5 The LoLaLi Map: Features ............................. 36
3.6 Discussion ................................ 39
  3.6.1 Dealing with more Relations ...................... 39
  3.6.2 More on Subtopics .................................. 40
3.7 Editing and Managing the LoLaLi Map .................. 43
  3.7.1 A Bit of History: First Attempts .................. 43
  3.7.2 Protégé and RDFS Modeling ....................... 44
  3.7.3 Accessing the Map through a Web Browser .......... 46
3.8 Conclusions ............................................ 46

4 Interacting with the LoLaLi Map .......................... 49
  4.1 Requirements for the User Interface ................ 50
  4.2 Related Work on Tree and Graph Visualization ......... 51
    4.2.1 Global views .................................. 51
    4.2.2 Local views ................................... 55
  4.3 A User Interface for the LoLaLi Map .................. 57
  4.4 User Studies ...................................... 61
    4.4.1 The Setting ................................ 62
    4.4.2 The Questionnaire ........................... 64
    4.4.3 Results ..................................... 65
    4.4.4 Discussion ................................... 77
  4.5 Conclusions ......................................... 79

5 Looking for Link Targets .................................. 83
  5.1 Passage Retrieval ................................... 85
    5.1.1 Structural Passage Retrieval .................. 86
    5.1.2 Fixed Size Passage Retrieval .................. 86
    5.1.3 Semantic Passage Retrieval ................... 87
  5.2 Linear Topic Segmentation: Two Algorithms ............. 88
    5.2.1 TextTiling .................................... 88
    5.2.2 C99 ........................................... 90
  5.3 Building the Ground Truth ............................ 91
    5.3.1 Inter-Annotator Agreement .................. 92
    5.3.2 The Handbook of Logic and Language .......... 93
    5.3.3 The Resulting Ground Truth .................. 94
  5.4 Evaluating Link Targets Against the Ground Truth ........ 96
    5.4.1 Evaluation Measures for Topic Segmentation ...... 96
    5.4.2 Results ..................................... 98
  5.5 Discussion ................................... 100
  5.6 Conclusions ......................................... 101

6 Connecting the Map and Link Targets .................. 103
  6.1 Annotation of Relevance Assessments ................ 103
  6.2 How to Evaluate a Link Target ........................ 104
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3 Evaluating a Collection of Segments</td>
<td>108</td>
</tr>
<tr>
<td>6.4 Experimental Setting and Results</td>
<td>109</td>
</tr>
<tr>
<td>6.5 Conclusions</td>
<td>112</td>
</tr>
<tr>
<td><strong>7 Conclusions</strong></td>
<td>115</td>
</tr>
<tr>
<td>7.1 How to Organize and Visualize a Domain Map</td>
<td>115</td>
</tr>
<tr>
<td>7.2 Linking the Map to the Handbook</td>
<td>117</td>
</tr>
<tr>
<td>7.3 The Bigger Picture</td>
<td>118</td>
</tr>
<tr>
<td>7.4 Directions for Future Work</td>
<td>120</td>
</tr>
<tr>
<td><strong>A Questionnaire for User Studies</strong></td>
<td>123</td>
</tr>
<tr>
<td>A.1 Computer Literacy</td>
<td>123</td>
</tr>
<tr>
<td>A.2 Information Gathering Strategies</td>
<td>125</td>
</tr>
<tr>
<td>A.3 Browsing the Interface</td>
<td>126</td>
</tr>
<tr>
<td>A.4 Search and Read</td>
<td>127</td>
</tr>
<tr>
<td>A.5 External Links</td>
<td>128</td>
</tr>
<tr>
<td>A.6 Features of the Map</td>
<td>128</td>
</tr>
<tr>
<td>A.7 User Preferences</td>
<td>129</td>
</tr>
<tr>
<td>A.8 User Wishes</td>
<td>130</td>
</tr>
<tr>
<td><strong>B Glossary</strong></td>
<td>131</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>135</td>
</tr>
<tr>
<td><strong>Samenvatting</strong></td>
<td>147</td>
</tr>
</tbody>
</table>