

# Table of Contents

|   |      |
|---|------|
| List of Abbreviations.....                                    | vi   |
| List of Tables .....  | vii  |
| List of Figures .....   | viii |
| <br>  |      |
| Chapter 1: Introduction .....                                 | 1    |
| 1.1 Motivation.....   | 1    |
| 1.2 HyperContext.....   | 2    |
| 1.3 Thesis overview .....                                     | 3    |
| 1.3.1 Chapter 2 .....   | 3    |
| 1.3.2 Chapter 3 .....   | 3    |
| 1.3.3 Chapter 4 .....   | 4    |
| 1.3.4 Chapter 5 .....   | 5    |
| 1.3.5 Chapter 6 .....   | 5    |
| 1.3.6 Chapter 7 .....   | 6    |
| 1.3.7 Chapter 8 .....   | 6    |
| 1.3.8 Chapter 9 .....   | 6    |
| 1.3.9 Chapter 10.....   | 6    |
| 1.4 Summary of contributions .....                            | 6    |
| <br>  |      |
| Chapter 2: Background .....                                   | 8    |
| 2.1 Preamble .....  | 8    |
| 2.2 Searching and navigating through information spaces.....  | 9    |
| 2.3 Organisation of information versus meta-information ..... | 11   |
| 2.4 Information re-use .....                                  | 13   |
| 2.5 Navigational difficulties in an information space.....    | 15   |
| 2.6 Assisting navigation in a digital information space.....  | 17   |
| 2.7 Summary.....  | 18   |

|   |    |
|---|----|
| Chapter 3: Literature Review.....   | 20 |
| 3.1 Introduction.....   | 20 |
| 3.1.1 General-purpose hypertext systems.....  | 20 |
| 3.1.2 General-purpose Information Retrieval systems.....                                    | 24 |
| 3.2 Towards Adaptive Hypertext Systems .....  | 26 |
| 3.3 Web-based Adaptive Hypertext Systems.....   | 31 |
| 3.3.1 Systems review.....   | 33 |
| 3.4 Reflecting short-term interests in a user model.....                                    | 38 |
| 3.5 Summary.....  | 40 |
| <br>  |    |
| Chapter 4: HyperContext: A Framework for Adaptive and Adaptable Hypertext ..                | 42 |
| 4.1 Introduction.....   | 42 |
| 4.2 HyperContext: A three-layer approach .....  | 44 |
| 4.3 Interpreting information in context: the Structure and Object Layers .....              | 46 |
| 4.4 A description of HyperContext in terms of the city metaphor .....                       | 54 |
| 4.5 Interacting with HyperContext: the Presentation Layer.....                              | 56 |
| 4.6 Summary.....  | 59 |
| <br>  |    |
| Chapter 5: Adaptability and Adaptivity: Putting the user in the picture .....               | 61 |
| 5.1 Introduction.....   | 61 |
| 5.2 User adaptation of a HyperContext hyperspace .....                                      | 61 |
| 5.3 A HyperContext implementation model (1).....  | 65 |
| 5.4 Browsing through a HyperContext hyperspace .....  | 66 |
| 5.5 Extending the metaphor .....  | 67 |
| 5.6 Adapting the hyperspace to the user .....   | 68 |
| 5.7 The TIR and IRC algorithms .....  | 71 |
| 5.7.1 Traditional Information Retrieval.....  | 71 |
| 5.7.2 Information Retrieval-in-Context.....   | 73 |
| 5.8 The short-term user model .....   | 74 |
| 5.8.1 Deriving a model of the user's interests from an interpretation .....                 | 74 |
| 5.8.2 Deriving a model of the user's short-term interests from the<br>context session ..... | 81 |
| 5.8.3 Abstracting a query from the user model.....  | 83 |
| 5.8.4 Openness of the short-term user model.....  | 83 |
| 5.9 Adaptive Information Discovery .....  | 83 |
| 5.10 Summary .....  | 84 |
| <br>  |    |
| Chapter 6: Context and HyperContext.....  | 86 |
| 6.1 Introduction.....   | 86 |

|   |         |
|---|---------|
| 6.2 Background.....   | 87      |
| 6.2.1 What is “context”?.....   | 88      |
| 6.2.2 A logic of contexts .....   | 89      |
| 6.2.3 The applications of context.....                                  | 91      |
| 6.3 Context in HyperContext.....  | 92      |
| 6.3.1 Interpreting information in context .....                         | 92      |
| Dynamic links .....   | 93      |
| Arbitrary interpretations of documents .....                            | 94      |
| Favouring predetermined links and interpretations .....                 | 96      |
| 6.4 The nature of context in HyperContext.....                          | 96      |
| 6.5 Relevance and context .....   | 98      |
| 6.6 Guiding users to relevant information .....                         | 100     |
| 6.6.1 Determining contextual relevance.....                             | 100     |
| 6.6.2 Top-down versus bottom-up context resolution.....                 | 102     |
| 6.7 Context and the short-term user model .....                         | 103     |
| 6.7.1 Deriving a salient interpretation.....                            | 103     |
| 6.7.2 Incorporating the salient interpretation into the user model..... | 107     |
| 6.8 Summary .....   | 108     |
| <br>Chapter 7: The HyperContext Prototype .....                         | <br>110 |
| 7.1 Introduction.....   | 110     |
| 7.2 HyperContext features included in the prototype.....                | 110     |
| 7.3 Prototype implementation details.....                               | 111     |
| 7.4 The HyperContext user interface.....                                | 112     |
| 7.4.1 The HyperContext Control Panel.....                               | 112     |
| 7.4.2 The HyperContext Document Frame.....                              | 113     |
| Link hiding versus link disabling .....                                 | 114     |
| 7.4.3 HyperContext's "See Also" Links .....                             | 115     |
| 7.4.4 HyperContext Status Information .....                             | 116     |
| 7.5 HyperContext Prototype Implementation Model (2).....                | 117     |
| 7.5.1 Communication between modules.....                                | 118     |
| 7.6 The HyperContext server .....                                       | 119     |
| 7.6.1 Interpreting a document in context .....                          | 120     |
| 7.6.2 Sending a document's contexts.....                                | 122     |
| 7.6.3 Satisfying a search request .....                                 | 122     |
| Traditional Information Retrieval (TIR).....                            | 123     |
| Information Retrieval-in-Context (IRC).....                             | 124     |
| Adaptive Information Discovery (AID) .....                              | 126     |
| 7.7 The HyperContext client .....                                       | 126     |

|   |     |
|---|-----|
| 7.7.1 The context session .....   | 127 |
| 7.7.2 The short-term user model .....   | 127 |
| 7.7.3 Adaptive Information Discovery: query formulation and link<br>recommendation..... | 127 |
| 7.8 A non-adaptive version of HyperContext .....  | 128 |
| 7.9 Summary .....   | 129 |
| <br>Chapter 8: The Conversion of a Web Site into a HyperContext Hypertext.....          | 130 |
| 8.1 Introduction.....   | 130 |
| 8.2 Preparing the mirror site.....  | 131 |
| 8.2.1 The HyperContext Web Robot .....  | 131 |
| 8.2.2 Determining a document's parents .....  | 132 |
| 8.2.3 HTML file preparation for HyperContext .....                                      | 132 |
| 8.3 Generating the HyperContext hypertext.....  | 132 |
| 8.3.1 Generating the Object Layer.....  | 132 |
| 8.3.2 Generating the Structure Layer .....  | 133 |
| 8.4 SWISH-E: Simple Web Indexing System for Humans - Enhanced .....                     | 134 |
| 8.5 The structure of the interpretations file .....                                     | 135 |
| 8.6 Creating the interpretations files .....  | 137 |
| 8.7 HyperContext hypertext statistics .....   | 137 |
| <br>Chapter 9: HyperContext Evaluation.....   | 140 |
| 9.1 Introduction.....   | 140 |
| 9.2 Background.....   | 140 |
| 9.3 Setting the boundaries .....  | 141 |
| 9.4 The short-term user model .....   | 142 |
| 9.4.1 Creating a test-bed.....  | 142 |
| 9.4.2 Automatically generating paths of traversal .....                                 | 146 |
| 9.4.3 Description of the experiments.....   | 146 |
| 9.4.4 Generating a query from the user model.....                                       | 148 |
| 9.4.5 Results.....  | 148 |
| 9.4.6 Prediction performance .....  | 150 |
| 9.5 Locating relevant information.....  | 151 |
| 9.5.1 Ideal empirical study scenario.....   | 151 |
| 9.5.2 Automatic evaluation of the HyperContext hypertext .....                          | 152 |
| 9.5.3 Description of the experiments.....   | 156 |
| 9.5.4 Results.....  | 157 |
| 9.5.5 Breakdown of the results according to path length.....                            | 159 |
| 9.5.6 Comparison of number of nodes visited by the different phases ...                 | 162 |

|   |     |
|---|-----|
| 9.6 Observations.....   | 164 |
| 9.7 Recommending documents using AID .....  | 165 |
| Chapter 10: Conclusion and Future Work.....   | 172 |
| 10.1 Introduction .....   | 172 |
| 10.2 Summary of main proposals and findings.....                                      | 173 |
| 10.3 Future Work.....   | 176 |
| 10.3.1 Prototype implementation.....  | 177 |
| 10.3.2 Extensions to the HyperContext framework.....                                  | 179 |
| Non-linear context sessions .....   | 179 |
| Aggregate nodes .....   | 179 |
| Link and node typing.....   | 180 |
| Adaptive Presentation techniques .....  | 181 |
| Link and interpretation ageing.....   | 181 |
| Adaptive Information Discovery recommendations based on<br>skim and deep reading..... | 181 |
| Dynamic data .....  | 182 |
| Link hiding and link disabling.....   | 182 |
| 10.3.3 Other research opportunities in HyperContext .....                             | 183 |
| Efficient context expansion .....   | 183 |
| Using adaptive methods to decide when a user model can be<br>trusted.....             | 183 |
| Ontology .....  | 183 |
| Object migration.....   | 184 |
| Automatic creation of interpretations .....   | 184 |
| Identifying new linking opportunities .....   | 184 |
| Using HyperContext with different external models of<br>Information Retrieval.....    | 185 |
| Alternative representations of interpretations .....                                  | 185 |
| HyperContext and the Semantic Web .....   | 185 |
| 10.3.4 Research opportunities in other domains.....                                   | 188 |
| Information Retrieval.....  | 188 |
| Intelligent Agents .....  | 189 |
| Information re-use .....  | 189 |
| Glossary of Selected Terms.....   | 190 |
| Bibliography .....  | 192 |

## List of Abbreviations

|      |   |
|------|---|
| ABF  | Adaptive Brute Force (Phase 3 of the HyperContext evaluation)     |
| AHS  | Adaptive Hypertext System   |
| AI   | Artificial Intelligence   |
| AID  | Adaptive Information Discovery                                    |
| AO   | Adaptive Ordered (Phase 4 of the HyperContext evaluation)         |
| CBR  | Case-Based Reasoning  |
| CGI  | Common Gateway Interface  |
| HTML | HyperText Mark-up Language  |
| HTTP | HyperText Transfer Protocol                                       |
| IDF  | Inverse Document Frequency  |
| IR   | Information Retrieval   |
| IRC  | Information Retrieval-in-Context                                  |
| ITS  | Intelligent Tutoring System                                       |
| NABF | Non-Adaptive Brute Force (Phase 1 of the HyperContext evaluation) |
| NAO  | Non-Adaptive Ordered (Phase 2 of the HyperContext evaluation)     |
| TF   | Term Frequency  |
| TIR  | Traditional Information Retrieval                                 |
| URL  | Uniform Resource Locator  |
| WWW  | World Wide Web  |
| XML  | EXtensible Mark-up Language                                       |

## List of Tables

|   |     |
|---|-----|
| 5.1 Weights of terms $t_0$ to $t_4$ in interpretations $I_0$ to $I_4$ of document I ..... | 75  |
| 8.1 Number of contexts per node .....   | 138 |
| 8.2 Number of out-links per node.....   | 138 |
| 9.1 Comparison of probabilities of recommending an unseen letter .....                    | 151 |
| 9.2 Comparison of results of ABF and NABF .....   | 157 |
| 9.3 Comparison of results of NAO and NABF.....  | 157 |
| 9.4 Comparison of results of AO and ABF.....  | 157 |
| 9.5 Comparison of results of AO and NAO.....  | 158 |
| 9.6 Comparison of results of AO and NABF.....   | 159 |
| 9.7 Comparison of results of ABF and NAO.....   | 159 |
| 9.8 Number of successes by phase by level depth of target node .....                      | 160 |
| 9.9 Summary of the three "invalid" experiments .....                                      | 161 |
| 9.10 Average number of nodes visited overall by phase .....                               | 162 |
| 9.11 Average number of nodes visited in each phase.....                                   | 162 |
| 9.12 Number of nodes at each level by hypertext .....                                     | 163 |
| 9.13 Average number of nodes visited during successful phase .....                        | 163 |
| 9.14 Successes by phase when target is at the same level.....                             | 164 |
| 9.15 Relevance judgements .....   | 166 |
| 9.16 A sample set evaluation log file.....  | 167 |
| 9.17 Aggregated results of the AID evaluation .....                                       | 169 |
| 9.18 Performance according to which recommended document was shown first                  | 169 |
| 9.19 Performance according to skim or deep reading a document set .....                   | 170 |
| 9.20 Performance according to skim or deep reading, and document order .....              | 170 |

## List of Figures

|  |     |
|--|-----|
| 3.1 A simple hypertext.....  | 21  |
| 3.2 Examples of navigational, index, and context-sensitive links.....              | 26  |
| 3.3 Examples of context-free or "See Also" links.....                              | 28  |
| 4.1 The three layers of the HyperContext framework .....                           | 45  |
| 4.2 The interpretation of node N-1.....  | 50  |
| 4.3 A different interpretation of node N-1.....                                    | 50  |
| 4.4 Node N-2 interpreted in the context bottom .....                               | 52  |
| 4.5 The Structure and Object Layers and their interaction with the outside world.. | 53  |
| 4.6 A context session without the user's visit to N3 in the context (N2, L2).....  | 57  |
| 5.1 Partial implementation model of HyperContext.....                              | 65  |
| 6.1 Siblings .....   | 101 |
| 7.1 The HyperContext user interface .....  | 112 |
| 7.2 The HyperContext Control Panel.....  | 112 |
| 7.3 Recommended links in an interpreted document are shown as HCTr buttons         | 114 |
| 7.4 The "See Also" References floating window.....                                 | 116 |
| 7.5 The HyperContext Status window .....   | 116 |
| 7.6 Implementation model of the HyperContext prototype.....                        | 117 |
| 7.7 A typical conversation between a HyperContext client and server.....           | 121 |
| 7.8 An HTML file fragment before (a) and after (b) it is interpreted.....          | 122 |
| 7.9 Example TIR results.....   | 123 |
| 7.10 Example IRC results displayed in the Document Frame.....                      | 125 |
| 8.1 Dividing an HTML file into blocks.....   | 134 |



|  |     |
|--|-----|
| 8.2 A SWISH-E decoded index file structure.....                                    | 136 |
| 9.1 The characters "A" (a) and "E" (b) shown divided into 15 regions (c), (d) ..   | 143 |
| 9.2 Example nodes and links to letters interpreted in context .....                | 144 |
| 9.3 "t" is accessible from "c" only when "c" is accessed in the context of "a".... | 145 |
| 9.4 The weights for "-" (value 3 in the numeric string representation) .....       | 146 |
| 9.5 Results in raw figures .....   | 149 |
| 9.6 Results in percentage terms.....   | 149 |
| 9.7 Locating a target node.....  | 154 |
| 9.8 Difference between pure and hybrid depth-first searches.....                   | 155 |
| 9.9 Effect of link deletion during the conversion process .....                    | 160 |