Working Paper Series
ISSN 1170-487X

Link as You Type: Using Key Phrases for Automated Dynamic Link Generation

by Steve Jones

Working Paper 98/16
September 1998

Department of Computer Science
The University of Waikato
Private Bag 3105
Hamilton, New Zealand

Link as You Type: Using Key Phrases for Automated Dynamic Link Generation

Steve Jones

Department of Computer Science
University of Waikato
Private Bag 3105
Hamilton
New Zealand
Tel: +64 7 838 4490
Email: stevej@cs.waikato.ac.nz

ABSTRACT

When documents are collected together from diverse sources they are unlikely to contain useful hypertext links to support browsing amongst them. For large collections of thousands of documents it is prohibitively resource intensive to manually insert links into each document. Users of such collections may wish to relate documents within them to text that they are themselves generating. This process, often involving keyword searching, distracts from the authoring process and results in material related to query terms but not necessarily to the author's document. Query terms that are effective in one collection might not be so in another. We have developed Phrasier, a system that integrates authoring (of text and hyperlinks), browsing, querying and reading in support of information retrieval activities. Phrasier exploits key phrases which are automatically extracted from documents in a collection, and uses them as link anchors and to identify candidate destinations for hyperlinks. This system suggests links into existing collections for purposes of authoring and retrieval of related information, creates links between documents in a collection and provides supportive document and link overviews.

KEYWORDS

Dynamic hypertext, automated link generation, key phrase extraction, information retrieval, user interface

INTRODUCTION

This paper is concerned with the automatic introduction of hypertext links into a collection of documents where no links previously exisisted. As long ago as 1945 Vannevar Bush suggested in his description of the Memex that associations between digital objects might be automatically created [2]. Digital libraries, and other means of collecting together enormous numbers of electronic documents give this goal more currency than ever. It is clear from our own experience in developing the Zealand Digital Library (NZDL) [19], and the observations of others [10] that source documents for digital library collections are rarely explicitly connected. This is hardly surprising given that it is often difficult for authors to predict which digital collections their documents will end up in. Manual creation of associations or links is impractical, and efforts in semi-automated link creation such as HEFTI [3] will not scale up to process collections of tens of thousands of documents in a reasonable amount of time.

Many information retrieval systems, such as digital libraries, provide full-text indexes which support ranked and Boolean key word searching. Although this is useful for querying, it requires users to form and specify their information needs in highly constrained ways. Browsing within and between documents is not directly facilitated and readers are unable to naturally follow trails through related material. Hence, the automated introduction of links is necessary to support the browsing activities of readers. Techniques to achieve this have focussed on generating links that are internal to a well-defined collection, and do so off-line prior to browsing [8, 11]. Links must be updated when new

documents are added, potentially impacting upon all documents in the collection.

Another approach is to generate links dynamically at the time of browsing. In this approach hypertext and information retrieval are integrated. Links are emulated by issuing queries containing text selected by readers, and using the most relevant items returned from the retrieval engine as candidate destinations [7]. This approach obviates the need for time-consuming manual link generation. It has a further advantage which we exploit in support of the writing task.

Use of digital libaries is often related to a writing task—authors wish to find related material to inform and contextualise their work. Although efforts have been made to blend reading, skimming, browsing and searching in interfaces to digital libraries [15], little has been done to more closely integrate the writing activity. We are interested in using query-mediated linking techniques to dynamically introduce hypertext links into an author's text. Ideally links to related material would be introduced as text is entered, blurring the distinction between writing and searching for related material, and removing disruptive transfer between writing and search activities.

A specific issue is the how appropriate link anchors within documents, and destinations available from those anchors can be identified. In common with other query based linking techniques [1, 7, 15] we use standard information retrieval measures of similarity between selected link anchors (a query) and documents in the collection to identify candidate destinations. We also use a statistical approach to identify important terms in documents to serve as link anchors. However, our work differs from that of others in that it exploits *key phrase* indexes rather than full-text indexes to determine statistical similarity. We believe that key phrases that concisely reflect the topic of a document can help to create effective links. Our approach moves towards the notion of "concept linking" presented by Cleary [4]

Our motivating examples are reference collections from the New Zealand Digital Library. One example collection, the Computer Science Technical Reports (CSTR) contains more than 40,000 technical research reports by tens of thousands of authors from hundreds of publicly-accessible repositories around the world. This collection and most of the twenty or so other NZDL collections contain no embedded links. Manual creation and maintenance of links between documents in the CSTR and within similar collections is precluded by the enormous amount of time and effort that would be required—there are several gigabytes of text.

In this paper we describe Phrasier, a system for dynamic automated generation of links between a source text and documents within a collection. The paper is organised as follows. In the next section we give an overview of the facilities provided by Phrasier. This is followed by a description of our use of automatically identified key phrases as the basis for generation of link anchors and destinations. We then describe the Phrasier user interface and software architecture. Finally we present directions for future work.

PHRASIER OVERVIEW

Phrasier offers support for authors and readers of documents by utilising automatically identified key phrases within document text as link anchors to related material in document collections of a digital library. Figure 1 shows the main panes of the Phrasier interface.

For authors using Phrasier, the act of writing is also the act of link creation. Text entered into Phrasier by users is analysed, and anchors are associated with key phrases in that text, which then link to related documents. In this case a document which is external to a document collection is linked into the collection. Lists of link anchor texts can then serve as effective document summaries to feed back into the writing process. Linked related documents can be viewed to inform the writing process. Key phrase identification and link insertion happen in real-time with no disturbance to the writing process.

Authors may wish to browse material from more than one collection. With key word searching it is difficult to determine if key words that are useful in one collection will be useful in another. Target documents for links from an author's text are likely to change when linking to different collections, as are the candidate anchors within the text. This introduces a requirement for collection-sensitive link generation. Phrasier supports linking of a document into multiple existing document collections. To do this a key phrase index is required for each collection, which forms the core of a collection server. These indexes are generated using an automated key phrase extraction process.

Phrasier blends the notions of querying and linking to generate multiple destinations for each link anchor. Ranked retrieval is carried out on key phrases to identify candidate destinations, and to offer alternatives to the readers of documents. Ranking of potential

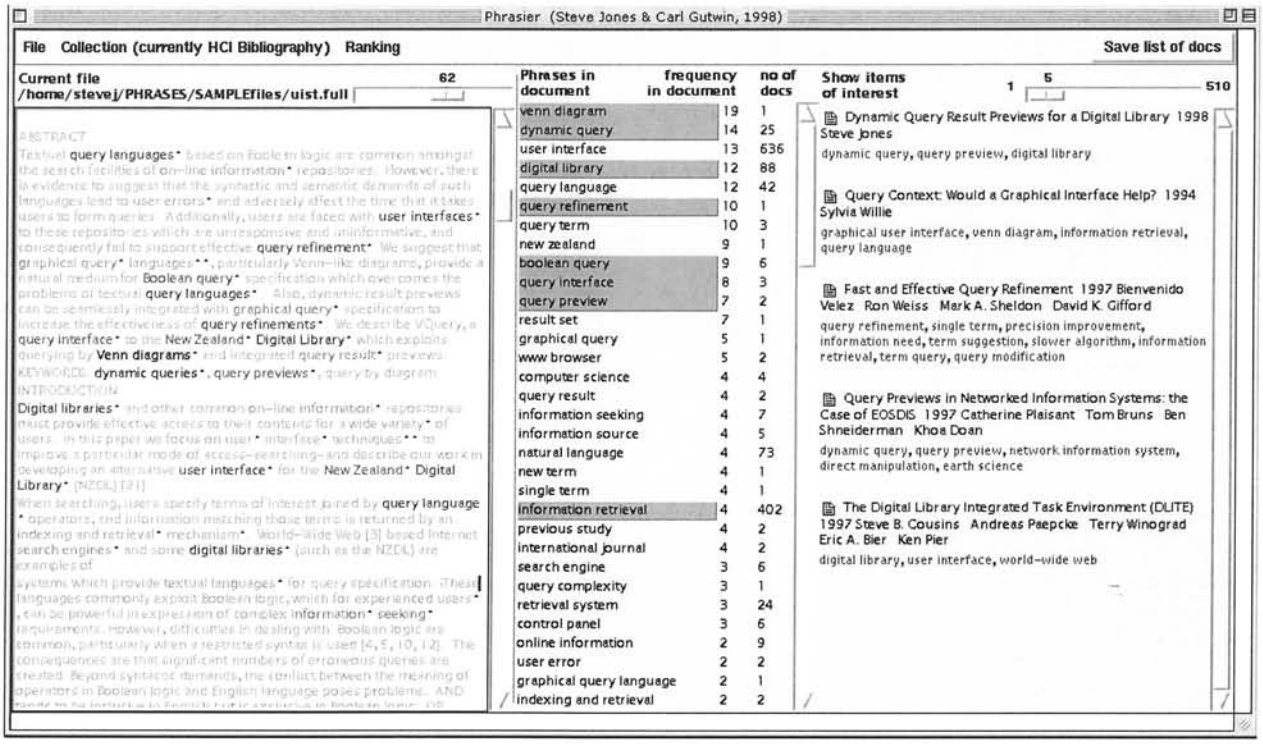

Figure 1: The main window of the Phrasier user interface. To the left is the document pane, in the middle is the key phrase pane, and to the right is the document summary pane.

destinations also offers the reader feedback about the relevance of each destination document.

Documents can be saved from Phrasier to persistent storage at any time in a Phrasier dependent format. Because links can be created to multiple collections from one document it is important to store the document and the sets of links within it independently. Each set of links is relative to a particular collection. One issue is how much information should be stored about each link. At one extreme we may store no link information at all! This means that when the document is read into Phrasier, link identification is relative to the current state of the collection key phrase index—changes to the collection since the document was last accessed are seamlessly integrated. At the other extreme we might store everything possible: location within the document, anchor text, lists of related documents, frequency within the document and so on. If the collection has changed, or the document has been amended in another application, the link information will be inaccurate. Both approaches are useful. The first is most suited to use when collections and documents are subject to change. The approach serves to give the document permanence, perhaps for use within a static hypertext.

We are currently considering how documents and links can be best represented in a portable standard format such as the XML Linking Language [21]

KEY PHRASE IDENTIFICATION

Some types of document (such as this one) contain key words and phrases specified by the author. These concisely summarise the subject matter of the document and have several uses, including classification or clustering of documents, and helping readers to quickly determine if a document is of interest to them. Unfortunately many documents, even scientific papers, do not contain key words or phrases. For example, of 15,000 or so entries in the HCI Bibliography fewer that 5,000 contain author specified key words or phrases. The source documents for the CSTR collection of the NZDL are in Postscript format, from which plain text is extracted for indexing. Author specified key words or phrases in these documents are not immediately identifiable. It would be useful to identify key words or phrases to support clustering, browsing (including linking) and retrieval, yet to do so manually is infeasible. To this end, the Phrase Research Group at the University of Waikato has developed automated techniques for key phrase identification [20]. In addition

to their use in link generation, the resulting key phrases have been used in a novel topic browsing interface [12]

The approach adopted applies machine learning techniques to the problem in a three stages process. The first stage entails extraction of candidate phrases and the number of documents that they occur in within a sample corpus of text. The sample corpus is a subset of the document collection for which phrases are to be identified. The maximum and minimum length of candidate phrases can be specified. The document frequency of candidate phrases is required in the next two stages.

The second stage is training. Training involves the use of an appropriate set of sample documents where key phrases have already been assigned—this might be done manually or use available author specified words and phrases. A "model" is derived for key phrase identification. The Naïve Bayes machine learning technique is presently used. Three attributes of candidate phrases are important within the model. Location within the document is used because more important phrases occur in titles, abstracts and introductory text. An entropy measure indicates whether a phrase, or a longer phrase in which it occurs should be used. A tf•idf measure reflects the frequency (or importance) of a phrase within a particular document, with respect to other documents. A phrase that occurs frequently in a given document but infrequently in others is likely to be useful.

Identification is the final stage, where the model is applied in turn to each document in a collection. The output for each document is a list of key phrases that are allocated to it and a measure for each phrase of the likelihood that it is a key phrase of the document. The number of generated phrases, their minimum and maximum length, and the probability threshold that candidate phrases must reach to be acceptable are all parameters to the process. The key phrases can then be stored in a key phrase index of a similar nature to a standard term index—a list of document identifiers is associated with each key phrase. However, the key phrase index is much smaller that a full-text index for the same collection, and its entries more accurately reflect the topics of documents.

This approach optimises identification for particular document collections, as opposed to others which try to provide general applicability. Initial evaluation indicates that it performs at least a well as state-of-the-art commercial systems such as Extractor [17].

UTILISING KEY PHRASES FOR LINKING

Automated link generation processes must identify relationships between documents which merit supportive links, and link anchors within those documents. Conventionally, information retrieval or browsing systems exploit indexes of terms within documents. Full-text retrieval systems (the NZDL uses MG [18]) typically create indexes which contain all terms from all documents within a collection (with the exception of commonly occurring but otherwise uninteresting 'stop-words' such as 'the', 'and', 'of' and so on). Indexed terms that are common across documents might identify relationships, and therefore may be candidate link anchors.

However, such indexes do not reflect how representative any given term is of a document (beyond simple frequency measures such as tf•idf). As such they are of restricted use in crafting term-based links between documents either statically prior to browsing, or dynamically during browsing. Attempts to do so have essentially utilised ranked querying to emulate linking: terms and some surrounding document context are used to issue ranked queries, with the most relevant documents used as targets for the notional link anchors [7]. In this approach, any selection in any document can be issued as a query and given the behaviour of a link by then displaying the most related document or documents. However, although all text components can act as link anchors, not all necessarily have sensible targets to link to. The quality of the link is unpredictable. Shneiderman et al observe the precision that phrases bring to queries [16]. This, along with the concise summarisation of document focus provided by key phrases, leads us to use them in link identification.

Key Phrase Relationships and Anchors

The key phrase index can be used to identify relationships between documents. Most simply we might assume that there is similarity between the items in the document list for a given key phrase and therefore potential for creating links between them. More realistically we are interested in the degree of overlap between sets of key phrases for given documents. Standard similarity assessments using the vector space model and the cosine measure can be applied to documents using key phrases rather than terms. Similarities can be computed ahead of time, prior to access to the documents. Degrees of similarity (or rankings) are produced allowing thresholds to be set which determine how similar documents must be before they should be linked. This takes much less time than when utilising full-text indexes, but is still highly resource intensive. Such an approach might be useful

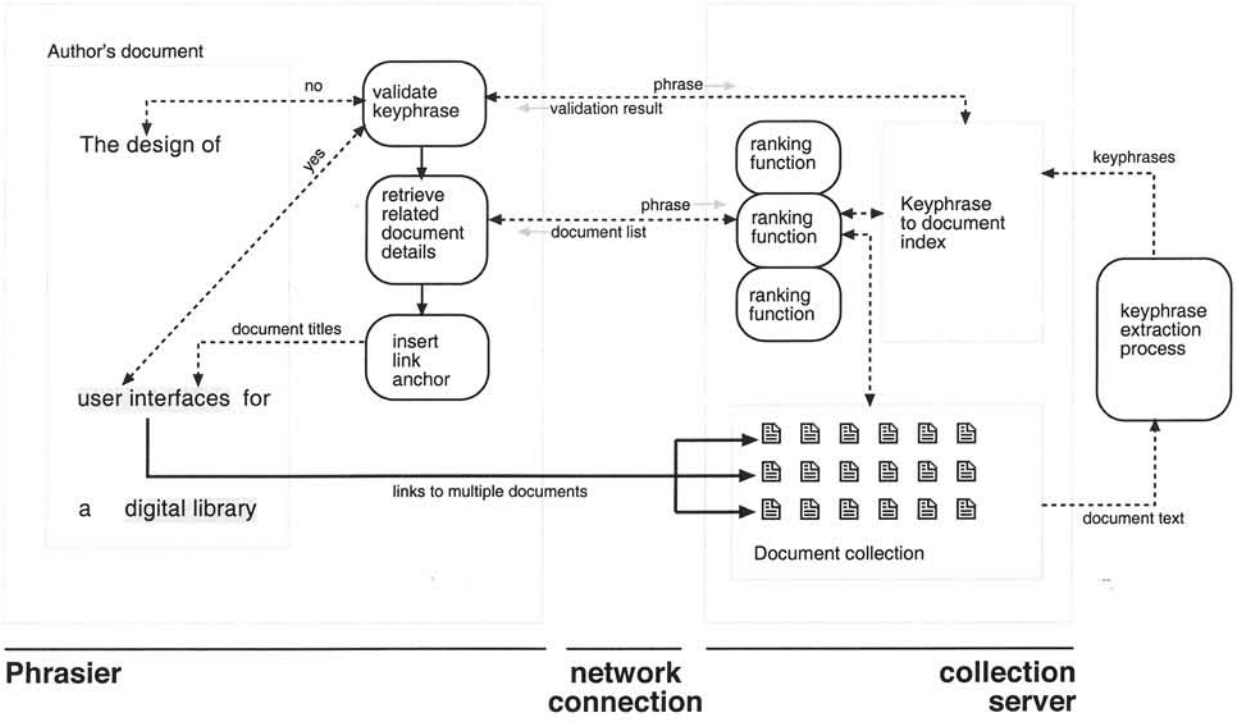

Figure 2: The relationship between Phrasier and a collection server, highlighting the automated link identification process.

for one-time-only identification of link candidates in a static hypertext.

This approach focuses on creating links within a group of related documents where the granularity of linked objects is at the document level. It is also desirable to provide linking at a finer level, from segments of document text to other documents. Key phrases from the index can also serve as candidates for link anchors in a given document. These key phrase anchors will point to only those documents to which the key phrase has been allocated (it is representative of), not just to a list of documents in which it appears as is the case with the term-based model. The distinction between occurrence in, and allocation to, is an important one that distinguishes the key phrase index from a term-based index. In addition, the list of potential target documents for each key phrase anchor is known, is a subset of the entire collection, and can be ranked using "on-the-fly" similarity measures.

PHRASIER

Phrasier is used in the context of a document collection within which users wish to browse, or to which they wish to link one or more existing documents. It uses a key phrase index and key phrase based document similarity measures to automatically generate links in real time. The Phrasier user interface consists of three panes: the document pane, a key phrases pane, and document summary pane (shown together in Figure 1).

Link Creation "On The Fly"

The *document pane* enables users to enter the text of a document that they are authoring, or to load a previously created document from disk for editing or viewing. Standard editing functions are provided. As a user enters text, or it is read from a file, a background process that handles link insertion (shown in Figure 2) compares a window of text (a phrase) either side of the insertion point to the key phrase index. If the phrase is in the index (it has been allocated to at least one document in the collection) two actions are taken. First the phrase is highlighted within the text by changing its colour. Second, the phrase is issued as a ranked query to a retrieval engine. The retrieval engine utilises the key phrase index to identify candidate targets within the document collection and ranks the candidates, returning them to Phrasier. A link anchor is then inserted into the text for the key phrase, pointing to the set of related documents. The continuous analysis of text entered by the user in order to submit queries to, and retrieve

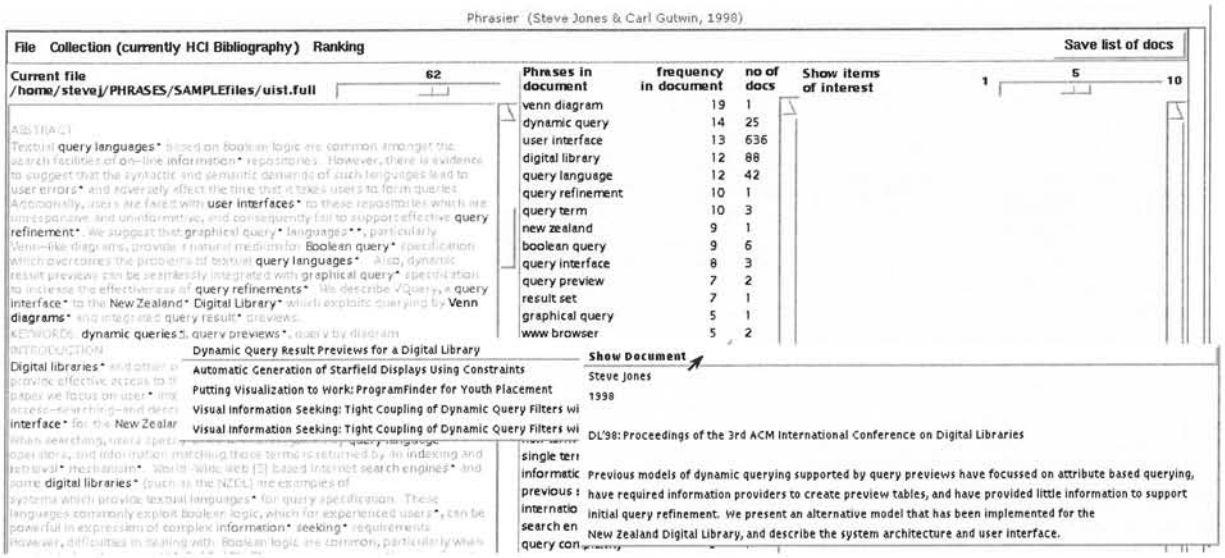

Figure 3: A selected link anchor with five candidate destinations. There is a two-level gloss for each destination document.

information from a retrieval engine is similar to the approach adopted in the Remembrance Agent [14]

Gloss. Each anchor has a *gloss* [23] associated with it. Glosses provide an indication of the nature of the destination content, the need for which has been discussed by Furnas (referred to as residues or scents) [6]. In Phrasier this takes the form of a two-level pop-up menu. The first level contains the titles of the target documents, presented in an ordered list with the most relevant at the top. The number of destinations to be included for each anchor defaults to a maximum of five, but can be specified by the user. The user can choose to pop up a secondary menu for each destination. This contains additional information about the destination document. For the HCI Bibliography collection of the NZDL, readily available metadata is combined to provide this detail (see Figure 3). For the CSTR collection short summaries extracted from the beginning of documents are shown. The glosses allow users to partly judge the utility of link destinations, using information derived from the destinations without the cost of a link traversal. When a user selects an item from the menu the associated document is displayed in a new window, and can be read into Phrasier for key phrase identification and link generation.

Link highlighting. The user can also toggle whether link anchors are displayed or not. Further, (as shown in Figures 1 and 3) key phrase anchors can be displayed in a range of grey levels. This presentation technique extends that used in XLibris [15] The grey levels correspond to a measure of importance of a phrase within the document text. Additionally a slider control allows the user to manipulate the contrast between key phrase anchors and the text of the document to emphasise or hide link anchors. At one extreme the anchors are indistinguishable from the rest of the document text, at the other only anchor text is visible. The link anchors can be used as "skimming" aids, to enable readers to rapidly consider key topics in the documents.

Multiple anchors, multiple destinations. The document pane enables the author or reader to move beyond the "one anchor, one target" model commonly used in linked structures such as the Web, by providing multiple targets where appropriate. It also supports "multiple anchors, multiple targets". Using a keyboard modifier multiple key phrase anchors can be selected in the document pane, and once selection is complete a list combining destinations for each link is formed and displayed.

Summary Overviews

As key phrases are identified in the text they are inserted into the *key phrases pane* (shown on the left of Figure 4). This contains a list of key phrase link anchors that have been identified within the document pane. Two items of information are displayed for each key phrase: the frequency with which each key phrase occurs in the text, and the number of the documents within the collection to which each key phrase has been allocated. This list therefore shows the link anchors

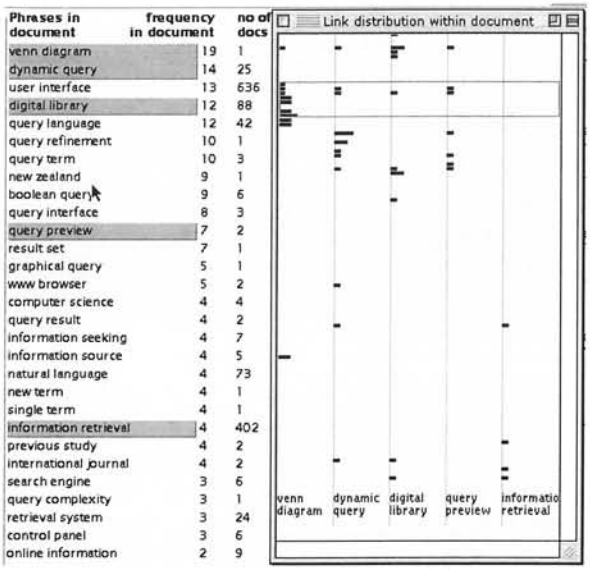

Figure 4: Summary overview information. To the left is the key phrase anchor list, and to the right is the link overview pane representing the distribution of links throughout the document.

present in the document. It can be sorted in three ways: alphabetically by key phrase, by descending frequency of each key phrase within the text, and by descending frequency of documents in the collection to which each key phrase has been allocated. Key phrase anchors in the list can be selected individually, in contiguous blocks, or in multiple disjoint blocks. Selections behave as they do in the document pane—the most relevant target document is retrieved and displayed, or alternatively a list of target documents is displayed.

The *link overview pane* (shown to the right of Figure 4) provides an overview of link locations for currently selected phrases within the document. The top of the pane reflects the start of the document and the bottom the end of the document. The document is mapped on to one hundred segments within this region. The black rectangles displayed in the region above the phrase labels show two attributes of each phrase anchor: the location with respect to the start and end of the document (percentage of the way through), and the frequency within each segment. This is similar in nature to the TileBars system [9]. Clustering of phrase anchors and consequently changes in focus of topic within the document become evident from this representation. A coloured rectangle that spans the width of the overview pane reflects the portion of the document text that is currently visible in the document pane. As text is scrolled in the document pane the position of this

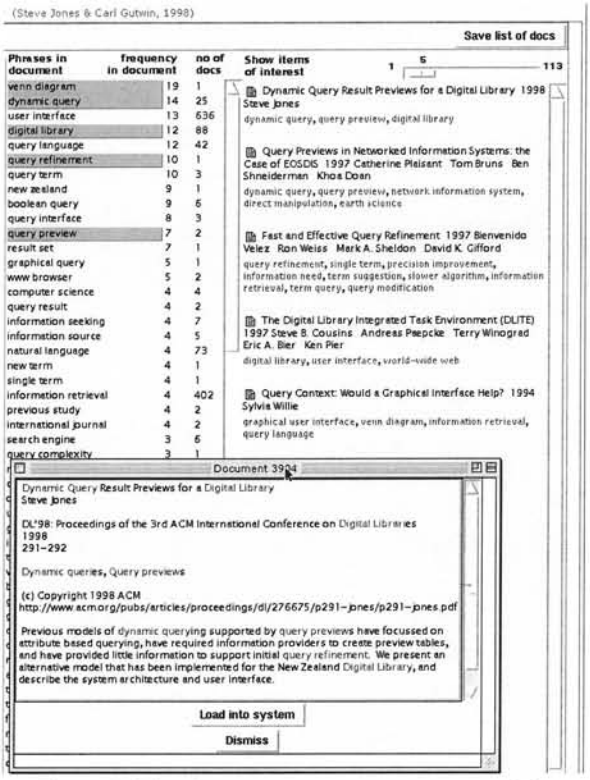

Figure 5: The document summary pane, and the text of a document selected for viewing via a link.

rectangle is updated. Dragging the rectangle vertically using the mouse results in scrolling of the text within the document pane, facilitating rapid movement to target portions of the document. Additionally each black rectangle responds to mouse selection, and the document pane text jumps to the location of the selected link anchor.

Document Display

The *document summary pane* (shown in Figure 5) displays summary detail of documents from within the collection which are related to sets of user selected key phrases. These phrases may have been selected in either the document or key phrase panes. When bibliographic information is available, some items (such as title, author, year, and keywords) are displayed in the summary pane. This is augmented by display of the key phrases that have been allocated to each of the documents by our key phrase extraction process. The number of document summaries displayed can be set by the user via a slider control. Currently displayed summary information can be saved to disk for later consideration. The text of a document can be retrieved and viewed by selection of its link anchor in this pane.

Initially the text of this document is displayed in a separate window to enable viewing of the author's text at the same time. If required the text of the document can then be read into Phrasier to continue the browsing process.

Implementation

The interface of Phrasier is implemented in Tcl/Tk and runs on systems with standard Tcl/Tk installations. In early versions the system was stand-alone, and needed to load large phrase indexes when starting up, which was slow and not practically portable. Calculation of similarity measures was also slow. We have moved to a client-server model, where the interface is still implemented in Tcl/Tk, but a Perl server process handles the key phrase and collection data more efficiently. Therefore the interface is more portable, data is centralised and servers can be tailored to the requirements of particular collections. The drawback is that this approach incurs the costs of transferring data over the network.

FUTURE WORK

We are currently undertaking a study to determine how effectively Phrasier identifies material related to an author's document. In this study authors are supplying published papers which are read into Phrasier to generate several ranked list of related documents (a variety of ranking algorithms utilising key phrases are being applied). The HCI Bibliography collection is being used in the study. Related documents returned by a term-based retrieval are also recorded. Authors are asked to rank each returned document for relevance, and to give an overall assessment of each ranked list. This will enable us to compare key phrase based retrieval against standard term based retrieval, and to identify the most effective of our set of ranking alternatives. At this stage our approach is to gather data from real potential users rather than to exploit data from an existing corpus, because of the opportunity for further interaction and investigation with the subjects. In the future we may use existing collections and data sets (such as TREC) for evaluation purposes, and will also carry out usability evaluations of Phrasier. Informal analysis indicates that it is indeed useful in identifying related material (to the extent that one user believed that Phrasier had merely identified the reference list from their paper and presented it as the related document list—of course, it had not).

The destinations of links within Phrasier are currently documents. Once users arrive at a destination document they must search through it to find the sections most related to the key phrase anchor that they selected to get there. We plan to use segments of document text as link destinations. As can be seen in the link overview we can identify clustering of sets of key phrases within a document, and such a visualisation might be useful in the gloss of a link anchor to allow the user to control where in a document they traverse to.

Phrasier currently connects to a single collection at a time. Users can easily switch between collections, but are unable to view links to multiple collections at the same time. Connecting to multiple collections at the same time is trivial, but raises several issues. In the interface link anchors must distinguish which collection they link into, and there may be so many anchors that the text is obscured (the ability to switch anchors off and grey level display of links provides some help here). Related document lists must distinguish from which collection the documents were sourced. Integrating the results of ranked retrievals from multiple collections is a particular problem [5, 22], as rankings are collection dependent. These are issues to be resolved in our future work.

ACKNOWLEDGMENTS

Many thanks to Carl Gutwin, who collaborated on the design and initial implementation of Phrasier. Thanks to Gene Golovchinsky for very useful comments on Phrasier and early drafts of this paper, and to the members of the Phrase Research Group at the University of Waikato.

REFERENCES

1. Allan, J. Automatic Hypertext Link Typing. In *Proceedings of ACM Hypertext'96* (Washington, DC, USA, March, 1996), ACM Press, 42-52.

2. Bush, V. As We May Think. *Atlantic Monthly*, 176, 1 (1945), 101-108.

3. Chignell, M.H., Nordhausen, B., Valdez, F. and Waterworth, J.A. The HEFTI Model of Text to Hypertext Conversion. *Hypermedia*, 3, 3 (1991), 187-205.

4. Cleary, C. and Bareiss, R. Practical Methods for Automatically Generating Typed Links. In *Proceedings of ACM Hypertext'96* (Washington, DC, USA, March, 1996), ACM Press, 31-41.

5. de Kretser, O., Moffat, A., Shimmin, T. and Zobel, J. Methodologies for Distributed Information Retrieval. In *Proceedings of the 18th International Conference on Distributed Computing Systems*

(Amsterdam, Netherlands, May, 1998), IEEE Computer Society, 66-73.

6. Furnas, G.W. Effective View Navigation. In *Proceedings of CHI'97: ACM Conference on Human Factors in Computing Systems* (Atlanta, GA, USA, March, 1997), ACM Press, 367-374.

7. Golovchinsky, G. What the Query Told the Link: the Integration of Hypertext and Information Retrieval. In *Proceedings of ACM Hypertext'97* (Southampton, UK, April, 19967), ACM Press, 67-74.

8. Green, S.J. Automated Link Generation: Can We Do Better Than Term Repetition? *Computer Networks and ISDN Systems*, 30, 1-7 (1998), 75-84.

9. Hearst, M. TileBars: Visualization of Term Distribution Information in Full Text Information Access. In *Proceedings of CHI'95: ACM Conference on Human Factors in Computing Systems* (Denver, CO, 1995), ACM Press, 59-66.

10. Kellogg, R.B. and Subhas, M. Text to Hypertext: can Clustering Solve the Problem in Digital Libraries? In *Proceedings of ACM Digital Libraries'96* (Bethesda, MD, USA, March 1996), ACM Press, 144-150.

11. Myka, A. and Guntzer, U. Automatic Hypertext Conversion of Paper Document Collections. In *Digital Libraries Workship: DL'94* (Newark, NJ, USA, 1994), Springer-Verlag, 65-90.

12. Nevill-Manning, C.G., Wittent, I.H., and Paynter, G.W. Browsing in Digital Libraries: A Phrase-based Approach. In *Proceedings of ACM Digital Libraries'97*. (1997), 230-236.

13. Price, M.N., Golovchinsky, G. and Schilit, B.N. Linking by Inking: Trailblazing in a Paper-like Hypertext. In *Proceedings of ACM Hypertext'98* (Pittsburgh, PA, USA, June 1998), ACM Press, 30-39.

14. Rhodes, B.J. and Starner, T. Remembrance Agent: a Continuously Running Automated Information Retrieval System. In *Acquisition, Learning and Demonstration: Automating Tasks for Users. Papers from the 1996 AAAI Symposium* (Stanford, CA, USA; 25-27 March 1996) AAAI Press, 122-5.

15. Schilit, B.N., Price, M.N. and Golovchinsky, G. Digital Library Information Appliances. In *Proceedings of ACM Digital Libraries'98* (Pittsburgh, PA, USA, 1998), ACM Press, 217-226.

16. Shneiderman, B., Byrd, D. and Croft, W.B. Sorting Out Searching. A User-Interface Framework for Text Searches. *Communications of the ACM*, 41, 4 (April 1998), 95-98.

17. Turney, P. Extraction of Keyphrases From Text: Evaluation of Four Algorithms. *Technical Report NRC/ERB-1051, National Research Council of Canada*, 1997.

18. Witten, I.H., Moffat, A. and Bell, T.C. *Managing Gigabytes: Compressing and Indexing Documents and Images*. Van Nostrand Reinhold, 1994.

19. Witten, I.H., Nevill-Manning, C., McNab, R. and Cunningham, S.J. A Public Library Based on Full-text Retrieval. *Communications of the ACM*, 41, 4 (April 1998), 71-75.

20. Witten, I.H., Frank, E., Paynter, G.W., Gutwin, C., Nevill-Manning, C. and Jones, S. Using Machine Learning to Extract Key Phrases From Text. *Working Paper, Department of Computer Science, University of Waikato, Hamilton, New Zealand.*, 1998.

21. XML Linking Language (Xlink). World Wide Web Consortium Working Draft WD-xlink-19980303. *http://www.w3.org/TR/WD-xlink.*

22. Xu, J. and Callan, J. Effective Retrieval with Distributed Collections. In *Proceedings of 21st Annual Conference on Research and Development in Information Retrieval* (Melbourne, Australia, August, 1998), ACM Press, 112-120.

23. Zellweger, P.T., Chang, B. and Mackinlay, J.D. Fluid Links for Informed and Incremental Link Transitions. In *Proceedings of ACM Hypertext'98* (Pittsburgh, PA, USA, June 1998), ACM Press, 50-57.