ELECTRONIC PUBLISHING, VOL . 2(3), 143–156 (OCTOBER 1989)

VORTEXT: The hard-backed screen

VICTORIA A. BURRILL* AND JOHN A. OGDEN

Department of Computer Science
The University, Whiteknights
PO Box 220, Reading RG6 2AX, UK

SUMMARY

With the current information explosion in the number of books and periodicals published annually coupled with the decreasing costs and availability of wordprocessors, it is authors, not publishers, who are becoming the main controllers of a document.

If a document is written using a computer then it obviously makes sense for it to be read on the same medium. But how will the ordinary man-in-the-street react to this? How will he react to sitting down at a terminal instead of browsing through bookshelves? How will he react to scanning a screenful of text rather than feeling the 'physicalness' of a real book? What facilities will he expect? What facilities will he want?

This paper is the result of three and a half years research using VORTEXT — VictORias TEXT reading system — a unique interface which begins to explore the limits, possibilities (and pitfalls!) of extending the 'real book' metaphor across from its traditional paper medium to its future computerized form.

KEY WORDS Hypertext Electronic books

INTRODUCTION

The information explosion has already begun. Vast numbers of books are published annually: in 1986 the British Lending Library housed nearly 3 000 000 serials and monographs with an annual increase of 102 000 titles [1]; the Oxford University Press published 21 000 titles [2], the Cambridge University Press more than 13 000 [3].

Simultaneously, with the increased availability of microcomputers and desktop publishing packages there is a definite trend towards the use of wordprocessors in the authoring of books. Many publishing houses such as Blackwells [4], IPC and John Wiley already issue guidelines concerning the submission of text not only on floppy discs but also by electronic mail. This trend is unlikely to diminish in the future.

As the writing, production and dissemination of books is increasingly handled by computer, it is perhaps only a matter of time before books are also read using computers. Rather than searching through a bookshelf, readers will sit down at a computer screen and select a book via the screen. Rather than feeling the 'physicalness' of the book, readers will be given other clues as to its quality and size. The contents and index pages will be there of course and each book can still be read page by page, but now these traditional paper features have been enhanced by the computer. Readers will be able to browse through the book in many different ways according to the type of information they are looking for and the type of information that the book provides. Computers are

*Now at: Informatics Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK
0894–3982/89/030143–14$07.00 *Received 17 November 1988*
 Revised 9 September 1989

also the ideal medium in which to incorporate dynamic features: the presentation of the book can easily be modified to suit a partially sighted person, for example, or writeable overlays may be used so that a reader may make notes in the margin without fear of defacing the book.

Of course the computerized version of a book is unlikely to replace the paper version — somehow the thought of snuggling under the bedcovers with a warm disc drive just doesn't have the right ring to it, and trying to read in the bath with power cables or batteries in the vicinity is still lethal! Therefore, rather than reading the entire book from the screen, perhaps most people would determine which sections are of interest to them and then request a hard-copy of just those pages. The computer would handle the copyright and printing fees and readers would be able to take home their own personal paper copy of the pages they requested.

But no matter how inevitable computerized books are, nobody will read them if they are more difficult to handle than the paper version, both physically and with respect to the book interface itself.

The most usual form of user input to a computer is the keyboard, but this is unwieldy and many people are unfamiliar with typing. Mouse-like devices are a possibility, but often require some tuition before they can be used effectively. The most intuitive input device is the human finger (or suitable stylus) used in conjunction with a touch screen, but this in turn implies firstly, a horizontal screen in order to avoid arm fatigue, and secondly, an interface capable of accurately resolving to the size of a human finger.

Perhaps the most obvious model for the interface to a computerized book is its paper counterpart. Two questions now arise: firstly, how much *should* a computerized book look like the paper version, and secondly, how much *could* a computerized book look like the paper version?

The overall philosophy behind VORTEXT is, firstly, to determine which characteristics of a real book people use and feel comfortable with, and secondly, to explore ways in which these characteristics may be realized on the screen. This is a user-orientated approach which is aimed at the man-in-the-street; it is not (necessarily) a good implementation of hypertext as such, but does incorporate some hypertext-like features into its design where appropriate. As far as the authors are aware, the only other system based upon this philosophy is the one being developed at the University of York [5], although a few systems incorporate isolated book-like features, such as the page turning mechanism in Dataland [6].

The advantage of deliberately presenting a familiar interface such as this is obvious. Providing that users have read real books (a not unreasonable assumption), then they are already equipped with a strong mental model of the computerized version and can use at least the basic features with the minimum of training. As readers gain experience with the system they can begin to incorporate some of the less book-like and more powerful computer-orientated features into their work as well. The first version of VORTEXT was evaluated using eight 'real' readers performing a given set of tasks [7]. Although such a small sample of people cannot be used to provide conclusive evidence, the uniform results of the trials proved sufficient to indicate that readers did indeed follow this general pattern of behaviour. As a form of debriefing, they then answered some general questions about VORTEXT. Throughout, reactions to the system were very positive: people consistently referred to 'the book' and used paper rather than computer terminology in their comments. This would seem to support the idea that the philosophy

behind VORTEXT is appropriate to this type of situation, and that the interface to VORTEXT agrees fairly closely with the user's mental model.

In order to determine how much a computerized book could look like a real book, we first discuss ways in which people read paper books. The support (or otherwise) of these methods when applied to computerized books is then explored with particular reference to ideas behind the VORTEXT system.

HOW DO PEOPLE READ BOOKS?

Anyone who has ever read a book must, by definition, know how they did it. For most people reading has become such a subconscious process that to ask them to explain what they are doing is almost impossible — rather like trying to describe how to tie shoe laces. From observations and common sense some behavioural patterns and methods do emerge however.

If the reader is looking along an unfamiliar bookshelf for a book he has never seen before, the search for that book is very deliberate. The reader walks alongside the bookshelf, head on one side to read the vertical titles on the spines, occasionally consulting a piece of paper on which the reference for the book has been written. Once the reader has found the required book he may reject it without even taking it off the shelf: "How am I going to carry that home on my bike?" or "I want an in-depth study not a childrens' ABC". If the book appears to be of the required type then the reader may take it from the shelf, open it at random, thumb through looking at the pictures, occasionally stopping to read something that has caught his eye. By doing this, the reader is gaining a subjective impression of the book and hence establishing whether or not it is likely to fulfil his needs. A book with the 'wrong' appearance may be rejected on that basis alone — imagine a novel in multi-column newspaper format for example, or a modern journal in a blackletter font.

If the reader requires not a general coverage of the subject but the answer to a very specific question, then the search process depends less on appearance and more on the content of the book. The reader first turns to the index in order to try and find a reference to the subject. If this fails, he tries various word and synonym permutations of it. If this also fails then the reader might consult the contents pages, but more often than not the book will be rejected and put back onto the bookshelf.

Contrast this behaviour with that of a reader looking for a familiar book on his own bookshelf. Now he does not need to examine each spine in turn but instead can go directly to the required book using external clues only: "The big red book with white lettering", "Top shelf near the pot plant" or "The little thin one between the two green ones". Rather than consulting the index and contents pages of the book, sections that are frequently referred-to are also remembered using subjective clues: "A few pages from the end" or "Opposite a large diagram".

So far, only text and reference books have been considered. The behaviour of readers differs when reading, say, a novel. Novels almost never provide an index and most do not provide a contents page since a listing of, say, 'Chapter 1... page 1', 'Chapter 2... page 14' and so on is singularly uninformative. In addition, since a novel is intended to be read sequentially from cover to cover, references to the middle of it seem highly inappropriate. To compensate for the decreased usefulness of the contents and index pages, readers of novels employ alternative techniques to aid the reading process. A

much greater emphasis is now placed on pre-knowledge of the physical organisation of the book. How many times have readers decided whether or not to continue reading a novel purely on the basis of the quality of its opening paragraph? How often do more impatient readers peek at the last few pages of a novel to satisfy their curiosity about *who dunnit* or whether they live happily ever after?

Finally, almost any book may be read in conjunction with some form of personalization. This may take many forms but perhaps the most common are, firstly, marginalia: "I've marked the relevant passages in red" (a practice which is often helpful, but definitely to be discouraged in library books); secondly, association: "It's the page where I spilled my coffee" and thirdly, actuality: "Don't lose the toffee paper — it's keeping my place!".

HOW DO PEOPLE READ COMPUTERIZED BOOKS?

Having established some ways in which people read paper books, the problems, implications and solutions to those methods are now explored within the context of computerized books.

Searching along the book shelf

The gallery in the Domesday System [8] is a good example of the type of visual imagery required in order to give the user a sense of walking through a building. Such imagery could also be applied within the book environment: as the reader walks towards a bookshelf, the detail of the titles on the spines gradually comes into focus. Although this is a very powerful method, it is perhaps rather inappropriate once the book has been opened since the metaphor of walking from page to page is very weak. Perhaps a simpler solution would be to give a 'distance' shot of the entire bookshelf. The colours and sizes of book spines could be displayed quite easily, although the titles are unlikely to be

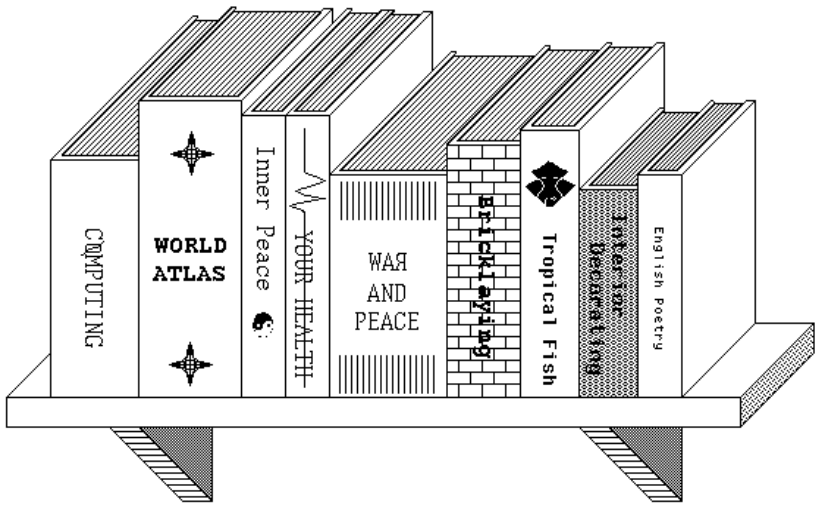

Figure 1. A bookshelf where the font and patterning on the book spine is used to indicate the subject of the book

legible. Images of pot plants may also be added to the shelves as locational clues for the reader to use in re-recognizing a book.

Within VORTEXT, the reader is presented with shelves of books to which various visual clues have been added. Firstly, since it is obviously impossible for the reader to pick up a computerized book, its we can be conveyed by the width of its spine — the greater the width, the greater the number of pages. Secondly, the title on the spine can be presented using a font and/or background pattern that suggests its subject — a blackletter font for a history book for example, a brick background for a book on house construction (Figure 1).

For a bookshelf in which the subjects of the books are many and varied, this latter scheme works very well. If all the books concern one particular subject, however, the scheme fails.

Opening the book

For most people, the book type is established almost immediately upon opening it. Novels tend to consist of large blocks of text with few headings but much indented dialogue. Text books tend to consist of more varied lengths of text with many headings, often interspersed with equations and diagrams. Modern paperbacks rarely include running heads but text books almost always do — the book title on the left-hand page and the book or chapter title on the right-hand page (Figure 2).

The recognition of a book type is a subconscious process and readers become aware of it only when they are irritated by some deviation of the format from the accepted (and therefore expected) standard. For example, people have become accustomed to seeing

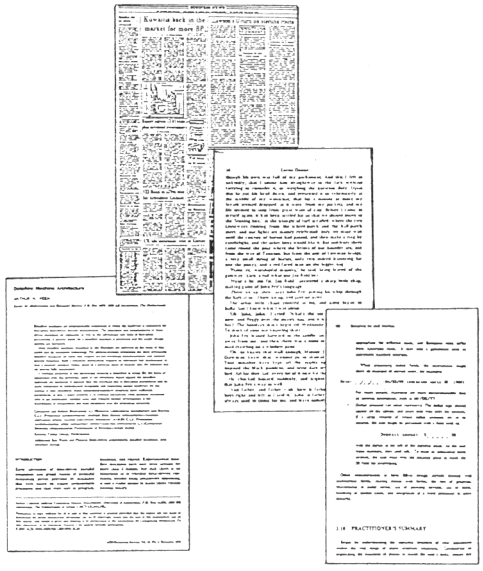

Figure 2. Pages from a newspaper, text book, novel and journal reduced to illegibility but still recognisable as such from their layout alone

fixed-width fonts within the context of typewritten letters, computer listings and household bills, but not within conventionally typeset books. Therefore, every effort should be made to maintain the resemblance of pages on the screen to their paper counterparts. This does present some problems, however.

Firstly, many books follow the Golden Section ratio (21:34 or 1:1.6 width to he). By contrast, most computer screens are landscape rather than portrait orientated, a character-based terminal typically displaying 80 characters by 24 lines (page aspect ratio 1:0.3) and a graphics workstation 1152 by 900 pixels (page aspect ratio 1:0.8). Although many people are as familiar with screens as they are with conventional books, the difference between the paper and screen ratios for use in displaying the pages of a book is considerable. Obviously, the area of the screen that is actually used for display can be clipped to present the usual ratio, but this may result in the useful area being reduced to an unacceptable size.

Secondly, each character on the computer screen is composed from pixels, typically 8 by 14 pixels per character, or 80 pixels per inch for a high resolution graphics screen. By contrast, typesetters can print at up to 5000 pixels per inch, so for any given point size it is very difficult to achieve the same degree of resolution and clarity on the screen as on paper. Most books require one font for the bulk of text, an italic and/or bold version for emphasis and one or two larger versions for headings. There may also be a requirement for incidental fonts such as footnotes or mathematical equations. The limited pixel-based screen character size makes such a variety very difficult to achieve, therefore many screens use only one basic font and rely on whole-character video inversion, high intensity or blinking to provide the variety required. Headings are indicated by the use of white space. Other more sophisticated methods of displaying fonts are available, such as anti-aliasing and the use of outlines, depending on the resolution of the screen and the skill of the designer [9, 10]. However, the calculations involved in the display of such fonts could increase the time taken to display an entire page of text to unacceptable levels.

The contents page and ‘Where am I?’

If a contents page is included in the book, it is presented as a hierarchical list of the chapter, section and subsection headings used. The contents pages give the reader a fairly good overview of the subjects covered in the book, but apart from the page number at which each section starts, no other direct information is given at all. The linear arrangement of headings is very uninformative beyond showing which sections follow which, and if the reader wishes to discover the approximate length of a section and hence the depth to which it is likely to treat the subject, then he has to first make a few mental calculations using the page numbers.

The contents page of a computerized book has a much greater scope for the inclusion of reading clues. For some books, there may be a strong locational element to each heading. For example, the contents page in a geography book about the British Isles might be more usefully displayed with the sections about Wales placed half way down the left-hand side of the page, Scotland placed at the top of the page, London near the bottom right-hand corner and so on. True, if the reader is unfamiliar with the geography of the British Isles then this arrangement will, at first, be as good (or bad) as the conventional, linear layout. Once the reader is familiar with the geography then this

arrangement constantly reinforces his mental map of the subject and hence is much more useful when retrieving information.

One apparent disadvantage of such a scheme is that a single screen is unlikely to be large enough to accommodate all the texts of all the headings in the book. But is it actually necessary for the headings to be displayed together in this way? Would it not be better to reduce the reader's information load by displaying only subsets of the headings? VORTEXT handles this problem by first presenting the reader with chapter headings only. By selecting one of these chapter headings the corresponding section headings within that chapter are displayed. By selecting a section heading, the corresponding subsection headings are displayed and so on, successively refining the scope of the search to as much, or as little, detail as the reader requires. At any point the reader can back-out and try an alternative search path, (and any headings of now rejected sections are removed), or select a specific section in order to turn to that page in the book and read it in full.

If the information is positioned on the contents page according to some known scheme then it seems logical for the reader to use this scheme when retrieving the information. This retrieval can be done in many ways: the information may be clued from the heading alone: "I want to read about Edinburgh so I'll select the chapter on Scotland", or the information may be clued from its position: "I can't remember exactly what the town was called but I do remember that its heading was displayed somewhere in the top section of the screen". Note that for the latter method the appropriate heading need not be currently on display — it is its position which is important, the words are an added bonus [6]. The reader may apply this *Method of Loci* within VORTEXT by selecting a blank area of the contents page. Headings that would normally be positioned in or near that area are then displayed one by one until the reader either selects a heading or moves to another area.

Obviously, once the reader has actually turned to a page in the book the contents page is no longer visible, but the reader still requires orientation information within the subject and within the book as a whole. Conventional paper books provide this information by the use of running heads and in the thickness of the paper itself. Computerized books could use the idea of running heads, but these would occupy one valuable line of screen space and would not solve all the problems anyway.

VORTEXT approaches this problem in a totally unique way by using the idea of closed pages. The text of a VORTEXT page is displayed between two sets of multiple vertical lines, the set to the left representing the thickness of paper of the closed pages of the book before the page currently on display, and the set to the right representing the thickness of closed pages after the page currently on display. Each vertical line pair is used to represent a section within the book and is spaced so as to give some indication of the relative length of that section. Each pair also includes the name of the section it represents and hence may be selected in order to turn directly to that page (Figure 3).

The advantages of this scheme are that the closed pages give the reader constant orientation within the subject of the book, and within the book as a whole. They are also behaving as concrete representational icons — they represent some concrete, physical feature (pages in a book) in addition to being an iconic indication of the action that will result upon selecting them (select the icon in order to initiate the 'turn to page' action).

The disadvantages of this scheme are that, firstly, the number of closed pages displayed at any one time must be restricted so as not to totally obscure the text area. One method in use displays all the chapter headings in the book together with the immediate

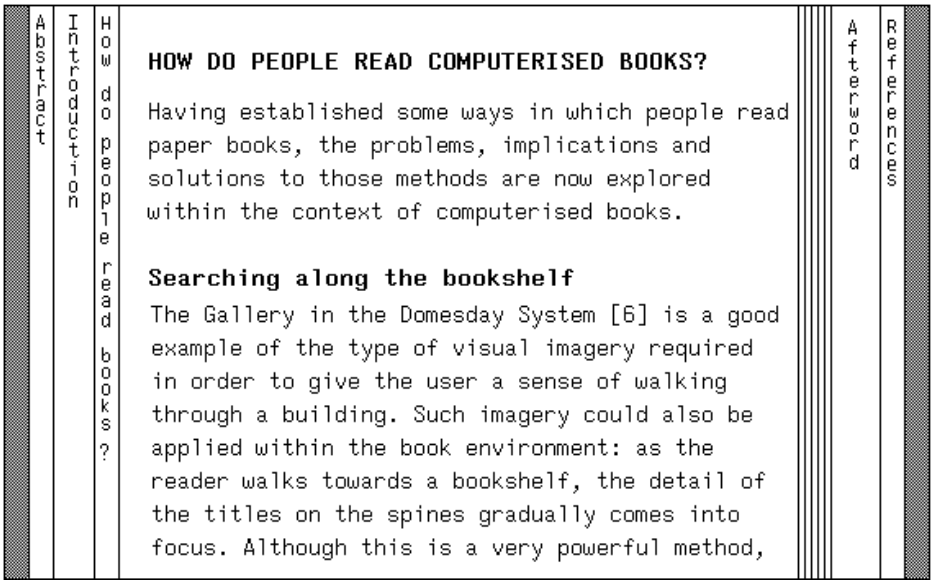

Figure 3. VORTEXT text page showing closed pages

child sections of the current section, the sibling sections of the current section and any parent sections in the hierarchy. This appears to work quite well for most text books. The second disadvantage of the closed pages lies purely in the difficulty of finding an effective method of displaying their heading text. Should the headings run down from the top of the closed page or up from the bottom, or should they be printed normally but with the letters arranged in a column? Should the headings be displayed against a background of vertical line shadings so as to indicate their chapter, section or subsection level, or should they be indented vertically? These questions require further investigation.

For books such as novels, which do not include a contents page, the closed pages can still be used but now they represent only the thickness of paper since it is not appropriate for them to contain any heading text. The user selects some point across the width of the closed pages in order to turn to the page at that proportional distance through the book.

In addition to orientation within the subject of the book and within the book as a whole, a third form of orientation is also required by the reader — that of orientation upon turning from page to page. When the reader of a (non-folio) paper book turns to the next page, the text of the current page is gradually obscured by the turning page from right to left, and the text of the next page is gradually revealed from right to left. The readers attention has time to be drawn to the new point of focus rather than suddenly being presented with a totally new display: visual momentum has been maintained [11]. By contrast, the conventional method of displaying text on a computer screen is from top to bottom, which may be appropriate for ancient scrolls but is not very book-like. It is a relatively simple task for a computer to take screen dumps of each page and then display them in conjunction with the closed pages from right to left as required, but this has significant implications on the software and hardware of the computer which may not always be practical to implement.

Although requiring further development, the concept of the closed pages is extremely simple and has proved to be the most intuitive and successful aspect of VORTEXT to expert and novice users alike.

Is it in the index?

Like designing a font, the creation of an index in a book is a specialist task. It requires a considerable understanding of the subject of the book, of the expectations of the target readers, and uses to which they will put the book. Traditionally, indexes have been created 'by hand' by a professional indexer. Although the final results are usually excellent, there are several disadvantages to this process. Firstly, the finished index is very much a matter of personal taste. No two indexers will index the same book in the same way; indeed, no one indexer is likely to index the same book in the same way twice. Secondly, there is a limit on the number of references any one person can handle. Although the index generally accounts for less than 3% of the total number of pages in the book, an average text book of 300 pages will generate about 300 indexable entries, each of which then has to be cross-referenced or associated with a list of page numbers.

Traditionally, the page numbers in index entries are classified as major, minor or as referring to an illustration. These are indicated by the use of various fonts, usually a bold type for the main reference and a lighter type for the others. The keyword *illus* is sometimes used to indicate a diagram, and *passim*, *ff*, *ll* and so on to indicate passing references. More recently, many books have been published using a much simplified form of this index that tends to include only the main references. Whether this change is for economic reasons or otherwise, it seems a pity that the secondary information that traditional indexes provide has been lost.

Perhaps then, the creation of indexes is a suitable candidate for computerization?

It is a very simple task for a computer to search through the text of a book (in machine readable form) and create a concordance — an exhaustive list of all the different wors in the text together with a list of all the locations in which each of those words occurs. (By contrast, an index is a human-generated list of keywords and phrases used in the book, chosen as being the ones most relevant to that book and its target reader.) Although concordances are very thorough, there are several difficulties associated with using them.

Firstly, stem, case and semantic synonyms have to be identified using a variety of tables and analysis methods, and their respective entries handled appropriately. Entries with multiple definitions according to context should also be treated appropriately.

The second disadvantage of concordances is the number of references given for each entry. In a conventional index, the indexer will ensure that about eight at most of the most relevant references are given. Without careful programming the computer is unable to determine the usefulness of each reference and hence the reader could be presented with dozens, if not hundreds, of references.

Many systems have been built to try to overcome the various problems of computer indexing [12–14]. Although these systems are very successful in their own right, the size of the databases they use and the sheer processing power they require make them impractical for general use in indexing books.

One advantage of concordances over indexes is in the type of reference each entry gives. Entries in a conventional index usually resolve to page numbers. Entries in a concordance could also resolve to page numbers, or even lines, but more usefully

perhaps, could resolve to chapters or sections. This would have the advantage of relating to the logical structure of the book rather than to its mapping onto pages. This would afford some measure of invariance when text is added or removed, and with a carefully designed section structure this invariance could even extend to the addition or deletion of sections.

Finally, one aspect of indexing that is rarely considered in any index, whether paper or computerized, is the idea of generic indexing, or information mapping [15]. In this, it is not the actual words of the text which are indexed, but the generic section-type such as 'Abstract', 'Summary' or 'Overview'. Such indexing could be very useful to readers who wish to read only certain types of section across the book as a whole.

Currently, VORTEXT does not attempt to computerize the process of creating an index. VORTEXT does however create a concordance for each book and then provides a (hopefully) flexible environment in which the human indexer can work.

As the book is being written, the author manipulates chunks of text. These chunks are normally a paragraph in length but may be as short as part of a line or as long as a complete screenful. The author also specifies the generic name of each chunk, typically 'Chapter heading', 'Section heading', 'Text', 'Summary' or 'Diagram'. Each of these names is associated with two attributes which determine whether or not the text in chunks of that name is to be concorded, and whether or not the chunk name itself is to be concorded for generic indexing.

As a result of the concording process, the author is presented with a list of all the unique concordable words and names used in the book. The concordance may then be manipulated by the author into its index form: common entries are removed, stem and case synonym entries are joined, and entries with multiple definitions are split. Finally, the author may write a short glossary definition for each entry in the index.

VORTEXT readers have two methods of accessing the index. Firstly, they can turn to the index pages and select the required entry. Secondly, they can point-to and select a single word of text anywhere on a non-index page. In both cases a glossary panel is then displayed at the bottom of the screen which contains the glossary definition of that word (or some appropriate message if the word has been declared common). If this is sufficient information, the reader may remove the glossary panel and continue reading. If the reader wishes to look-up references to that word in the book then he must use the Mmi icons. The Mmi icons are the (up to) three icons contained in the glossary panel and are intended to provide the same form of information as the use of bold fonts, light fonts and *illus* in an index. The **M**ajor Mmi icon indicates that the selected word is used within heading texts in the book; the **mi**nor Mmi icon indicates that the selected word is used within the body of (concorded) texts in the book, while the **i**llustration Mmi icon indicates that it is used within the caption to a diagram.

When one of these Mmi icons is selected, a bookmark is displayed in the upper right-hand corner of the page. Like the main text page, the bookmark also contains closed pages which indicate both the number of references found for the selected entry and the readers' relative position within them (Figure 4).

By selecting the closed pages of the bookmark the reader may browse through the references. At any time the reader may select the closed pages of the main text in order to read the book in the usual way again, or may remove the bookmark in order to select and then browse through the references of a different word.

There are several aspects of the use of the VORTEXT index that require further

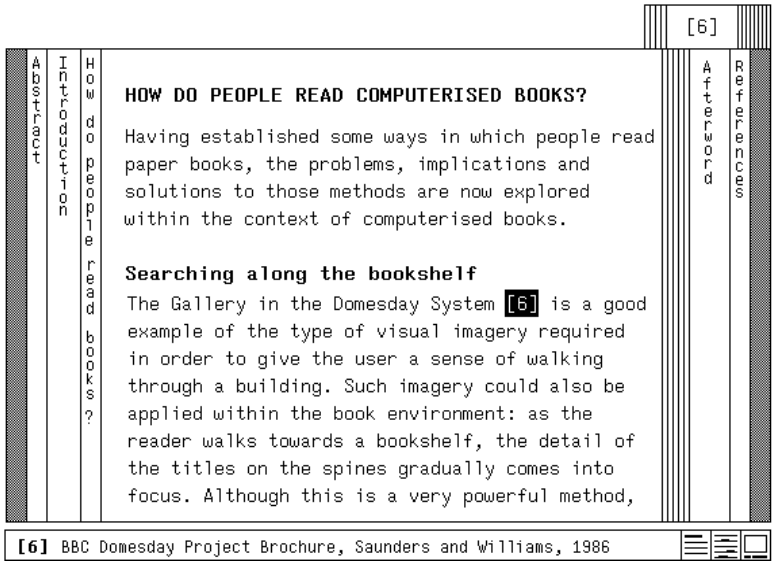

Figure 4. VORTEXT text page showing a glossary panel, Mmi icons and the bookmark

consideration. Firstly, although the concept of selecting an entry in the index is familiar, the concept of selecting a literal item such as a word of text is very unfamiliar, even to experienced computer users. Some extra emphasis is therefore needed in order to overcome this.

Secondly, if the VORTEXT reader selects both the major and minor Mmi icons for a single index entry then it seems logical for the bookmark to represent their merged combination. But what if the reader selects the major references to one word and the minor references to another? Should the bookmark represent their logical combination (so the reader is effectively browsing two lists at once), or should it represent their intersection (so in this case the reader is effectively browsing the second word as a subset of the first)? Since readers of conventional books normally browse only one reference at a time, research is needed to determine what, if anything, should be done.

The third aspect of indexing within VORTEXT is in the representation of page references by the bookmark. The closed pages of the bookmark indicate the number of references found and the readers' relative position within this list, but give no clues whatsoever about the position and distribution of the references with respect to the subject of the book. In order to try and overcome this problem an alternative form of bookmarking has been suggested. Now, when an Mmi icon is selected rather than displaying a bookmark of references, filing tabs are positioned down the edge of the appropriate closed pages (Figure 5).

In addition to releasing valuable space at the top of the screen, the tabs give an instant visual indication of the number of references found, their distribution over the book as a whole and their distribution within the subject of the book. The disadvantage of this method is that if the number of tabs is very large they could obscure the text of the closed pages.

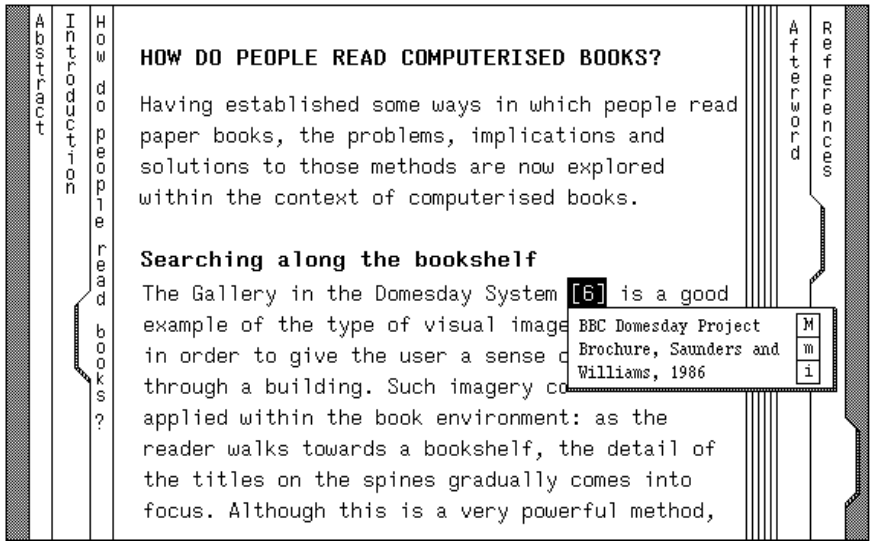

Figure 5. VORTEXT text page showing glossary panel and filing tabs

Finally, although by tradition there is only one single, static glossary definition associated with each word, there is no reason for the computerized version of a glossary to follow this restriction. Rather than containing a definition and providing access to Mmi references, a series of glossary panels could contain successive parts of a lengthy definition and could include closed pages with which to read them. Consider, for example, a novel in which the plot is highly complex. The name of each character in the novel could be associated with a series of glossary panels which successively detail the key events in that characters life: ‘Born on a dark and stormy night’, ‘Did research into reading methods’, ‘Wrote a magnificent paper for EP-odd’ and so on. Readers may then browse through the pages of the glossary panels in order to refresh their memories about that character. This would appear to be a unique and potentially very useful feature, although how far it could be applied within a computerized book remains to be seen since there is no ‘real book’ equivalent.

Can I make notes in the margin?

Personalization within a much used book is an extremely important part of the reading process. However, personalization by marginalia, association and actuality do depend very greatly upon the physicalness of the book and hence are instantly lost when the book is computerized. Currently, VORTEXT only implements a computerized form of personalization by actuality, but ideas for the implementation of the other two forms have been suggested.

Given a touch-sensitive screen and virtual screen overlays, personalization by marginalia could be implemented in a form similar to that used in the Reos project [16]. Since the marginalia screen is virtual it can be superimposed upon the ‘real’ page for its creator, or removed so that other readers are presented with a clean copy of the book. If a

variety of styli are provided then the reader has the choice of making comments in the margin (as if with a fine pen), or of highlighting sections of text (as if with a marker pen). Taking this idea a step further, if each stylus were associated with a particular author of that book then co-authors could annotate their own work, browse through the changes they have made, browse through the changes their co-authors have made or browse through all the changes together.

Some form of personalization by association with a coffee stain would be somewhat pointless — creating the image of the stain is simple enough, but how could the reader specify 'This is the page where I'm going to spill my coffee'? Perhaps a more useful idea would be to incorporate page viewing frequency counts. Each time a page is read a few more random pixels are drawn onto that page in order to give it that 'well-thumbed' look. In addition, the page is given an increasing bias equivalent to breaking the spine just that little bit more so that the book seems to fall open naturally at the pages that are referred-to frequently.

Finally, VORTEXT currently implements a form of personalization by actuality using a fingers icon as if the reader had used their fingers as temporary page markers. When the fingers icon is selected, the current page number is added to the list of references represented by the bookmark. The reader may then use the bookmark to browse through these kept pages which are displayed with the fingers icon video-inverted. This method is reasonably satisfactory but perhaps a more natural way of keeping a temporary place in a book is to turn down the corner of the page. This could be implemented in the computerized version by the reader dragging the pointing device across one corner of the screen. A shaded triangle could then be drawn to indicate that that page has been marked, and readers could turn directly from one marked page to another by selecting the appropriate corner (Figure 6).

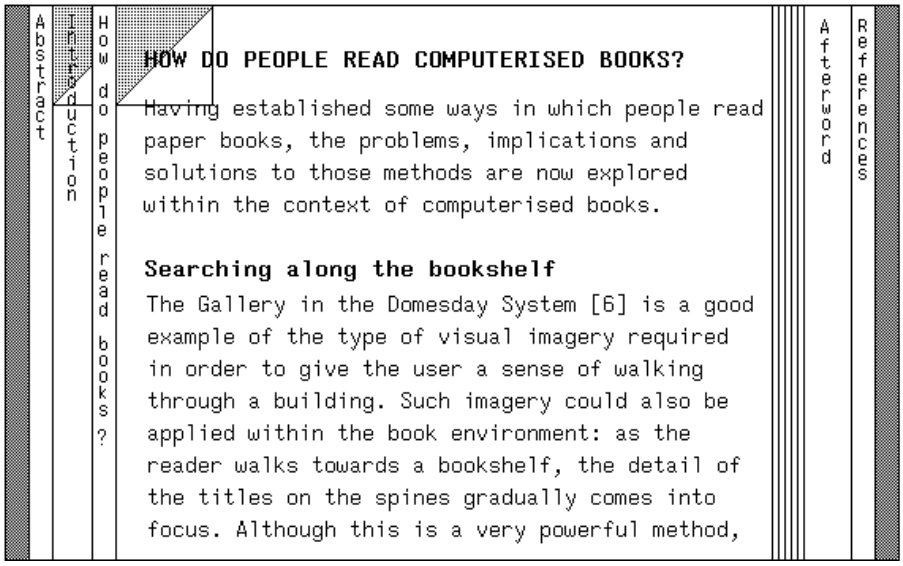

Figure 6. VORTEXT text page showing a marked page

Can I get a paper copy of this?

Until now, the representation of paper books on computer has been discussed, but what of the representation of computerized books back onto paper again? Screen dumps are very easy to create, but could it be guaranteed that the resolution of the printer matched that of the screen so as to give an equally readable paper copy? Since the closed pages were originally intended to substitute for the thickness of paper, are they actually needed on the paper copy? Parts of the contents page and the index references are effectively hidden on the screen, so how should they be revealed on the paper copy? Other than providing screen dumps, this is one area which VORTEXT has not attempted to address.

AFTERWORD

VORTEXT was developed as a PhD research project on a PERQ-1 at the University of Reading between October 1984 and December 1987. The ideas presented in this paper are a combination of the original design plus the enhancements suggested by 'real' experimental usage of the system. As such, no firm conclusions can be drawn beyond the fact that VORTEXT incorporates some potentially very interesting ideas, and that some ideas require considerable further work! For the future, it is intended to implement VORTEXT-2 on an Amstrad 1640.

REFERENCES

1. *The British Library Facts and Figures*, BLL Document Supply Centre, W. Yorkshire, April 1986.
2. *Oxford University Press General Catalogue*, Oxford, 1986.
3. *Cambridge University Press Catalogue and Index*, Cambridge, 1986.
4. B. Blackwell, *Guide for authors*, Basil Blackwell Publisher Ltd, Oxford, 1985.
5. I. Benest, G. Morgan, and M. D. Smithurst, 'A humanised interface to an electronic library', *Proceedings INTERACT '87*, Elsevier Science Publishers B.V. (North-Holland), (1987).
6. R.A. Bolt, *The Human Interface: where people and computers meet*, Lifetime Learning Publications, California, 1984.
7. V.A. Burrill, *VORTEXT: an experimental hypertext reading system*, PhD Thesis, Reading University, 1988.
8. *BBC Domesday Project Brochure*, Saunders and Williams, Croydon, November 1986.
9. I.H. Witten, 'Elements of computer typography', *Int. Journal of Man–Machine Studies*, **23** (6), 623–688 (1985).
10. R. Rubinstein, *Digital typography: an introduction to type and composition for computer system design*, Addison-Wesley Publishing Co, Reading, Massachusetts, 1988.
11. D.D. Woods, 'Visual momentum: a concept to improve the cognitive coupling of person and computer', *Int. Journal of Man-Machine Studies*, **21** (3), 229–244 (1984).
12. G. Salton and M.E. Lesk, 'The SMART automatic document retrieval system — an illustration', *Communications of the ACM*, **8** (6), 391–398 (1965).
13. J.D. Bishop and G.J. Smith, 'Experiences with a STATUS database', *Software — Practice and Experience*, **11** (12), 1315–1329 (1981).
14. D.C. Blair and M.E. Maron, 'An evaluation of retrieval effectiveness for a full-text document retrieval system', *Communications of the ACM*, **28** (3), 289–299 (1985).
15. R.E. Horn, 'Information mapping for learning and reference', *Datamation*, **31** (1), 86–88 (1975).
16. R. Kreuzer, 'Reos — a new document image processing system', *Electronics and Wireless World* 622–626 (1987).