

WinRefer: A Reference Management System for Windows

Version 1.0-beta, September 1992.

Si Dobson

Department of Computer Science
University of York
Heslington
York YO1 5DD

e-mail: simon@minster.york.ac.uk

18 September 1992

1. Introduction

Until quite recently, Unix-based platforms were in general the systems of choice throughout the academic community. Recently, however., there has been a significant shift towards PC systems, especially those using Microsoft Windows™ graphical environment.

A disadvantage of using Windows-based word processors is that they offer comparatively little support for the management of references, which is a major headache for academic users.

WinRefer is a simple program for searching through databases of references and formatting them according to a number of popular conventions, ready for insertion into a paper or book.

Copyright and Licensing

WinRefer is brought to you by the Pink Software Company - a group of academic computer scientists in several British Universities and research laboratories - which is dedicated to reducing the need for the sorts of over-priced rubbish encountered all too frequently in the PC world. We specialise in useful little tools of a high quality at a low price. Watch out for some Visual Basic custom controls in the near future!

Personal users are free to use and distribute WinRefer as you see fit. There is no license fee for personal users (if any organisations want to use it, let me know and we'll work something out), and no nagging dialogue boxes. The software isn't crippled - not deliberately, anyway! - it's a fully-functional application which we hope will be useful to many people.

However, should you find WinRefer useful, a small contribution to the author would be greatly appreciated. £20 (which was about \$30 last time I looked at the exchange rate)

would seem reasonable, and will encourage the development of WinRefer to the detriment of thesis-writing. We like to call this style of distribution "conscience-ware."

In addition, we'd like to hear peoples' experiences with, and comments on, WinRefer. Any very popular suggestions may find their way into future versions. E-mail contact is preferred, but my e-mail address is likely to change sometime in November 1992. If you can't get hold of me, try mailing ken@minster.york.ac.uk (Ken Tindell) instead.

2. Manifest

WinRefer is distributed either on a single disc or in a single ZIP archive. The following files are included:

setup.exe	}	
setup1.exe	}	
setup.lst	}	installation system
vbrun100.dl_	}	
setupkit.dl_	}	
winrefer.exe		the WinRefer program itself
refer.dll		the reference formatting library
commdlg.dll	}	
cmdialog.vbx	}	support files
winrefer.for	}	
default.for	}	
wordlet.for	}	additional format sets
word.for)	
winrefer.wri		documentation (Windows Write)

3. Basics

This section contains instructions for the basic use of WinRefer. More advanced topics are covered in §4.

3.1. Installation

If you obtained WinRefer as a ZIP archive, un-ZIP it into an empty directory (you're presumably done this already...).

WinRefer comes with an installation program, which will create the directories needed to hold the various files. Starting **setup.exe** from either Program Manager or File Manager will run the setup process.

Some words of explanation are in order. WinRefer is composed of two files an executable file and a supporting dynamic link library (DLL). Windows must be able to

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT2

find the DLL when the WinRefer main program is executed, which means that the DLL must live in the WinRefer start-up directory or in a directory which is on the search path. **setup** will offer to add the directory containing the WinRefer executables to your **autoexec.bat** file.

In addition to these executables, WinRefer must be able to find three additional files: the internal and default format sets and the default reference file (these terms will be explained in a minute). These can live in the WinRefer home directory, or in some other place. On my system, I have a directory called \usr\dict where the format sets and reference database files live, *à la* Unix.

Once installed, WinRefer may be included into a Program Manager group, dragged onto the desktop (if you're using Dropper, hint hint...), or whatever. When it is executed for the first time, WinRefer will request that the paths to the necessary format set and reference database files are supplied:



These files live in the directory which you specified for format sets and reference files to **setup**. The default file names for these items are **default.for**, **winrefer.for** and **refs.ref** respectively, although any format file may be used for the first and any reference file for the last. Once they have been supplied, the main WinRefer window is displayed.

3.2. A Quick Diversion: Reference Databases

The most common format for reference databases is undoubtedly the format used by Unix' **refer** program. In such a file, references consist of a number of lines each holding an element of the reference its title, author, journal name and the like. Each line normally begins with a keyfield consisting of a percent sign and a single letter indicating which element of the reference follows. References are separated by blank lines. For

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT3

anyone who isn't familiar with this form of reference file, a more complete description appears in Appendix A.

The other common format is used by the **bibtex** pre-processor to **latex**. This is considerably more complex.

WinRefer can use **refer**-format reference databases without modification. We've tested it on several large databases - including our local reference database and the (*very* large) commercially-available Inspec™ database - without a hitch. WinRefer should be able to access databases across a network, but we're not kitted-up to test this - any experiences welcome.

WinRefer *can't* read **bibtex** databases, and we have no plans to provide such support unless there's a lot of call for it. I believe most sites store their main databases in **refer** format anyway, so this hopefully shouldn't be too serious an omission.

3.3. WinRefer in Use

WinRefer appears to the user as a single window:



The window consists of three functional areas: the *search term* area, the *current search* area, and the control buttons. The search term area is where details of the references being sought are entered; the current search area allows references found to be browsed or restricted, or new searched to be started. The two control buttons give access to the WinRefer control panel, or exit the program.

3.3.1 Searching for References

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT4

WinRefer's *raison d'être* is to search for, retrieve and format references suitable for inserting into papers. Once retrieved, references can be formatted and pasted into papers.

Expressions and Keywords

References may be sought in two ways: by supplying a list of keywords or by using a more complex matching expression.

The simpler method is to supply a list of keywords together with the way in which they should be combined: a reference may be selected if any one keyword is present in it, or only if all the keywords are present.

To use keyword searching, the keyword button should be selected in the search term area. Up to four keywords may then be supplied into the boxes. The connective - any keyword or all keywords - may be set by selecting the appropriate button.

Expressions are a more complex, but more flexible, method for retrieving references. A matching expression takes the form of a number of keywords connected by logical terms: for example, the following are all matching expressions:

real-time
dobson and smith or transputers
dobson or tindell
(dobson or tindell) and not (burns or wellings)

As with keywords, the case of a search term is ignored during the search. Evaluation is left-to-right in the absence of brackets.

The "Start search..." button becomes enabled when data has been supplied for a search either an expression or some keywords. Clicking it will start a search of all the reference databases known to WinRefer. A progress box is displayed showing the file currently being searched and the number of references searched so far in that file:

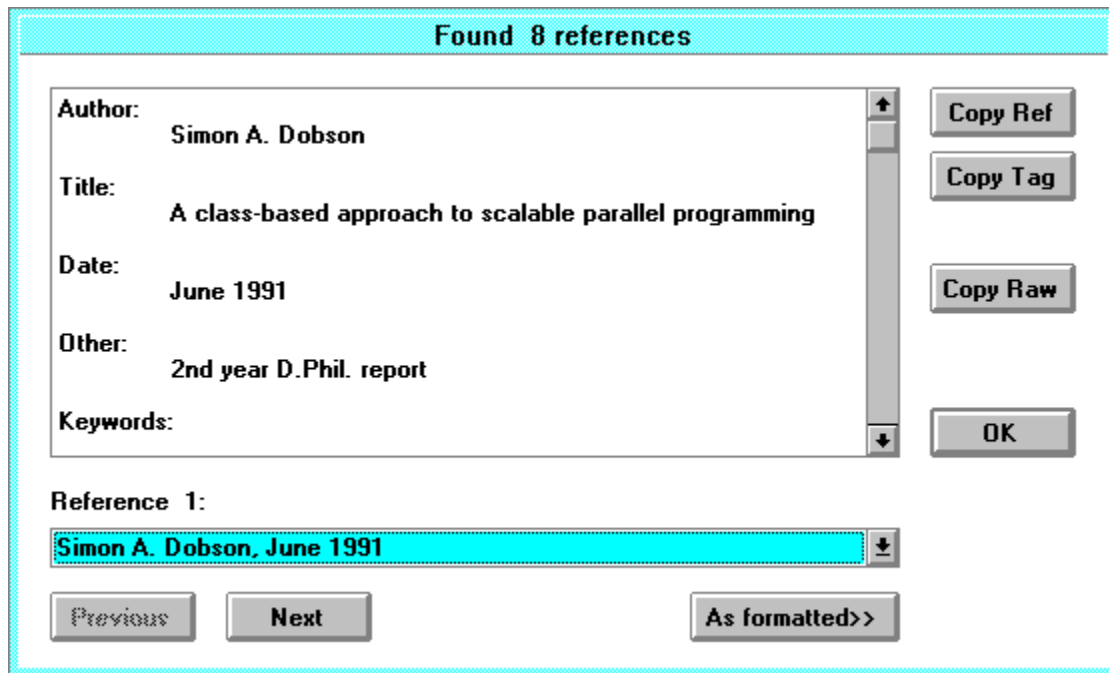


Hitting Cancel in this box will abort the search; otherwise WinRefer will search all references in all database files for matching references. Once the search is complete, if any matching references are found, the browsing dialogue will be displayed.

Browsing and Copying

The references found as the result of the latest search may be examined at any time by clicking the "Examine..." button. It appears automatically when a search is completed:

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT5



The main window contains the current reference, in a (largely) unformatted state. The combo box displays (by default) the authors and data of publication of the current reference. A particular reference may be selected for display by choosing it from the combo list; alternatively, clicking the "Next" or "Previous" buttons will move between the references.

The three copy buttons on the right-hand side of the dialogue allow the reference to be copied to the clipboard. "Copy Ref" will format the selected reference according to the current format set (see later) and place it on the clipboard. "Copy Tag" will format and copy the selected reference's tag. "Copy Raw" will copy the reference to the clipboard in its raw (Unix **refer**) form.

The "As formatted>>" button extends the browsing dialogue to include a box which shows the selected reference as formatted according to the current format set. Clicking "Okay" closes the dialogue and returns to the main window.

Narrowing a Search

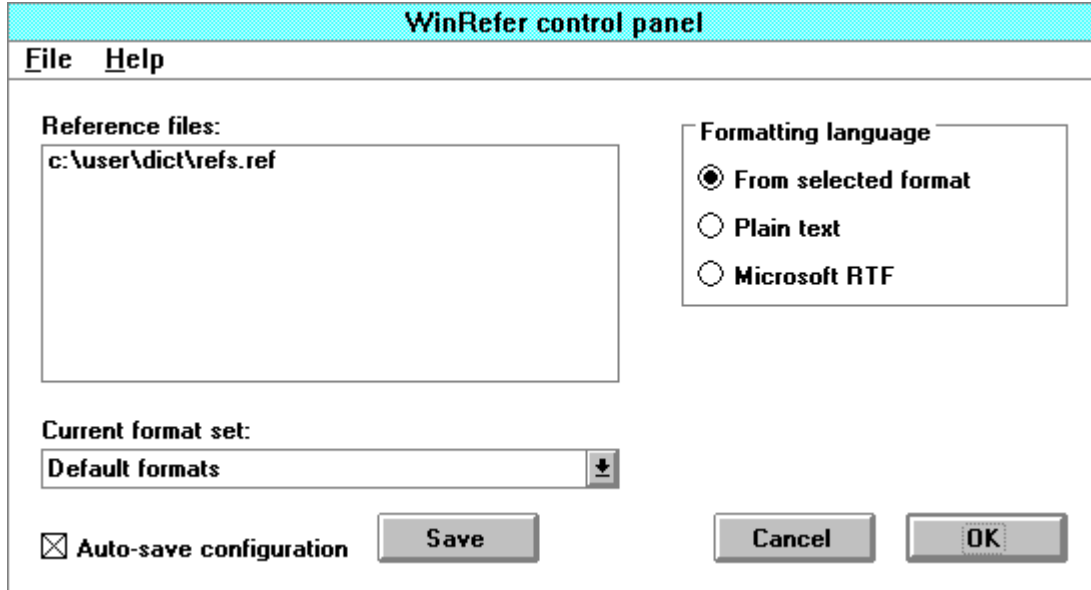
If too many references are found as the result of a search, the search may be narrowed by changing the selection criteria - expression of keywords - and clicking the "Narrow search..." button. This applies the search only to those references found as a result of the *last* search, not to any reference files.

3.3.2 System Settings

The system settings are controlled using a single dialogue box obtained by clicking

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT6

the "Setup..." on the main window. This brings up the following dialogue:



The list box labelled "Reference files:" contains a list of all the reference databases which WinRefer will search for references. A new file may be added to this list by selecting the "Open reference file..." option from the File menu: this will display a standard file-chooser dialogue. Several reference files may be chosen from this dialogue, and each will be added to the file list.

To delete a database file, select it with the mouse and choose the "Close reference file" option from the File menu.

The combo box below displays the description of the current format set being used. The displayed format set is used to lay-out references which are found as a result of searches, both for browsing and for insertion into documents. Several format sets may be available for selection, although only one is actually in use (and in memory) at any one time. To add a new format, choose the "Add set of formats..." option from the File menu, and select the appropriate file(s); to remove a format set, select it into the combo box window and choose "Remove current set of formats" from the File menu.

The set of buttons titled "Formatting language" determine how text styles such as emboldening are implemented. A preferred formatting language is defined in every format set, and this default may be used by selecting the From selected format button. Alternatively, format sets may be forced to use plain text or Microsoft RTF to format references. Selecting plain text will lose all stylistic elements such as italics and emboldening; using RTF will allow formatted references to be pasted into (for example) Word for Windows.

The check box labelled "Auto-save configuration," if selected, causes WinRefer to store any configuration changes (reference files, format sets, auto-save *et cetera*) between executions.

Clicking "OK" causes the configuration to be committed for use by WinRefer. Clicking "Cancel" ignores changes and reverts to the previous configuration. The "Save"

button behaves like OK, and also saves the current configuration immediately.

4. More Advanced Use

The above covers the basic use of WinRefer, but there are more advanced modes of use. Chief amongst these are the ability to use different reference formatting conventions, and to define new conventions.

4.1. Format Sets

The first thing which anyone will notice on looking at a paper is that references may be handled in many different ways. There are several conventions for citing references in the text: as [1] or [Dobson91a] or¹. The actual references may appear at the bottom of a page or at the end of the paper (or chapter, or book), and may be formatted in different ways. Different styles are also used for books *versus* papers in journals *versus* papers in proceedings *et cetera*.

A *format set* is a file which describes the way in which references are formatted. It describes both the way references are cited in the text (the reference's *tag*) and the way in which the reference itself appears.

Supplied Format Sets

WinRefer comes supplied with three format sets. One is the default and two are specialised for use with Word for Windows. Each format set can identify references which to papers in journals, books, papers in books and technical reports, and can format each type of reference appropriately.

The Default Format Set

The default format set **default.for** is the most basic format set. It formats references in plain text, and does not generate reference tags.

Word for Windows Format Sets

There are two format sets for Word for Windows. Both generate references and tags using Microsoft's Rich Text Format, but generate citation tags (and the corresponding tags in the formatted reference) differently.

The first set, **word.for**, tags all references with a sequence number in square brackets according to the reference's position in the final reference list, *i.e.* [1]. This list may be sorted as required, and the sequence numbers will reflect their associated reference's position.

The other set, **wordlet.for**, generates tags using the major author's surname, the year of publication, and a unique letter, *i.e.* [Dobson91a].

Sometimes you may find that a document has "**Error! Bookmark not defined!**" warnings appearing. This is because the WinRefer's cross-referencing mechanism is

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT8

implemented using bookmarks. To get rid of the errors, select the whole document (Ctrl-Shift-Keypad 5) and hit F9 a couple of times to unpadte all the fields. If you've still got warnings, it means that you've pasted in a tag without a corresponding reference to resolve it to.

4.1.2. Custom Format Sets

A format set is stored as a simple text file, and it is relatively straightforward to create new format sets to implement personal preferences or the policies of particular publishers, or to interface to other word processors.

Creating or Modifying a Format Set

The exact grammar used in defining format sets is given in appendix B: here we shall give an overview. See the supplied format sets for examples.

Basics

A format set provides WinRefer with several pieces of information:

- o a description of the format set;
- o the formatting language to be used by default;
- o a collection of formats, each of which contains information on when to use this format and how to format a particular reference; and
- o a tag description showing how tags are to be generated.

The processing of such a file occurs *en bloc* when a format set is selected from the control panel (or when WinRefer is first loaded).

Description and Language

The description line contains atextual description of the format set, and is displayed on the WinRefer control panel in the format set combo box. If no description is present, this box will contain the format set's file name.

WinRefer is capable of generating output in a number of print description languages. At present plain text and Microsoft rich text formats are supported. The language description specifies which language should be used: **plain** or **rtf**. This language will be used by default, but may be over-ridden from the control panel.

Parts of a Reference

WinRefer considers a reference to be composed of several parts, or *fields*. Each field loosely coresponds to a line from the database (described in Appendix A), although this

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT9

need not be the case. All WinRefer formats are expressed in terms of combination of reference fields.

Within a format description file, the following names are used to describe parts of a reference:

Title	Authors	ReportNumber
Book	Journal	Editors
Publisher	Volume	Number
Pages	City	Date
Others		

The Inference Part

Since a format description file may contain many formats, WinRefer must be able to determine which format to use for a particular reference. This function is performed using the inference parts of formats.

The inference part is a collection of field names, which may be necessary or optional. Collections of optional fields are delimited by square brackets. In inferring which format to use for a reference, WinRefer checks the fields of the reference against the inference parts of each available format. It considers only those formats whose necessary fields are all present within the reference; of these, it chooses the format with the most matches in optional fields.

If no matching format can be found, WinRefer does not format the reference and displays **[No format appropriate for this reference]**.

The Format Part

The format part of a format controls how a reference is printed. Once a format has been selected for a reference using the inference mechanism, the reference is formatted according to these instructions.

A format consists of a set of instructions, which are expended to generate the formatted reference. They are taken from the following:

- o simple fields*, which expend to the field taken from the reference;
- o literal characters*, which are inserted as-is;
- o optional clauses*, which are only formatted if present in the reference;
- o font clauses*, which allow the typeface used in formatting to be set; and
- o formatted fields*, which contain special formatting instructions to format a single field.

Each possible instruction is described below.

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT10

Simple Fields

A simple field is simply the name of a field from the reference. It is expanded to the text contained in that field. The following are all simple fields:

Authors **Date** **Pages**

In addition there are two special fields **MajorAuthor** and **MinorAuthors**, which extract the major and minor authors of a paper from the author list.

Literal Characters

Literal characters are inserted into the formatted reference as-is, without further formatting. They are delimited by double-quote characters. These are a number of special symbols which may be used to insert literals not available directly from the keyboard: **66** or **OpenQuotes** and **99** or **CloseQuotes** insert double open- and close-quotation marks respectively; **Tab** inserts a tab character; and **Para** begins a new paragraph. The following are all valid literals:

"," **","** **66**
CloseQuotes

Optional Clauses

Optional clauses may be used to deal with optional fields from a reference. An optional clause is displayed only if there is a field within it which is present in the reference.

Optional clauses are delimited by square brackets, and may contain any set of instructions including further optional clauses. The following are all valid clauses:

[Date] **["(" Number ")"]**
["," Volume [" " italic { Number }]]

In formatting an optional clause, WinRefer checks to ensure that there is a field within the clause which is present in the reference. If such a field exists, the optional clause is formatted.

In the last example above, formatting would occur as follows. If the reference did not contain a **Volume** field, no action would occur; if a **Volume** field was present, it would be formatted (as, for example, ,2). Next the nested option would be tested: if the reference contained a **Number** field, this would be formatted (as 3); otherwise, no further action would occur until the end of the clause.

The simplest way to view optional clauses is that they are only formatted if there is a field within them to print; otherwise, no formatting takes place. It is not possible to express if-else conditions, only if conditions.

Font Clauses

Some parts of a reference may be best printed in a different typeface: a font clause allows this to be specified. A font clause contains a typeface description and a set of instructions: all the formatting of the instructions occurs in the given typeface. Four typefaces are available: **normal**, **italic**, **bold** and **bolditalic**. The following are both valid font clauses:

italic { "et alia" } bold { "(" Number ")" }

Formatted Fields

Some of the fields mentioned above have a considerable degree of internal structure, and it may be desirable to perform some formatting internally. This is particularly the case with people's names, which may be formatted in several ways. Formatted fields allow this.

A formatted field is a field name followed by some subsidiary instructions in round brackets. Only certain types of field have this feature **MajorAuthor**, **MinorAuthors**, **Authors**, **Editors** and **Date** at present. The permitted instructions are given below: literals may also be present.

From **MajorAuthor**, **MinorAuthors**, **Authors** and **Editors** fields the following items may be extracted: **Surname**, which expands to the named person's surname; and **Initials**, which expands to their given names and/or initials. For **Date** fields, the **Century** option returns the year within the century of the reference (so 25th May 1992 would be reduced to 92).

The following are valid formatted fields:

Date(Century) MajorAuthor(Surname "," Initials)

Field Truncation

A field may be truncated by appending a colon followed by a width. In this case, the formatted field will contain at most the given number of characters. Negative numbers may be used to acquire characters from the right-hand side of a field; positive numbers acquire characters from the left-hand side. Field truncation is only available on simple and formatted fields (not optional or font clauses).

The following are all truncated fields:

Title:2 MajorAuthor(Surname):5

The Handling of Lists

Some fields may have several components for example there may be several minor authors of a reference, or several sets of pages. WinRefer handles such lists by inserting

the appropriate separation characters (a comma or full-stop).

Tags

Tags are used in documents to uniquely identify references. They are generated, according to a number of different conventions, from the information contained within the reference.

Tags may contain all the elements described above. In addition, the simple field **Unique** is expanded to a (pretty much) unique number identifying the reference. This is used when generating sequence letters: see the **wordlet.for** format set for an example.

The tag is placed onto the clipboard whenever the "Copy Tag" button on the reference browser dialogue is clicked.

An Example

To illustrate the use of format sets, consider the following simple format set file:

```
description "Default formats"
```

```
/* use plain text for output */  
language { plain }
```

```
/* tags */  
tag { "[" MajorAuthor(Surname) Date (Century) "]" }
```

```
/* paper from journal */  
format { Title Authors Journal [ Vol No ] [ Date ] [ Pages ] [ Others ] }  
as { MajorAuthor [ ", " MinorAuthors ", "  
      ", " 66 italic { Title ", " } 99  
      " " Journal [ " " Vol ] [ "(" bold { No } ")" ] [ "(" Date ")" ] [ " pp." Pages ]  
      ". "  
      [ " " Others "." ]  
      Para Para }
```

```
/* book */  
format { Title Authors Publisher [ City ] [ Date ] [ Others ] }  
as { MajorAuthor [ ", " MinorAuthors ]  
      ", " 66 italic { Title ", " } 99  
      " " Publisher [ ", " City ] [ "(" Date ")" ]  
      ". "  
      [ " " Others "." ]  
      Para Para }
```

This format set appears in the WinRefer as Default formats, and is formatted using plain ASCII. Tags are generated in square brackets from the major author's surname and

the date of the paper's publication, *i.e.* [Dobson91]. Clicking "Copy Tag" from the browser dialogue will place this tag onto the clipboard, from whence it may be placed into a document.

There are two possible formats. The first is used on references which must contain a title, author and journal and may contain fields including date, volume, number, page number and others. The second is used for references without a journal but including a publisher. Any references which match neither pattern will not be formatted.

A reference to a journal paper (the first possible format) might appear as follows: the reference

```
%A Mark A. Linton
%A John M. Vlissides
%A P.R. Calder
%T Composing user interfaces using InterViews
%J IEEE Computer
%V 22
%N 2
%D February 1989
%X HCI:user interfaces:toolkits
```

would be formatted as:

```
Mark A. Linton, John M. Vlissides and P.R. Calder, ``Composing user
interfaces using InterViews," IEEE Computer 22(2) (February 1989).
```

and would be tagged in the text as [Linton89]. The formatted reference is placed on the clipboard when Copy Ref" is clicked.

The Internal Format Set

The internal format set **winrefer.for** is used to generate the one-line description of each reference displayed in the combo box of the browser dialogue (§3.3.1). It is a format set like any other, albeit a very simple one, and may be changed if required. An example would be to change it to display the title and date of each reference rather than the authors and date:

```
format { Title [ Date ] }
as { Title[ ", " Date ] }
```

5. Future Intentions

This is the release of WinRefer, and we fully intend - work permitting! - to improve it in time. We have two major intentions for the immediate future.

The first is to provide an indexing capability. At present all searches are performed linearly through the selected reference files, which can be very time-consuming on a slow

machine. We intend to provide a facility for generating cross-indexes to key words within references, like those generated by the Unix tool **pubindex** (although probably in a different format). These would be generated and maintained automatically.

The second major intention is to provide some form of automatic control. At present, it is the user's responsibility to generate the list of references for a paper: WinRefer, unlike **refer**, does not perform bibliography generation. The reason for this is simply that there are many different possible implementations of this feature, and it is not obvious which would be of most use.

The preliminary idea is to control WinRefer down a DDE link. The problem is that, within the Windows model of data linking, control would be vested in the word processor, not in WinRefer, and this would make life difficult for some word processors (notably Write).

If anyone has any clear ideas about how they'd like bibliography generation to work, we'd really like to hear them.

6. Acknowledgements

My thanks go to Ken Tindell, Martin Atkins and Paul Butcher, without whom this program would never have been written. Ken made numerous suggestions about the user interface, and tried out all the combinations which we came up with: I'd like to think that the result is quite usable.

WinRefer is written in a combination of Zortech C++ and Microsoft Visual Basic, and uses the Microsoft Visual Basic Professional Toolkit.

A. refer-format Reference Databases

A **refer-format** reference database is composed of sets of lines describing a reference, with references being separated by blank lines. Each line begins with a code indicating what the rest of the line contains: the title, author, publisher *et alia*. A line without a code is assumed to be a carry-on from the previous line.

WinRefer recognises the following codes:

Code	Meaning
%T	title
%A	author
%B	book title
%E	editor
%I	issuer (publisher)
%C	city (of publication)
%J	journal name
%V	volume number
%N	number in volume
%D	date (of publication)
%R	technical report number
%O	other (appended)

%P	pages
%K	keywords
%X	other (not printed)

Any other code letters are treated as %X lines. (Unix **refer** actually recognises a few more codes than these, but they are very seldom encountered.)

B. The Grammar of Format Set Files

The grammar is presented in EBNF form. Brackets are used for grouping. Items in square brackets are optional (as far as the grammar is concerned); items separated by vertical bars are mutually exclusive; items in bold are literal elements which may appear in the text. An elipsis denotes zero or more instances of the preceding item.

formatfile	=	element [element...]
element	=	description language format tagdescription comment
description	=	description literal
language	=	language { (plain rtf) }
format	=	format { inferencepart } as { formatpart }
inferencepart	=	(name optionalinference) [(name optionalinference)...]
name	=	Author Title Date Journal Volume Number Book Editors Pages Publisher City Others
optionalinference	=	[inferencepart]
formatpart	=	formatelement [formatelement...]
formatelement	=	simplefield literal optionalclause fontclause formattedfield
simplefield	=	(name Tag MajorAuthor MinorAuthor) [: n]
literal	=	" c [c...] " 66 OpenQuotes 99 CloseQuotes Tab Para
optionalclause	=	[formatpart]

fontclause	=	font { formatpart }
font	=	normal bold italic bolditalic
formattedfield	=	namefield (nameformat [nameformat...]) datefield (dateformat [dateformat...])
namefield	=	Authors MajorAuthor MinorAuthor Editors
nameformat	=	Surname Initials literal
datefield	=	Date
dateformat	=	Century
tagdescription	=	tag { tagformatpart }
tagformatpart	=	tagformatelement [tagformatelement...]
tagformatelement	=	formatelement Unique
comment	=	/* c [c...] */

Note that some elements have alternative names, for example **66** and **OpenQuotes**. Capitalisation is optional, as is the use of plurals.