

# WinRefer: A Reference Management System for Windows

Version 3.0, August 1993.

Si Dobson

Informatics Department  
SERC Rutherford Appleton Laboratory  
Chilton, Didcot, Oxfordshire OX11 0QX, UK.

*e-mail:* sd@inf.rl.ac.uk

25 August 1993

## 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 reference citations and bibliographies, which is a major headache for academic users.

WinRefer is a program for searching through databases of references and formatting them according to a number of popular conventions, ready for citation in a paper or book. Its main features are:

- o simple but flexible user interface
- o support for a number of popular reference file formats
- o search for references using up to four keywords, or using more complex expressions
- o automatic file index generation and maintenance
- o user-defined formatting conventions for references and citations, with several of the most popular supplied as standard
- o automatic generation of bibliographies
- o handles plain text and Microsoft Rich Text (as used by Word for Windows)
- o a DDE interface for remote control from other applications
- o a template for use with Word for Windows

- o e-mail contact with the author

The development of WinRefer is very much an on-going project, so comments, "wish lists," bug reports and the like are welcome (preferably by e-mail). Some ideas for the near future are described in §5. Please enclose a (small) sample database which illustrates any bugs reported: *please don't* send a "Dr Watson" trace, as they don't tell me anything!

## 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.

You are free to use and distribute WinRefer as you see fit. There is no license fee 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.

## 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
index.dll            the index sub-system

troff.rml           }
chemabs.rml         }
medline.rml         } reference management libraries
bids.rml            }

test_ref.ref        }
test_ca.ca          } example reference databases
test_med.med        }

commdlg.dll         }
cmdialog.vbx        }

```

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

winrefer.for	}	support files
default.for	}	
hitlist.txt	}	
wordlet.for	}	
word.for	}	additional format sets
wordseq.for	}	
winrefer.dot		Word for Windows template
winrefer.wri		documentation (Windows Write)
tutorial.wri		tutorial document (Windows Write)

### 3. Basics

This section contains instructions for the installation and 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. If you installed an earlier version of WinRefer on your system, I suggest you erase all traces of it before running **setup**.

Some words of explanation are in order. WinRefer is composed of several files, an executable file and some supporting dynamic link libraries (DLL's). Windows must be able to find the DLL's when the WinRefer main program is executed, which means that they 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 the internal and default format sets (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. **setup** will make the location of these files, wherever you install them, available to WinRefer automatically.

**setup** will also offer to install various optional files, such as additional format sets, reference management libraries and example databases.

(There is a known problem with installing WinRefer on systems running Norton Desktop for Windows - **setup** may cause NDW's program groups to be corrupted, although it will *not* damage any files.)

Once installed, WinRefer may be included into a Program Manager group, dragged

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

onto the desktop or whatever. When run, the main window appears.

### 3.2. A Quick Diversion: Reference Databases

The most common format for reference databases is probably that 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.

Other common formats are that used by the **bibtex** pre-processor to **latex**; the **ProCite** format; and the format used by things like Medline - there are many others.

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. The program will happily access database files across a Windows for Workgroups network. WinRefer also provides support for Chemical Abstracts, Medline, and BIDS (ISI), although it doesn't recognise all the information. Brief descriptions of all these file formats may be found in Appendix A.

### 3.3. Terms

We shall use several terms in the following discussion which may need explanation.

A *reference* is a collection of details about a piece of published work - its title, authors, the journal or book in which it appeared *et cetera*. A citation to a reference appears in the text - for example [1] or [Dobson1992a] - whilst the full text appears elsewhere. In the sciences, reference texts are usually gathered together into a bibliography at the end of the paper; in the humanities the texts frequently appear as footnotes on the page on which they are cited.

WinRefer calls the citation marker a *tag*. A tag is inserted wherever a reference is cited. The full text of the reference is *formatted* to appear in the bibliography, footnote or wherever, with the tag being used to identify which text is being cited. WinRefer allows many different conventions to be used when formatting texts and creating tags for citation (the two styles above - [1] and [Dobson1992a] - are supported as standard). A formatting convention is called a *format set*, and WinRefer stores such format sets in files on disc. A format set file contains instructions on how to format references and tags, and may format different *type* of reference differently - references to books may appear differently from references to papers in journals or private communications, for example.

WinRefer searches for references in a file containing descriptions, called a *reference database*, retrieves them, and formats their texts and tags for insertion into papers and bibliographies.

### 3.4. WinRefer in Use

WinRefer appears to the user as a single window consisting of three functional areas:

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

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.

### 3.4.1 Searching for References

WinRefer's *raison d'être* is to search for, retrieve and format references suitable for inserting into papers. Once retrieved, references and citations 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. Up to four keywords may be supplied.

Expressions are a more complex, but more flexible, method for retrieving references. A matching expression takes the form of a number of single 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)

Valid logical connectives are "and" and "or"; "not" may be used to invert an expression; and expressions may be bracketed as required (there is no precedence amongst operators). 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 search mode used may also be controlled through the "Options" menu.

The "Clear" button clears all the search terms or expressions, and resets WinRefer to begin a new search. The "Clear references" option of the "Search" menu does the same.

Clicking the "Start search" button (or selecting "Start search" from the "Search" menu) 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.

There is a limit to the number of references which WinRefer will return as the result of a single search - currently this limit is 100. If more than this number are found, a

warning box will be displayed: any additional references will be ignored.

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 or by selecting "Examine references..." from the "Search" menu. It appears automatically when a search is completed.

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 form, much as it appeared in the database file.

The "As formatted>>" button extends the browsing dialogue to include a box which shows the selected reference as formatted according to the currently selected 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 or by selecting "Restrict search" from the "Search" menu. This applies the search only to those references found as a result of the *last* search, not to any reference files.

### **3.4.2 System Settings**

The system settings are controlled using a single dialogue box obtained by clicking the "Setup..." on the main window, or by selecting "Setup..." from the "File" menu.

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.

It is possible to open reference files from the "File" menu in the main window, without going through the Setup dialogue.

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.

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

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 at end of session," if selected, causes WinRefer to store any configuration changes (reference files, format sets, auto-save *et cetera*) between executions.

The check box labelled "Preserve bibliography between sessions" causes WinRefer to save its current bibliography into a file. The bibliography is automatically re-loaded when WinRefer next runs. On the "File" menu is an option labelled "Bibliography save file..." which allows the name of the auto-save file to be set (although the default is probably adequate for most purposes).

Clicking "OK" causes the configuration to be committed for use by WinRefer. Clicking "Cancel" ignores changes and reverts to the previous configuration. The "Save now" 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. These include the use of indexes to improve search times, the ability to define new reference formatting conventions, WinRefer's support for multiple reference database file formats, bibliography generation, the control of the application *via* DDE, and other goodies.

### **4.1. Indexing**

When you press the "Start search" button WinRefer searches all the selected reference databases for the keywords you have selected. This may involve searching a lot of data, which can be a very time-consuming process.

To accelerate searches, WinRefer includes the ability to create and maintain indexes to reference files. An index file cross-references all the important words found in a reference database file with their positions, so references containing particular keywords may be retrieved far more quickly. Index files take up more space on disc and in memory, but offer a great increase in efficiency.

WinRefer's index management is designed to be almost completely transparent in use.

All index selection, location, creation and maintenance are handled automatically. It is controlled by a single set of buttons on the control panel headed "Index generation."

There are four ways to use indexes. The first is to ignore all indexes and use sequential searching, which may be useful in machines with little memory or disc space. Slightly more useful is to use an index only if it present and up-to-date (meaning that the reference database file has not been edited since the index was created).

The main indexing modes, however, re-build indexes if they are missing or out-of-date. The two options are to prompt before re-building the index, or to re-build automatically. In both cases, the index will be (offered to be) re-built before searching the reference file which it indexes. WinRefer will *never* use an out-of-date index.

Another option is to cache indexes in memory. If selected, this casuses an index never to be discarded once it has been loaded (unless it becomes out-of-date). This means that it is immediately available for use, and does not have to be loaded from disc whenever searching is performed: the disadvantage is that cached indexes are always using memory, and so caching is only really appropriate on machines with a lot of memory. If memory is tight, turn caching off.

Indexes are built using most of the fields in a reference - notably title, author, journal and abstract. Each keyword is entered into the index, and may be searched upon. WinRefer allows a "hit list" of words which should *not* be indexed - words such as "a", "of", "not" *et cetera* - to be supplied (from the "File" menu). This list may be customised for each user (to exclude words which are common in particular domains, for example, or to add words in French). A default hit list file is supplied.

In use, indexes reduce search times to a tiny fraction of that of sequential seach. By way of example: my 20Mhz 386SX takes around two minutes to search for a single keyword in a reference database of 800 references, but the same search using indexing takes around a second. On faster machines, this speed increase renders searching effectively instantaneous.

## 4.2. 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] or <sup>1</sup>. 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 and sorted 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 n the text (the reference's *tag*) and the way in which the reference itself appears. It does *not* define where a citation or a reference appears in a document: this remains the province of the word processor.

### 4.2.1. 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

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



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 three format sets for Word for Windows. All 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 second set, **wordlet.for**, generates tags using the major author's surname, the year of publication, and a unique letter, *i.e.* [Dobson91a].

The final set, **wordseq.for**, generates tags and references which may be sorted into inclusion order - the first reference in the bibliography will be the first reference cited, and so forth. (More instructions for using this format set may be found in **winrefer.dot**.)

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

## **4.2.2. 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. Unfortunately you really need to think like a computer scientist to get any benefit from the following description! (I'm considering writing an editor for creating format sets graphically.) See the supplied format sets for examples.

### *Basics*

A format set provides WinRefer with several pieces of information:

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

- 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 a textual 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 corresponds to a line from the database (described in Appendix A), although this 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:

<b>Title</b>	<b>Authors</b>	<b>ReportNumber</b>
<b>Book</b>	<b>Journal</b>	<b>Editors</b>
<b>Publisher</b>	<b>Volume</b>	<b>Number</b>
<b>Pages</b>	<b>City</b>	<b>Date</b>
<b>Others</b>		

WinRefer actually distinguishes other elements in a reference, such as copyright notices, addresses and affiliations, but these cannot be included into a format.

### *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.

### *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 what is hopefully 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 year 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 (refer-style) 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 (using the "paper in journal" format) 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 cited 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 ] }
```

### **4.3. Extra Database File Formats**

Additional reference database file formats are handled using a dedicated *reference management library*. All reference management libraries are held in files with the extension **.RML**. At present four are supplied with WinRefer - supporting **refer**, Chemical Abstracts, Medline and ISI style reference databases - but more should hopefully become available in time. The default **refer**-format library is automatically installed; the others may be selected for use in WinRefer from the Setup dialogue, by selecting "Management libraries..." from the File menu. This brings up a dialogue box.

The dialogue allows new RML's to be installed into WinRefer, and allows the descriptions of installed libraries to be changed. A list shows the currently-installed RML's: pressing "Install..." brings up a file selector in which a new RML file may be selected and installed.

Each RML deals with files having a particular extension. The extension handled by each RML, and the short description of what sort of references such files contain, may be altered by selecting an RML from the list and then editing the fields of the dialogue. Each RML has default values for its extension and description which are installed along with the library. These defaults may be re-installed by pressing the "Defaults" button.

Selecting an RML and pressing "Remove..." removes the library from WinRefer (but not from disc).

### **4.4. Bibliographies**

A bibliography is a list of formatted references. WinRefer can automatically generate and format such a list using the standard search and format mechanisms, ready for

insertion into a paper.

WinRefer maintains a single bibliography. References may be entered into this bibliography by pressing the "Select" button in the browser dialogue. Such references are then stored in the bibliography for future use, even when the dialogue is closed.

The current bibliography may be viewed by pressing the "Biblio..." button on the browser dialogue. A new dialogue then appears containing the bibliography which has been generated in a formatted form - which may be incomprehensible, depending upon the format set being used!

The "Load..." and "Save..." buttons respectively load a bibliography from a file (making it current and erasing the current bibliography) and save the current bibliography into a file for later use. The bibliography file is not directly useful to other programs, so a bibliography must be loaded back into WinRefer before use. Note that saving a bibliography this way is independent of the bibliography-preservation mechanism mentioned in §3.4.2.

The "Copy" button on this dialogue copies the entire bibliography to the clipboard. The "Clear..." button erases the bibliography. "OK" returns the the browser.

An important point about bibliographies which should be borne in mind is that the references are stored in formatted form. If you alter the format set being used by WinRefer, all *future* additions to the bibliography will be formatted using this format - but all *existing* entries will be unaffected. This is due to the need to save bibliographies to disc, which makes it impossible to store references unformatted. Make sure you've settled on a format set before building a bibliography! Also note that it is perfectly possible to store a reference in a bibliography several times.

#### 4.5. Controlling WinRefer *via* DDE

DDE - Dynamic Data Exchange - is a mechanism by which Windows applications may exchange data. This is not the place to describe the workings of DDE but, for the knowledgeable, WinRefer can act as a DDE server.

When initiating a DDE link, the application name should be set to **WinRefer** and the topic to **Search**. The link *must* be a cold link - it makes no sense to update such a link *at any time*. WinRefer recognises four topics on links:

- o ReferenceTag - the tag of the last-selected reference
- o FormattedReference - the last reference formatted
- o RawReference - the last reference raw
- o Bibliography - the current bibliography

The service also accepts five commands (using the DDE "execute" style):

- o PasteReferenceTag }
- o PasteFormattedReference                    } paste data to the
- o PasteRawReference } clipboard
- o PasteBibliography }
- o WinRefer - bring WinRefer to the foreground



The DDE service works closely with the bibliography service. The reference available for DDE is always the references selected *last* into the bibliography - *i.e.* the reference which was displayed when you last pressed "Select."

#### **4.6. WinRefer and Word for Windows**

For users of Word for Windows, the **winrefer.dot** template contains some macros which use the DDE mechanism to allow reference selection from inside Word, without using the clipboard.

Full instructions for the use of this template are contained within it.

#### **4.7. WinRefer on a Network**

Many sites may use some form of Windows network (*e.g.* Windows for Workgroups), with reference databases living on a file server which is remote from the workstation running WinRefer. Just to set peoples' minds at rest, WinRefer will work quite happily in this setting.

When installing a new reference database (from the "Setup" dialogue), you may select a remote site using the universal naming convention (UNC): on the file-chooser dialogue, enter "\\<server>\<share>" in the "File name" field and press return (where <server> is the name of the remote machine and <share> is the exported directory hierarchy containing the database). This should give you access to the remote machine, from which databases may be selected as normal. You *don't* need to have mounted the remote directory onto a local drive letter, although this is more sensible most of the time.

(Thanks to Kent Sullivan of Microsoft for working this out.)

To my knowledge - on the basis of a *very brief* trial - WinRefer works quite happily in conjunction with the latest beta-releases of Windows NT. I intend to port WinRefer to run as a native NT application as soon as Visual Basic becomes available for these platforms.

#### **4.8. Saving Desktop Space**

The WinRefer main window is designed to present the necessary options in a clear and concise manner. For this reason, options controlling WinRefer's operation have been sequestered into a separate dialogue.

However, for some users it may be more advantageous to reduce the amount of desk space occupied by WinRefer: if, for example, all searches always use one keyword, there is little point in displaying the expression window or additional term boxes.

If the WinRefer window is re-sized below its default (maximum) size, various items are de-selected and "fall off" the window. At minimum, the window will simply display the menu and a single term box (or small expression box). This occupies only a very small amount of desktop space.

All the functions of WinRefer are accessible from the menus, so there is no loss of

functionality by using this approach.

#### 4.9. Memory

WinRefer will potentially use a lot of memory. Every index occupies memory, especially if cached; all references found in the current search are held in memory until cleared; the formatted text for a bibliography is kept in memory for the duration of a WinRefer run.

Memory management in Windows is not brilliant, and WinRefer's own scheme is by no means as sophisticated as it could be. WinRefer frequently "trades space for speed": things are kept in memory rather than being re-created at need. This improves performance, at a price of increased memory usage.

Memory errors are invariably fatal in the current version of WinRefer. Running out of memory will generate a warning message from the WinRefer memory management sub-system, and this message will immediately be followed by a system crash. I'm trying to handle this more gracefully, but for the moment, be warned!

In order to avoid running out of memory, the following tips may be helpful:

- o don't cache indexes between searches
- o try and be selective in searches, so fewer references are found
- o have a larger number of smaller reference databases, rather than a single monolithic file
- o buy more memory
- o increases the size of your swap file

Memory fragmentation - a rather technical subject which complicates memory management beyond all redemption - is also a big problem at present. Future versions of WinRefer - certainly any port to Windows NT - will be more memory-friendly.

At present *any* WinRefer crash (generating a general protection fault dialogue) will require Windows to be re-started before WinRefer can run again. This is because such crashes leave loaded libraries in an inconsistent state.

#### 5. Future Intentions

There are several ongoing improvements to WinRefer in the pipeline, work permitting.

The first is to support additional reference database file formats, to save users having to re-enter or duplicate their existing reference collections. The next new format will probably be **BibTeX** (which is well under way - I know I've been promising it for over year, but I'm trying, honest!), followed by **ProCite** (there's a new amendment to ProCite due out soon, and I'm waiting for it to be announced before doing anything about it). These new formats will be supported simply by distributing additional reference management libraries.

The second is to provide a mechanism for using a file of standard abbreviations for journal names: some journals require that publication names are abbreviated in a standard way, whilst others require a full name to be used. This feature would allow a format set to specify whether full or abbreviated names are to be used, with the translation being done through look-up into a user-supplied file.

Some improvements are pending for the formatting language, including better support for lists, improved tag generation, improved name formatting, and the ability to include additional fields in a formatted reference.

The final intention is to provide an additional application, separate from (but integrated with) WinRefer. It will allow format descriptions to be constructed graphically, rather than textually.

## 6. Acknowledgements

Many people have contributed to the (hopeful) usability of this release of WinRefer.

My thanks go first to Ken Tindell, Martin Atkins and Paul Butcher, without whom this program would never have been written. Ken especially made numerous suggestions about the user interface, and tried out all the combinations which I came up with. He is also chief bug-finder in the **troff** RML.

Many members of the InterNet community made very important suggestions, and were invaluable as bug-finders. At the risk of singling people out, I'd particularly like to thank the following people:

- o Mark Lee, for introducing me to several new reference file formats
- o Jim Garven, for hanging-in there when my stupidity introduced bugs
- o Tom Carter, for Chemical Abstracts
- o David Middleton, for Medline
- o Sue Stigleman, for some free publicity
- o Russell Malmberg, for pointing out a problem with index generation
- o Kent Sullivan, for confirming that WinRefer does indeed work across a network
- o Trevor Jenkins, for pointing out my gratuitous disregard for UI standards

Thanks in general to everyone who commented on WinRefer. I hope you'll keep using it, and keep the comments coming!

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

### A. The Various Formats of Reference Database

WinRefer supports three sorts of reference database: refer, Chemical Abstracts and Medline. Each of these is described briefly below.

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT19

A common feature of all the formats is that they are plain-text files - no "special" character codes are recognised, although this may change in the future to allow unusual characters to be included.

***PLEASE NOTE*** that there are copyright restrictions on the storage and use of references received from many commercial reference databases, and it is ***YOUR RESPONSIBILITY*** to remain within the law appertaining wherever you are. The author completely disclaims any responsibility for the mis-use of data obtained from any database.

You should ensure that all reference database files end with a couple of blank lines. WinRefer may otherwise ignore the last reference in a file.

In general, WinRefer tries to maintain as much information as possible about the references which it manipulates, but differences in the contents of the various styles of database make absolute consistency impossible. It should be noted in particular that the "Copy raw" button does not *necessarily* produce exactly the same data as was read-in from the database file.

### **A.1. refer**

A **refer**-format reference database is composed 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. The file **test\_ref.ref** contains a sample **refer**-format database.

WinRefer recognises the following codes:

- %Ttitle
- %A author
- %B book title
- %Eeditor
- %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 abstract
- %Zignored

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

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT20

%X contains an abstract: this is *not* defined explicitly by **refer**, but is tacitly assumed in all the major databases I'm aware of.

A variant of **refer** uses long names instead of letters - %Author instead of %A, for example - and such files will be handled correctly, as only the character immediately after the % is significant: other characters up to the first space are ignored.

## A.2. Chemical Abstracts

The Chemical Abstracts format is similar in many respects to the **refer** format, but uses different tags. The file **test\_ca.ca** contains an example of a Chemical Abstracts search. WinRefer recognises only a small sub-set of these tags:

L3	copyright
AU	authors, delimited by semi-colons
TI	title
PY	year of publication
SO	journal, volume, number and pages
CS	affiliation
LO	address
LA	language

Other fields are ignored.

Please note that this library is largely untested, and is thus very experimental.

## A.3. Medline

Medline is similar to Chemical Abstracts, but uses different tags. A sample may be found in the file **test\_med.med**. There are actually several formats which go under the Medline title, which may cause problems in some cases: WinRefer supports what I believe are the most common.. The tags recognised are:

AU	author
TI	title
LA	language
DP	date of publication
SO	source (journal name)
PG	pages
VI	volume identifier
MH	keywords
AD	address
AB	abstract

Other fields are ignored.

## A.4. BIDS ISI

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT21

The Bath Information and Data Service (BIDS) provides a citation index for various scientific and social journals and sets of conference proceedings. The data is supplied and owned by the Institute for Scientific Information Inc., and uses the same format as other ISI databases elsewhere.

Due to copyright restrictions, there is *no* example file supplied in this format.

The ISI format is similar to Medline:

TI	title
AU	authors
NA	authors' contact address
JN	journal and other details
KP	extra keywords
WA	author-supplied keywords
RF	research fronts
VO	volume
NO	number
PG	pages
PY	year of publication

Other fields are ignored, and WinRefer accretes all keyword fields (KP, WA and RF) into a single block.

ISI records are, in general, supplied in upper case. This makes the final formatted reference look very awkward. To save having to change the database, the RML automatically performs case conversion on ISI references. It tries to do a sensible job, but may fail to get things completely correct, so it's wise to double-check the formatting.

## 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	= <b>description</b> literal
language	= <b>language</b> { ( <b>plain</b>   <b>rtf</b> ) }
format	= <b>format</b> { inferencepart } as { formatpart }

inferencepart	= ( name   optionalinf ) [ ( name   optionalinf)... ]
name	= <b>Author</b>   <b>Title</b>   <b>Date</b>   <b>Journal</b>   <b>Volume</b>   <b>Number</b>   <b>Book</b>   <b>Editors</b>   <b>Pages</b>   <b>Publisher</b>   <b>City</b>   <b>Others</b>
optionalinf	= [ inferencepart ]
formatpart	= formatelement [ formatelement... ]
formatelement	= simplefield   literal   optionalclause   fontclause   formattedfield
simplefield	= ( name   <b>Tag</b>   <b>MajorAuthor</b>   <b>MinorAuthor</b> ) [ : n ]
literal	= " c [ c... ] "   <b>66</b>   <b>OpenQuotes</b>   <b>99</b>   <b>CloseQuotes</b>   <b>Tab</b>   <b>Para</b>
optionalclause	= [ formatpart ]
fontclause	= font { formatpart }
font	= <b>normal</b>   <b>bold</b>   <b>italic</b>   <b>bolditalic</b>
formattedfield	= namefield ( nameformat [ nameformat... ] )   datefield ( dateformat [ dateformat... ] )
namefield	= <b>Authors</b>   <b>MajorAuthor</b>   <b>MinorAuthor</b>   <b>Editors</b>
nameformat	= <b>Surname</b>   <b>Initials</b>   literal
datefield	= <b>Date</b>
dateformat	= <b>Century</b>
tagdescription	= <b>tag</b> { tagformatpart }
tagformatpart	= tagformatel [ tagformatel... ]
tagformatel	= formatelement   <b>Unique</b>
comment	= /* c [ c... ] */

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

## C. Changes Between 3.0 and 2.1

The changes are quite minor, though important. As usual, remove all traces of the old version of WinRefer before installing the new version (but keep a copy...). This includes

C:\USER\DOC\TEMPLATE\MYPAPERS.DOT23

old index files, as indexing has changed subtly.

Bibliography saving has been added, both manual and automatic. A consequence of this is that bibliographies are now stored in formatted form: this is a difference from 2.1, where bibliographies were formatted on demand (*i.e.* when copied).

The notorious "Out of stack space" bug in the bibliography generator has now gone.

Several internal limits have been increased at the request of people with reference databases of awesome size. There are still problems indexing very large databases, due to the issues mentioned in §4.9.

A sometime bug in sequential searching has been removed.

The main window may now be re-sized. All button functions are duplicated as menu items so as to be available in smaller windows. This means that all the functions are also available *via* keyboard short-cuts. I'm not sure that I've selected the most appropriate strategy to deal with re-sizing, but it seems to work okay.

Some minor re-arrangements have been made to some of the dialogue boxes.

Indexing hit-lists are now stored in a file.

I no longer have any intention of writing a reference database editor. This function is available in a program called Refs, and I'm more than happy to stick to formatting problems. This seems better than a monolithic, empire-building approach, and is entirely in keeping with the general Windows model of applications.