5th BILETA Conference
British and Irish Legal Technology Association

A Guide to Justus: an overview of a hypertext legal database

Eve Wilson

Computing Laboratory, University of Kent at Canterbury
Canterbury, Kent, CI2 7NF, England
Telephone: +44227 764000 exL 3628
Fax: + 44227762811
E-mail: ew@ukc.ac.uk

Keywords: hypertext, information retrieval, law, user interface, electronic dictionaries, statute law, law reports, indexing, cross-referencing.

Abstract: This paper discusses the need for an integrated electronic database for law, and.examines the interface requirements of a practising lawyer. It goes on to show how the system aims have been realised in the Justus system running under Guide, the hypertext system devised and developed at the University of Kent at Canterbury by Professor P J Brown. It concludes by looking to the future in research and development of a useful system.

1. The need for an electronic system for law

Law has long been a prime candidate for electronic information retrieval systems. There are many reasons for this:

1.1. Availability

All relevant primary sources in legal research are natural language texts which are printed, dated and of known origin. Consequently the database is comprehensive and uniform to a degree rarely found in other disciplines, in which research may depend upon observations and experiments in the physical world. To-day publishers use electronic media to produce traditional printed texts so recent material is available in computerised form even before it is published. The improvement in techniques for character recognition means that the contents of earlier printed sources can be captured by this means with little effort or expense.

1.2. Size

The volume of data is enormous and growth is exponential. The reign of Elizabeth I (1558-1603) produced only 272 statutes; the text of which fills some 200 pages in Ruffhead's *The Statutes at large* [Ruffhead 1786-1800]. In the reign of Elizabeth II, this is surpassed in a single year. For example 1985 saw the passage of 75 public general acts, running to some 3250 pages.

The explosion in law reporting has been even more spectacular. The proceedings of the courts from 1535-1580 are covered by only five sets of published reports: *Plowden, Dyer, Brooke, Bendilowe* and *Dalison.* [Abbott 1973]. The number of cases reported in this forty-five year period is surpassed in a single year by the commercial information retrieval system *Lexis,* which takes forty-two different series of law-reports totalling some 800-1200 cases a year. In addition to these *reported cases* transcripts from court stenographers add another 2500-3000 cases, the *unreported cases.*

EEC law and the activities of the European Court of Justice are growing with equal vigour. There is no prospect of amelioration.

1.3. Longevity

Legal data has a long life. The oldest statute still in force in England dates from 1267 *(Statutes of Marlborough 52 Henry III).* Statutes that have been repealed may still be required for cases and appeals pending, or for legal research, or to increase understanding of contemporary law reports. With law reports there is no equivalent of repeal: their life-span is indeteriminate. For example, *Methuen-Campbell* v *Walters* [1979) QB *525* CA considered Plowden's report *Hill v Grange* (1556)1 P1 164. A central on-line database can reduce the physical storage needed, and make all data available at any time however remote the source.

1.4. Maintenance and updating

The statutes are a dynamic data base the content of which is constantly being updated, frequently by amendment in a later Act. Keeping a printed manual up-to-date is arduous and prone to error. Loose-leaf bindings have been introduced but have only alleviated the problem.

Law reports are not updated but if they are cited in later cases the observations of the courts may be of crucial importance to a lawyer reading a report to find analogies to his current case. Sirnilarly, a lawyer fighting a case on a specific section of a statute needs precise information on any interpretation in the courts. Inverse referencing of this sort is currently achieved (if done at all) by sticky labels, and by subsequent volumes of case citations, which may come too late to be of use. A central databank where updating and cross-referencing are carried out automatically would be an obvious boon to busy practitioners.

1.5. Integration

The sheer volume of printed paper in a traditional system makes access wearisome. To research thoroughly even a straightforward case, a lawyer may need to access a dozen different locations in the law library. Some documents may not be available locally. To find all relevant data, he does not only need a thorough knowledge of the law, but a good understanding of legal bibliography and the relationships between disparate source materials. Hypertext provides a means of integrating documents so that for practical purposes they seem to be a single seamless information universe. To prove that such a universe was realisable was the prime aim of the Justus project; the second aim was to make the resulting system attractive to practising lawyers with no computer training or expertise.

2. Disincentives for the electronic system user

The growth in the use of electronic systems by practising lawyers has not been as rapid as the system providers would like. The first experiments in legal information retrieval were carried out in the 1950's, although serious use was probably not feasible until Butterworth made *Lexis* commercially available in 1980. Even since then, growth has not been rapid, particularly among smaller firms where the partners have no specialist librarian to do computer research for them. One reason is undoubtedly cost, but lawyers without easy access to a good law library find traditional methods of retrieval both costly and time-consuming. More fundamental disincentives are:

i) the need for a dedicated terminal which can be used for nothing else;
ii) difficulty in transferring the information they discover to another computer system such as a word processor for easy integration with other case notes;
iii) the complexity of the user interface.

3. The user interface

Lawyers are busy people for whom time is money. They have a high investment in the traditional skills of literacy and documentary research. They have run their practices without computers (save in the guise of wordprocessors) and are under no compulsion to use computers now unless they see benefits in doing so. They epitomize discretionary us ers. What qualities then must a user interface have to appeal to lawyers?

3.1. Ease of use

Lawyers want to devote l00 per cent of their intellectual effort to solving their legal problems; they do not wish to be diverted into solving computer problems that arise incidentally during their legal research. So a user interface must be easy to use.

Lawyers are going to be intermittent users; a lawyer with a case in court may not use the computer for days or even weeks. The system must be easy to learn and easy to remember. It must also be consistent in use: similar actions must produce similar outcomes regardless of context.

3.2. Multi-task capability

Consistency is not simply a requirement within an individual process. The system must use the same interface for all the computer tasks in a lawyer's office: wordprocessing, information retrieval, document drafting or expert systems. To-day lawyers are confronted with numerous free-standing systems, each of which performs a different task in complete isolation. Information obtained from a client and typed up on a word-processor has to be retyped for entry into a document drafting program. The draft document must be retyped into the wordprocessor before it can be formatted for printing. Information from the retrieval system cannot be easily transferred to another system for further processing. Expert systems have as many different interfaces as they have implementors. There is no incentive for the occasional user to invest effort in learning any of these; it is more cost effective to adhere to the traditional techniques of documentary research. If the system is to compete and win it must use the same interface for all the computer tasks in a lawyer's office: wordprocessing, information retrieval, document drafting or expert systems. Consistency is not simply a requirement within an individual process.

3.3. Familiarity

A lawyer's investment in the traditional techniques of legal research should not be underestimated. Methods have evolved over centuries, and new gimmicks with a computer are unlikely to modify them in the short term. Lawyers know what they want to do and how they want to do it. Computers ought to support them in their task, not demand changes in method to support computerisation. Many systems do not take account of the user profile and expectations. Easier access to familiar material in its traditional format may not be as technically exciting as a completely new knowledge structure to support an expert system but it may be more acceptable to the user.

Familiarity may not accelerate technological advance but it aids understanding and integrates more easily with the paper system (which is going to be around for a long time to come). For these reasons, it was decided to preserve the structures of the original documents in the hypertext system. A user should have at least the resources of an experienced reader in a good library but need have little investment in computing expertise.

4. Development of the Justus hypertext library

4.1. The Guide hypertext system

Guide [Brown 86], which was devised and developed at the University of Kent at Canterbury by Professor P.J. Brown, was chosen as the host hypertext system because it is ideally suited to supporting natural language texts of diverse structure. Not all hypertext systems can do this easily and naturally. What distinguishes a hypertext system from other document display systems is its ability to support a directed graph. Within that broad definition there are many realisations, not all of which lend themselves to our purposes.

A directed graph is composed of nodes; each node may be linked to any other node; the pattern of linkages is unpredictable. A familiar example of a directed graph structure in traditional printed books is an encyclopaedia. An encyclopaedia consists of articles of widely varying lengths. The linear ordering imposed on the articles by the traditional book format is entirely artificial. Once the reader has found his initial entry point he may be directed to another entry by the compiler of the encyclopaedia, or choose another article on his own initiative because he wants further information on a topic in the original entry. His path through the book is unpredictable, but he usually wants to refer to only a small portion of the whole at any one time. In a hypertext system, each article in the encyclopaedia could be represented as a node; the cross references to other material are links in the directed graph. It is these links that determine the true structure of the encyclopaedia, not the artificial linear ordering which is only used to determine the initial access point. For the hypertext system to provide a natural mapping of the encyclopaedia, it must be able to support linkages within and between documents. Ideally it should also support a variable node size. Some entries may be terse:

> ABDUCTOR: see MUSCULATURE.

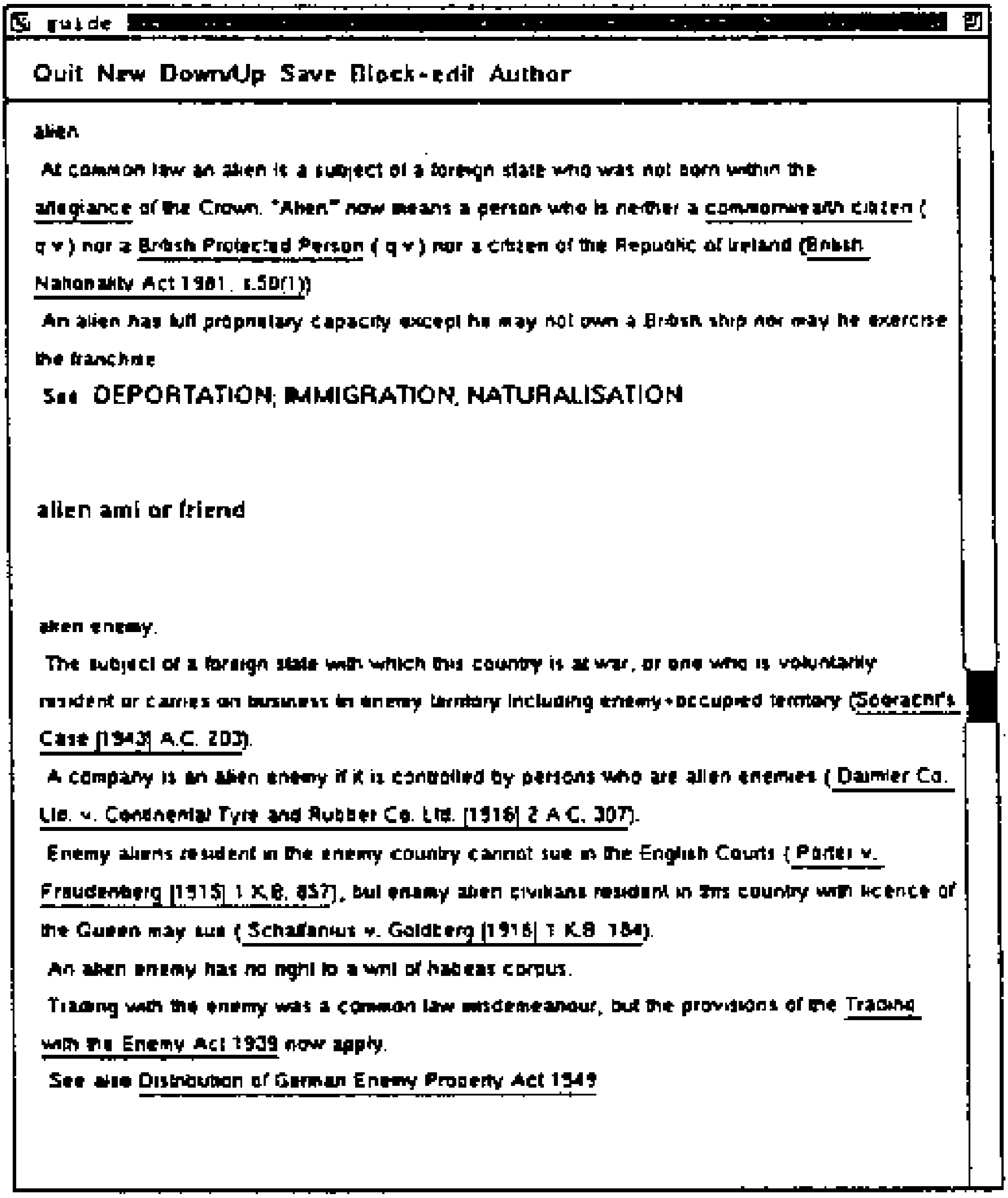
Others may run to many pages.

Not all hypertext systems support a variable node size or permit backwards and forwards scrolling of text. Those that do are better for mirroring the forms and facilities of traditional text.

4.2. The source documents: the dictionary

Because the integration of documents into a single hypertext book was a prime aim of the system, and because this integration can only be achieved automatically by recognising, by computer program, cross-references already in the text, I decided that the first book I should try to convert should be - of the machine readable texts available to me - the text with the greatest density of cross-references. This was Osborn's Law Dictionary generously made available by the publishers Sweet and Maxwell. Not only are there many cross-references to other entries in the dictionary; the reader is frequently referred to statutes and to case reports.

The entries in figure 1 show the density and variety of the cross-referencing.



alien

At common law an alien is a subject of a foreign state who was not born within the allegiance of the Crown. "Alien" now means a person who is neither a commonwealth citizen (q v) nor a British Protected Person (q v) nor a citizen of the Republic of Ireland (British Nationality Act 1981, s.50(1))

An alien has full proprietary capacity except he may not own a British ship nor may he exercise the franchise

See DEPORTATION; IMMIGRATION, NATURALISATION

alien ami or friend

alien enemy.

The subject of a foreign state with which this country is at war, or one who is voluntarily resident or carries on business in enemy territory including enemy-occupied territory (Sovfracht's Case [1943] A.C. 203).

A company is an alien enemy if it is controlled by persons who are alien enemies (Daimler Co. Ltd. v. Continental Tyre and Rubber Co. Ltd. [1916] 2 A.C. 307).

Enemy aliens resident in the enemy country cannot sue in the English Courts (Porter v. Freudenberg [1915] 1 K.B. 857), but enemy alien civilians resident in this country with licence of the Queen may sue (Schaffenius v. Goldberg [1916] 1 K.B. 284).

An alien enemy has no right to a writ of habeas corpus.

Trading with the enemy was a common law misdemeanour, but the provisions of the Trading with the Enemy Act 1939 now apply.

See also Distribution of German Enemy Property Act 1949

Figure 1: Dictionary showing cross-referencing

In the two entries that have been expanded, *alien* and *alien energy* there are references to other entries in the dictionary (e.g. *allegiance* DEPORTATION), to Acts of Parliament (e.g. *British Nationality Act 1981, s.50(1)),* and to cases (e.g. *Daimler Co. Ltd* v *Continental Tyre and Rubber Co. Ltd* [1916] 2 A.C. 307). These have been converted by program into *link icons* or, in Guide terminology, *buttons.* When the button is selected with the mouse by the user, the hypertext system searches for the text associated with that button. The most frequently used buttons are shown in figure 1.

1) Definition buttons

Each entry in the dictionary has become a text node in the hypertext system. The *label* or, in Guide, the *definition button* associated with each node is the dictionary headword for that entry. When the definition button is selected the associated text is expanded in place on the display and the screen reformatted to accommodate it.

2) Usage buttons

Definition button names or icons can appear anywhere in the text. If they are recognised and converted into *link icons* or, in Guide, *usage buttons* then when the user selects the usage button the hypertext system searches for the definition button with that name so that it can display the associated text. Two types of usage button are shown in figure 1:

> i) Replacement buttons
>
> These are shown as bold face. A copy of the text corresponding with the button will be substituted for the button in the display and the screen reformatted to accommodate it - as in a definition button.
>
> ii) Glossary buttons
>
> It is not always appropriate to replace a button in situ; it interrupts the flow of text and could easily mislead or confuse the reader. *Glossary buttons* are used to prevent this. When a glossary is selected, Guide creates a subwindow and the replacement text for the button is displayed in the subwindow, not the main window.

Work on the dictionary is described in more detail in [Wilson 88a].

5. Statutes

As a testbed for developing a program to recognise cross-references automatically the dictionary was a great success. References to cases and to Statutes could be automatically recognised and highlighted. References to statutes in particular may be complex in structure and in semantics. As well as references to a single text node:

> *Trading with the Enemy Act 1939*
> *British Nationality Act 1981 s50(1)*

there are structurally complex reference lists:

> *Ecclesiastical Jurisdiction Measure 1963 ss2,8,13,82*
> *Companies Act 1967 ssl4 (8) (c), 130, Sched 8, Part III*

and extended references:

> *Consumer Credit Act 1974 ss43-58*
> *Supreme Court Act 1981 s108 (1) - (5).*

There may also be combinations of these types:

> *Children and Young Persons Act 1969 ss 35-59, Sched 8, Sched 6*

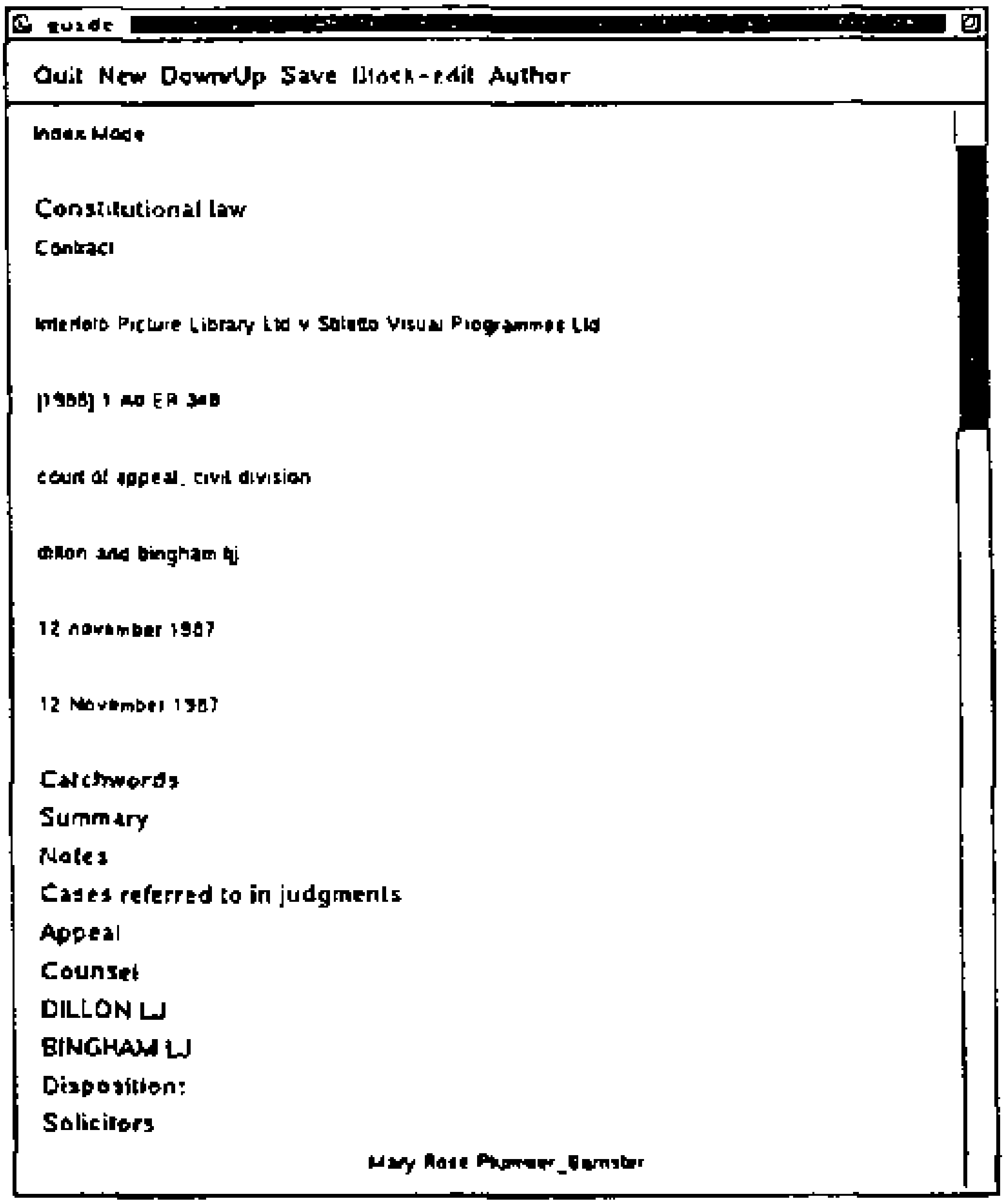
The Justus programs patently produced buttons for all these nodes and the buttons were, as far as could be determined by manual checking, correctly aliased. (Aliasing is a way of mapping different phrases or terms with the same semantic meaning on to the same text node.) Frustration set in when demonstrating the system because only references internal to the dictionary could be satisfied when selected. So that the multifaceted references to Statutes could be seen to work, it was necessary to add statutes to the database.

I chose to add industrial relations statutes because I had by me a collection of statutes and cases from an earlier project in the 1970's. The statutes were easy to convert to hypertext form. They are highly structured documents and already labelled by the parliamentary draftsmen down to sentence level. Guide's variable node size was invaluable here. A statute can be represented as a single definition node; but the node is composed of other definition nodes which represent parts. The part nodes are aggregations of section nodes, and the section nodes comprise subsection nodes, and so on, until the text is reached. While structurally more interesting than the dictionary, statutes lacked variety in explicit cross-references. A statute that has not been annotated refers only to itself and to other statutes. Statutes were invaluable for proving implicit referencing methods.

Statutes often contain *statutory definitions,* i.e. terms that have been assigned a precise meaning within that statute. A reader cannot properly understand an Act if he is unaware of special meanings assigned to terms used in it. Justus recognises statutory definitions by their form, then converts all occurrences of the defined terms into glossary buttons. When the user selects one of these glossary buttons the corresponding definition is displayed in the glossary subwindow.

6. Cases

The first case reports introduced to the system were invaluable in proving the capacity of the system to integrate data from different sources and with different formats. They did not excite me as data because they were not from a reputable publisher in a standard format and there was no hope of more from the same source. At this point, the law publishers, Butterworth, gave me access to a machine readable version of the *All England law reports 1988.* These were exciting. The All England law reports started in 1936 and, with the reprint series of selected cases from the sixteenth century onwards, form a comprehensive collection. They are highly structured and the structure is so comprehensive that reports from other series could easily be mapped onto subsets of the structure. A skeleton structure is shown in figure 2.



Index Mode

Constitutional law
Contract

Interfoto Picture Library Ltd v Stiletto Visual Programmes Ltd

[1988] 1 All ER 348

court of appeal, civil division

dillon and bingham ljj

12 november 1987

12 November 1987

Catchwords
Summary
Notes
Cases referred to in judgments
Appeal
Counsel
DILLON LJ
BINGHAM LJ
Disposition:
Solicitors

Mary Rose Plummer_Barrister

Figure 2: Case skeleton

Short components are shown in full: name, report reference, panel, hearing date, date (ie. date when the decision of the court was delivered). Other components appear as buttons. The replacement text can be obtained by choosing the appropriate button. The *catchwords* section is a list of manually assigned key terms, and these are also used to produce an index that can be used by the reader to select an appropriate entry point to the case database. *R* v *Sharp* was found from the index term *criminal evidence.*

The opinion of each judge can be retrieved by selecting the appropriate bold-face name, e.g. **LORD MACKAY OF CLASHFERN LC.**

A complete description of this work is in [Wilson 90].

7. Multiple versions and reference inversion

Up to this point great care had been taken to preserve the text in its original form without addition or deletion. The user was to have access to exactly the information available to him from printed books but without the inconvenience of having to locate the books, access an index at the back of the book, and find the place in the text that he wanted. Clearly, information could more readily be incorporated into an electronic system than a printed book system. The opportunity to resolve two problems intractable in the manual system was too tempting to resist. The problems arise out of referential legislation and case citations.

7.1. Referential legislation

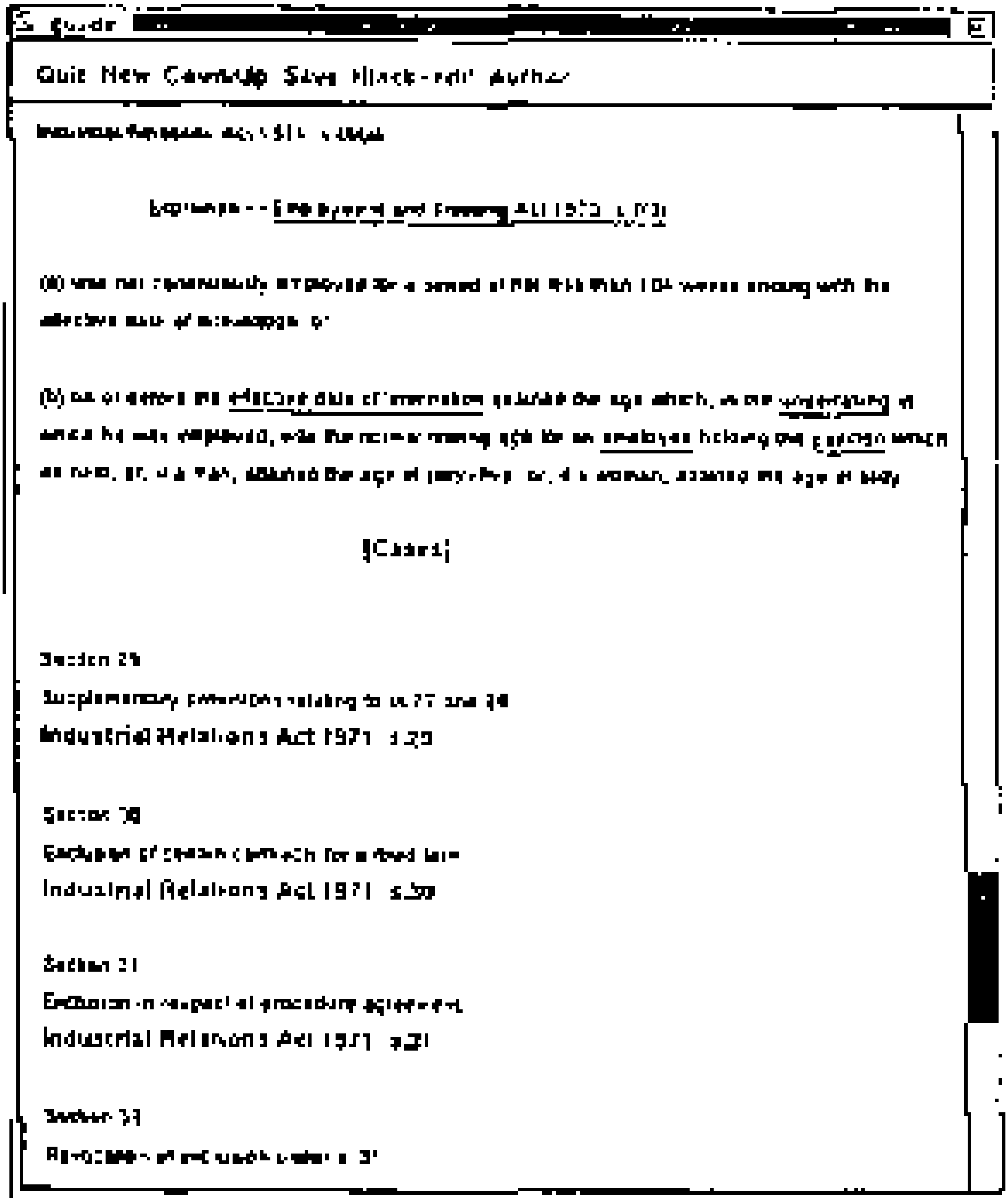
Much legislation in Britain is amended merely by reference from a later statute. The difficulty of incorporating these amendments into printed books can easily be appreciated. Moreover, the original text of amended or superseded legislation cannot be discarded because cases still pending may turn upon the original text. For example, a Finance Act is passed annually but tax claims may be made up to six years in arrears and disputes may involve access to even earlier enactments. Justus can keep any number of versions. The current version is shown first; the reader can have instant access to previous versions on request.

7.2. Case references

In a common law system it is of paramount importance to a lawyer that he should be aware of cases that refer to statutory legislation relevant to his current case, or of cases that refer to an earlier relevant case. This can easily be incorporated into an electronic database. In the statutes it has been done by giving each node referenced a *[cases]* button. When selected this will give access to relevant cases - which are themselves buttons that may be selected and so on. There is no limit to the length of the reference trail.

Figure 3 shows s.28(a) of the Industrial Relations Act. Figure 4 shows the screen after the two buttons *Employment and Training Act 1973 s.7(3)* and [cases] have been selected. Note that since *Employment and Training Act 1973 s.7(3)* is a glossary button the text appears in the subwindow.

For a fuller description see [Wilson 89].



Industrial Relations Act 1971 s.28(a)

Exceptions -- [illegible] Act 1973 s.[illegible]

(a) was not continuously employed for a period of not less than 104 weeks ending with the effective date of termination, or

(b) on or before the effective date of termination attained the age which, in the undertaking in which he was employed, was the normal retiring age for an employee holding the position which he held, or, if a man, attained the age of sixty-five, or, if a woman, attained the age of sixty.

Section 29

Supplementary provisions relating to ss.27 and 28

Industrial Relations Act 1971 s.29

Section 30

Exclusion of certain contracts for a fixed term

Industrial Relations Act 1971 s.30

Section 31

Exclusion in respect of procedure agreements

Industrial Relations Act 1971 s.31

Section 32

Revocation of exclusion order under s.31

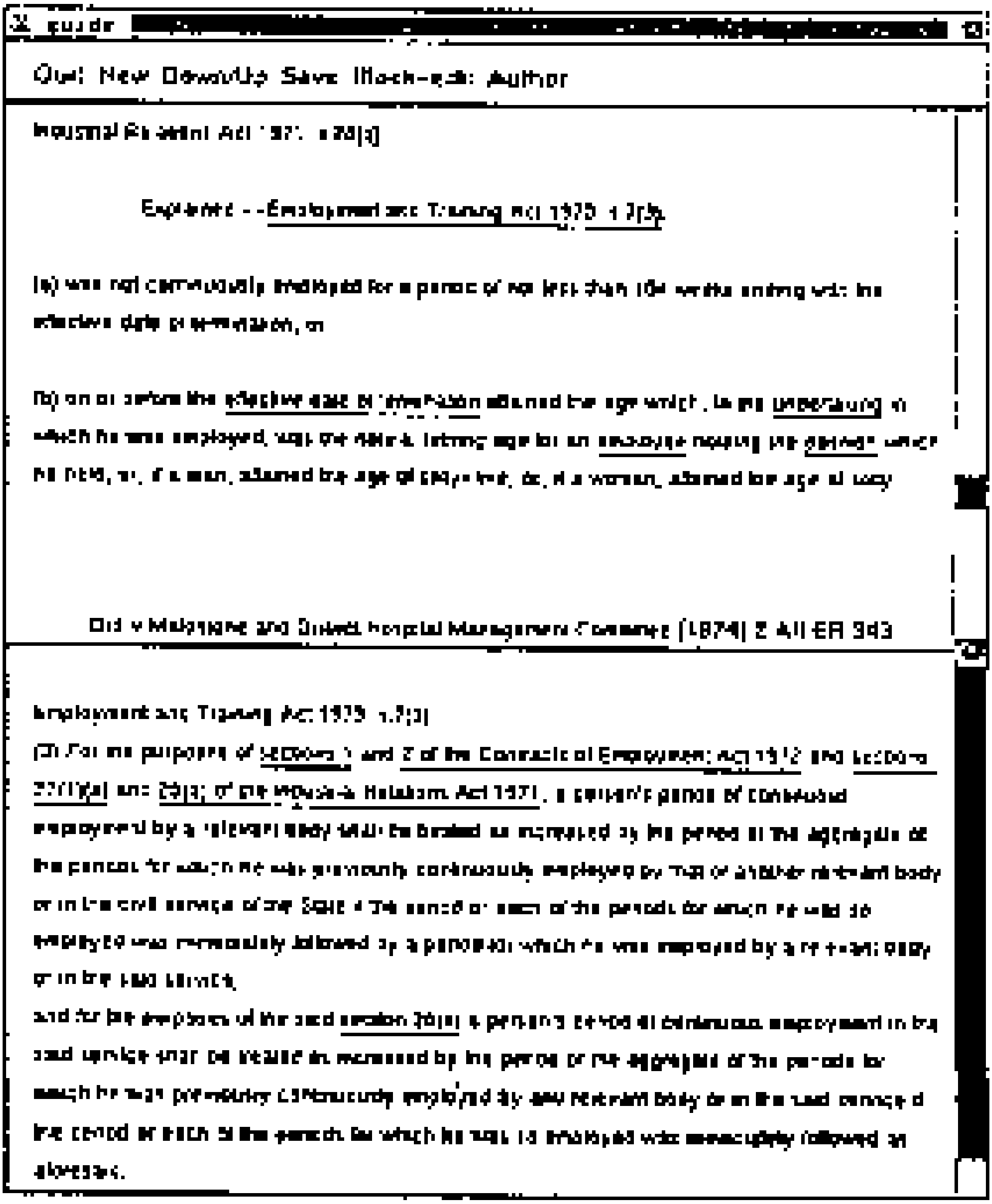


Figure 3: Statute *with* added references Figure 4: References have been selected

8. Textbook

Academic lawyers who had seen the Justus system realised its potential as a learning environment. To make this application more evident portions of a textbook, Selwyn's *Guide to the Industrial Relations Act 1971,* have been incorporated in the database. The mode of incorporation mirrored that of statutes, with chapters, sections, subsections and paragraphs as the basic node types. The only point that I would like to make here about the textbook is that its conversion to a hypertext document was enormously simplified because intra-text references were already given in paragraphs - a logical unit of the text - rather than by page number, which wholly depends on the physical medium of reproduction. Editors and publishers please take note!

A description of how the textbook integrates with the system in use appears in [Wilson 88b].

9. Precedents and interactive drafting

The legal texts currently under investigation are precedent books, in particular, and with the encouragement of the publishers, Butterworth's *The Encyclopaedia of Forms* and *Precedents.* Precedents are different from the previous texts. With those it was tacitly assumed that the lawyer would use them as he used printed texts but with greater ease of access. He may want to take a copy of a paragraph of a textbook, a section of a statute or a judge's opinion. He could, of course, add his own notes at any point and save the file for further research on another occasion or have it printed out for later perusal, but such gleanings were entirely for his own use and not expected to form polished coherent documents in their own right. However, a lawyer consults a book of precedents because he needs to create an original document by combining relevant clauses in the appropriate order. Into these clauses he wants to introduce material of his own. At a minimal level this may be the names and addresses of parties to a contract, dates, values, or descriptions, but he may also want to add whole clauses of his own devising. This requires a greater degree of interaction and flexibility than a straightforward information retrieval/browsing affords. It can be achieved only by embedding an 'expert system' within the hypertext interface. The initial version of this system is a drafting tool, which will:

1. help the user choose the precedent most suited to his present purpose.
2. allow him to insert information into present clauses.
3. deal with pronoun gender problems for named parties to a contract.
4. allow him to reject clauses he doesn't want or to add new clauses while the system maintains a clause count.

10. Conclusion

While data available to the system still falls short of the requirements for use in the real world enough has been done to prove the techniques and to show the possibilities of the system as a tool for practising lawyers and as a teaching resource in further education. Response from lawyers within and without the academic community has been enthusiastic. Even computer-unfriendly diehards in the legal profession have enjoyed experimenting with it. Without substantial investment a comprehensive system is not realisable, but Justus on a workstation dedicated to a cohesive subset of law, such as Wills and Probate, is a realistic and worthwhile goal. Meanwhile many interesting options remain to be explored. A priority is the further development of the interactive drafting tool. Work on making more efficient use of the manually produced key words and on improving retrieval is also underway.

Acknowledgments

I am indebted to Butterworths Ltd. for permission to use material from the *All England law reports* and the *Encyclopaedia of Forms and Precedents,* and, in particular, to Mr. L. Schapira, of Butterworths, for support and encouragement. Thanks are also owing to: Mr. R.C. Saunders, for discussions and help throughout the course of the Justus project; Dr G. Raines, for reading and offering advice on the original draft.

Bibliography

Abbott, L. W., 'Law Reporting in England 1485-1585', Athlone Press, 1973.

All England law reports, Butterworth, 37 vols., 1558-1935, issued 1957-67.

All England law reports, Butterworth, 1936-

Bell Laboratories, 'UNIX Time-sharing System: UNIX Programmer's Manual, Seventh Edition, Volume 1, Murray Hill, N.J., 1979.

Brown, P. J., 'Interactive Documentation.' *Software Practice and Experience,* Vol.16(3), March 1986, pp.291-299.

Encyclopaedia of Forms and Precedents, Fifth Edition, [edited by] Sir Raymond Walton, Butterworths, London, 1986.

Ruffhead, O., *The Satutes at large,* Eyre & Strahan, 1786-1800.

Wilson, E., 'Integrated Information Retrieval for Law in a Hypertext Environment' Proceedings of the SIGIR/ACM 11th. International Conference on Research and Development in Information Retrieval, Grenoble, 1988a.

Wilson, E., 'Electronic Books: the automatic production of hypertext from existing printed sources', *Proceedings of the Fourth Annual Conference of the UW Centre for the New Oxford English Dictionary, Information in Text,* Waterloo, 1988b.

Wilson, E., 'Reference and Reference Inversion in Statutes and Cases; a hypertext solution', *Informatics 10,* York, 1989.

Wilson, E., 'Cases for Justus: preparing a case database for a hypertext information retrieval system.' To be published later in 1990 in the *Journal of the Association of Literary and Linguistic Computing*.