Dynamic Tailoring of Law Related Documents To User Needs

C.A. Royles and T.J.M. Bench-Capon
LIAL - Legal Informatics at Liverpool
Department of Computer Science
The University of Liverpool
Liverpool,
England

Abstract

In this paper we describe a system which allows for documents available on the World Wide Web to be dynamically tailored to meet the particular needs of different users. We describe the need for such a system, and the architecture of our system. We give details on some of the particular techniques used in our system. We illustrate our system with an example relating to a UK Social Security benefit. We conclude with some discussion of the place of such systems in the future of Legal Information Systems.

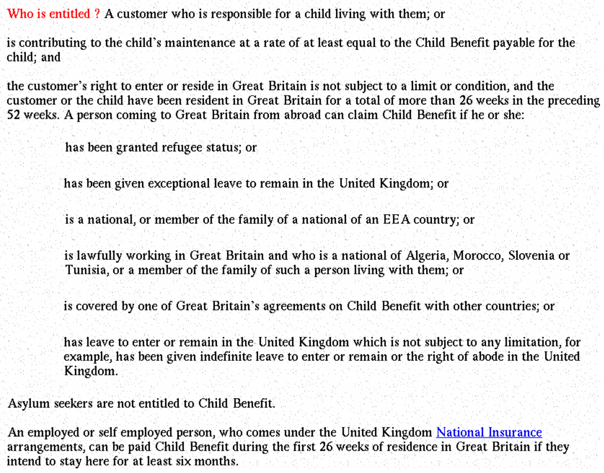
1. Introduction

An increasingly large volume of law related material is available on the World Wide Web (WWW). For example, information relating to UK Social Security Benefits is supplied by the Benefits Agency [1]. Such pages suffer from many of the problems associated with information leaflets which serve a similar purpose, identified in [2]. Because they must cater for a wide range of different readers with different information needs, different levels of skills, and with different circumstances, they must attempt to be all things to all people. The result is an inevitable compromise which:

- omits information required by a minority of readers; either because their task has some esoteric needs or because their background knowledge is such that they require more detail and explanation than the targeted reader;
- includes information irrelevant to a given reader. Because the document attempts to cover all circumstances, readers must be able to identify what does, and what does not, apply to their particular circumstances; because the document attempts to support all tasks, the reader must select the pertinent information;
- uses terminology which may be inappropriate to certain readers;
- is organised in a structure which is suited to some readers and tasks more than others.

The readers must therefore exert considerable effort to extract the information required, and to apply that information to their particular circumstances and needs.

In the case of paper documents there is little that can be done about this, short of producing a range of documents which are broadly similar but directed at different readers and tasks. This does, of course, occur to some extent, but even so such documents remain a compromise, and there is the added danger that particular readers may find themselves with inappropriate versions. With electronic documents, however, there is the possibility of processing the document so as to give it a different appearance, depending on some characterisation of the reader. In this paper we describe a system which will undertake such dynamic tailoring. We will illustrate the system by reference to documents providing advice on Social Security benefits.



Who is entitled ? A customer who is responsible for a child living with them; or

is contributing to the child's maintenance at a rate of at least equal to the Child Benefit payable for the child; and

the customer's right to enter or reside in Great Britain is not subject to a limit or condition, and the customer or the child have been resident in Great Britain for a total of more than 26 weeks in the preceding 52 weeks. A person coming to Great Britain from abroad can claim Child Benefit if he or she:

has been granted refugee status; or

has been given exceptional leave to remain in the United Kingdom; or

is a national, or member of the family of a national of an EEA country; or

is lawfully working in Great Britain and who is a national of Algeria, Morocco, Slovenia or Tunisia, or a member of the family of such a person living with them; or

is covered by one of Great Britain's agreements on Child Benefit with other countries; or

has leave to enter or remain in the United Kingdom which is not subject to any limitation, for example, has been given indefinite leave to enter or remain or the right of abode in the United Kingdom.

Asylum seekers are not entitled to Child Benefit.

An employed or self employed person, who comes under the United Kingdom National Insurance arrangements, can be paid Child Benefit during the first 26 weeks of residence in Great Britain if they intend to stay here for at least six months.

Figure 1: Part of Child Benefit Page

Figure 1 shows part of the WWW page relating to Child Benefit. What is immediately apparent is that the reader is confronted by a large number of rarely applying conditions. These separate the basic entitlement conditions from such more commonly sought information as the definition of "child" and the rate of the benefit. If the reader's circumstances were known, all the conditions about coming from abroad could be suppressed, except any that actually applied. The result would be a much more accessible piece of text.

2 System Overview

What we require to tailor the document is the ability to gather information from the user and to relate this information to appropriate modifications to the text. We do this using knowledge based systems techniques. The document to be tailored is supplied with additional special HTML tags expressing rules which state that certain *effects* are to be applied to the block of text if certain conditions are met. The rule may also establish certain other conditions that are used by other rules. When a condition is encountered, if its status is not already known, the user is queried for the information. Thus in the example above it would be first established if the user had come from abroad. If not all the special conditions applying to arrivals would be suppressed. If the user was a recent arrival, further questions would determine which particular paragraph was applicable, and the others would be suppressed. We have a two pass process, the first pass determining the effects to be applied, and the second applying these effects to display the tailored document to the user.

The key modules in the system are:

- *Proxy server*. - In order to establish a complete and fairly standard method of intercepting internet documents in transit, the network communication side of the application is through a proxy server. Using a proxy server is perhaps the most technically interesting decision, and this is discussed further in section 5.
- *Cache and File Manager*. - Files retrieved will require caching in order to improve the latency of file requests. As well as file caching, preferences and session files recording the information gathered from the user need to be stored. These operations are handled by a file manager.
- *Document Parser*. - A general file parser is used to identify the different HTML tags within a document. A separate module within the parser then passes the processing of specific tags off to a separate tag handler routine that can be dynamically loaded into the system. This allows us to extend the tags that can be dealt with if we need to do so.
- *Knowledge Base Manager*. - The knowledge base and its manager are the core of the application. They store the dynamic knowledge base and determine how the rules are handled and which queries will be asked of the user. The inference engine is included here to determine how the rules are handled and combined.
- *Query Manager*. - The query manager determines from the Knowledge Base Manager which queries are required. The corresponding HTML file is then loaded. On submission of the form supplied by the query, the Query Manager parses the information and returns it so that it can be added to the Knowledge Base.
- *Effect Manager*. - The Effect Manager is the final stop for the HTML document, Here the original document is loaded from cache and the appropriate effects are applied.

A diagrammatic view is shown in figure 2.

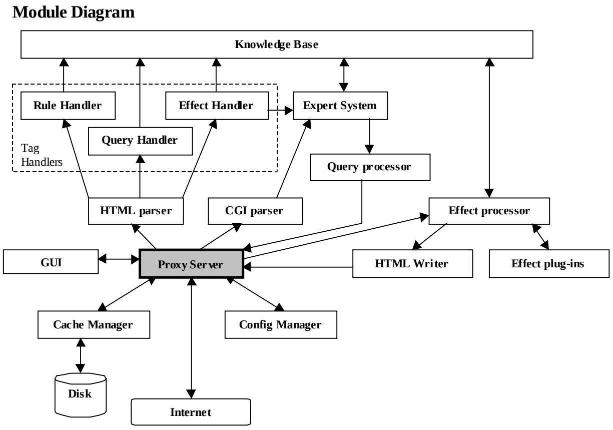


Figure 2: System Architecture

A dynamic view of the process is shown in the event diagram in Figure 3. Using this diagram we can trace the progress of a request from a user. From the perspective of the user; a document is requested, some questions are asked and answered and then the tailored document is displayed on the internet browser. From the perspective of the WWW Server, first the original document is requested, and then potentially some further documents representing queries may be requested. Note that requested documents are cached by the Proxy Server to improve response time and minimise net traffic. The proxy server parses the documents, identifies the queries to pose, applies the rules, identifies and processes the effects, and in general acts as a go-between for the browser and the WWW server, whilst tailoring the document to the required form.

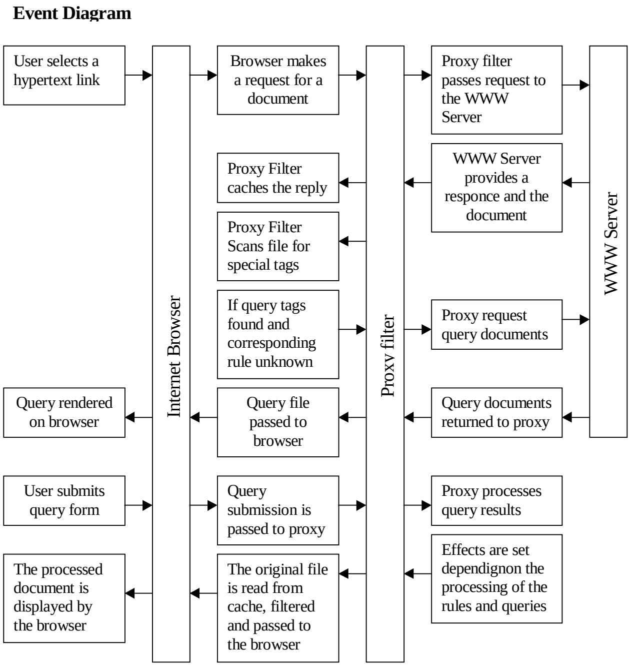

Event Diagram
User selects a hypertext link
Internet Browser
Browser makes a request for a document
Proxy filter
Proxy filter passes request to the WWW Server
WWW Server
Proxy Filter caches the reply
WWW Server provides a responce and the document
Proxy Filter Scans file for special tags
If query tags found and corresponding rule unknown
Proxy request query documents
Query rendered on browser
Query file passed to browser
Query documents returned to proxy
User submits query form
Query submission is passed to proxy
Proxy processes query results
Effects are set dependignon the processing of the rules and queries
The processed document is displayed by the browser
The original file is read from cache, filtered and passed to the browser

Figure 3 Event Diagram

3. Rules

The rules used in the system are based on propositional logic. This is considered adequate for the information that we need to represent. A rule consists of :

- a set of *conditions*. These may be either conjunctive or disjunctive.
- a set of *effects*. These determine how the text will be treated if the conditions are satisfied. A more detailed discussion of effects is in section 4.
- zero or more *consequences*. These are propositions which can be deduced from the truth of the conditions, and which there is therefore no need to query the user for.

Each condition which the user may be requested to supply is associated with a *query*. This is a document which offers flexibility in the form in which the question is posed, and further allows for the posibility of including guidance on the interpretation of the query. For example, if the user is asked if he has a child, we may wish to include in the query document the information that a child has to be either under 16, or under 19 and in full-time education. Because the query is itself a document there is the option to include rule tags to tailor the presentation of such information on the basis of information previously supplied. Queries are posed to the user only when they are required to evaluate the conditions of some rule, and of course, only when the information is not already available as the answer from a query or as the consequence of some satisfied conditions.

4. Effects

The effects are the modifications that can be made to the document. At present they include the following:

- Text can be *suppressed*: ideally readers should see only those parts of the document that are relevant to their particular circumstances, and their particular information needs at the time they access the document;
- Text can be *highlighted*: Devices such as bold font, smaller or larger font sizes, use of white space and the like can greatly help readers to understand a document by drawing their attention to the critical parts of a document; what needs to be emphasised can again vary;
- Text can be *included*: Sometimes it may be helpful for text from another document to be imported into a document to supply critical background information, or to expand certain cryptic references;
- Text can contain *links*: If we don’t want to overload the readers with background information in the document itself, it may still be helpful to indicate that background material is available, and to give easy access to it. What background material is best made available will again depend on the particular reader;
- Text can be *substituted*: Different readers may have preferences for different vocabulary and phraseology. Readers understanding will be enhanced if the vocabulary and phraseology is appropriate to them.
- Text can be *re-organised*: Different readers will start with different background knowledge, and may find different orderings of the material presented helpful.

These effects can be deployed individually or combined, so that a piece of included text could be highlighted and given a link, for example.

The architecture diagram in Figure 2 also shows a module for effects “plug-ins”. This allows for the possibility of including multimedia effects such as sound or video, should they be thought desirable.

5. Why a Proxy Server?

In this section we will discuss our decision to use a proxy server. There are a number of options for providing interactive services across the Internet, including: Common Gateway Interface (CGI), Java, javascripts (either client side or server side), and remote browser control through an API. The major differences in these techniques turns on whether the program code is executed on the remote or local machine. There is a third option for providing control and interaction with the user and that is to use a proxy server. A proxy server is a program that can be asked by a browser for some information, which then forwards that request to the correct Internet file server. The response is passed back through the proxy to the browser. Proxy servers are typically used to provide services such as firewalls, chaching, filtering and redirection. Normally a proxy is located on a remote machine, but in our case the server will be located on the client machine.

We locate the proxy server on the client's machine for several reasons. The proxy server has access to the user's machine hardware such as memory, disk, and processor. This access allows the program to save files or session information, provide more functionality and a complete graphical user Interface and run at the speed of the client's machine. The client machine is usually slower in real terms than a dedicated Internet file server, but faster than a shared Internet file server which may share its processor power between several users at the same time.

Advantages from this solution include:

- Dynamically generated session information can be stored on the users machine. This information is thereby unique, secure and easily accessed. In other solutions storing session information is difficult .
- Information storage on the users hard disk allows for a greater volume of storage and the ability for the user to selectively save and load particular session files into the proxy engine.
- There is a trade off regarding file transfer over the Internet. Either the source information is downloaded and processed in its entirety on the clients machine, or each query has to be sent over the internet producing a delay between each request.
- Information can be cached on the users machine by the proxy providing much faster access to information, once retrieved over the internet, it is available locally until removed from the cache. This approach allows pre and post processed files to be stored and provides an easy way of monitoring alterations or versions of the same file.
- The application may be processor intensive and it is impractical to allow the internet file server to handle several parallel requests where such processing will be required.

There are some small disadvantages to using a proxy server:

- A very small delay in file transfer as the information is passed through, but this delay would occur in any filtering system.
- The user must download and install the proxy server on their own machine before they can benefit from its functionality.
- The browser or the links to the documents have to include information that directs the request through the proxy server. If the proxy server is not installed this could result in a bad link.

We feel that these disadvantages are far outweighed by the advantages of using a proxy server.

6. Potential of Document Tailoring Systems

We believe that the style of system we have described here will have a considerable impact of the ways in which legal information systems will be conceived, especially those targeted at members of the public.

The important consequences of integrating database and hypertext (and hence WWW) technology into legal knowledge based systems have been recognised for some time. A fairly comprehensive discussion can be found in [3], which discusses a range of integration options. Some specific issues relating to the delivery of expert systems over the WWW are discussed in [4]. As yet, however, expert systems made available over the WWW provide little more functionality than stand alone consultative expert systems, other than to provide hypertext links in the explanation phase. In fact the consultative model has not proved particularly successful for legal expert systems. What is required, we believe, is not simply the integration of available technologies using the existing expert systems model, but rather a rethinking of the model, in the light of the new possibilities opened by the technology.

Perhaps the most commercially successful fielded system, in the UK at least, is the Retirement Pension Forecast and Advice system[5], originally developed for the Department of Health and Social Security in the late 1980s. What distinguishes this system from other knowledge based advice systems is its interaction. The user requests advice by post; no further questioning is required because all the relevant information is already held on databases within the Department. The user then receives back, again by post, a letter (typically running to about three pages) giving a forecast of likely pension entitlement and advice, customised to the user's own particular situation. Thus the interaction is essentially a request for a *tailored document*, with the document being constructed from pieces of boiler plate text determined by the rules of a knowledge based system. It is essentially this model that we are adopting for our proposed system, with the written documents distributed by mail being replaced by electronic documents delivered over the WWW, and the role of the database being taken by user replies to dynamically generated questions. Of course, were the information available from an accessible database, the queries could be sent to the database rather than to the user.

The advantages of this style of system derive from the fact that what the user sees is in a familiar form, but which is nevertheless directed at a personal set of circumstances, while the mechanics by which the tailoring is determined do not intrude. Thus the approach combines the merits of the general advice document and the specific expert system, while avoiding the pitfalls of the over generality of the advice documents, and the clumsy interaction of the expert system.

Advice documents are not the only type of documents that we might wish to tailor. A long standing source of applications in legal information systems concerns document generation. Perhaps the earliest of such systems is the will processor of Sprowle [6]. Other examples are provided in the fields of contract drafting [7] and show-cause orders [8]. Our approach could be applied to this kind of task also, although the difference in our approach is that the text is not regarded as a set of boiler plate paragraphs used as data by the system, but that the rules are embedded within the text itself. [8] distinguishes between the representation of rhetorical and illocutionary structure, both represented as rules; in our approach the rhetorical structure is supplied by the document itself, and only the illocutionary structure by rules.

We could go further: many systems which supply decision support in the legal domain such as HYPO [9] and PLAID [10] present their information in the form of a structured document or brief. Again interesting possibilities could arise if we conceive of such systems as tailoring rather than generating a document.

As we move from advice documents through standard documents to briefs we increasingly emphasise the knowledge based content at the expense of the text. None the less, all these applications can be seen as tailoring an output text, and our approach could be used to accommodate them. This approach has particular and obvious advantages if the systems are intended to be delivered over the WWW.

7. Conclusion

In this paper we have described an approach to integrating knowledge based and WWW technology to providing individual advice through tailored documents. We believe that the approach has potential which goes beyond advice giving applications, and which could be used as the model for other legal information system tasks.

References

[1] The URL for general benefit information is http://www.dss.gov.uk/ba/GBI. This gives access to more specific pages dealing with individual benefits.

[2] Soper, P., and Bench-Capon, T., (1994), *Coupling Hypertext and Knowledge Based Systems: Two Applications in the Legal Domain*, Artificial Intelligence and Law, Vol 2, pp293-314.

[3] Greenleaf, G., Mowbray, A., and van Duick, P., (1995), *Representing and Using Legal Knowledge in Integrated Decision Support Systems: DataLex Workstations*, in Artificial Intelligence and Law, Vol 3 Nos 1-2, pp 97 -142.

[4] Greenleaf G., Mowbray, A., King, G., Cant, S., and Chang, P., (1997), *More Than Wyshful Thinking: AustLII's Legal Inferencing via the World Wide Web,* Proceedings of the Sixth International Conference on AI and Law, Melbourne, ACM Press, pp 47-55.

[5] Spirgel-Sinclair, S., and Trevena, G., (1988*), The DHSS Retirement Pension Forecast and Advice System*, in Duffin, P., (ed), Knowledge Based Systems: Applications in Government Administration, Ellis Horwood, pp34-40.

[6] Sprowl, J.A., (1980) *Automating the Legal Process: A Computer that uses Regulations and Statutes to Draft Legal Documents*, in Niblett, B., (ed) Computer Science and Law, Cambridge University Press, pp195-205.

[7] Daskalopulu, A., and Sergot, M., (1995), A *Constraint-Driven System for Contract Assembly,* in Proceedings of the Fifth International Conference on AI and Law, Maryland, ACM Press, pp 62-70.

[8] Branting, L.K., Lester, J.C., and Callaway, C.B., *Automated Drafting of Self-Explaining Documents,* in Proceedings of the Sixth International Conference on AI and Law, Melbourne, ACM Press, pp 72- 81.

[9] Ashley, K.A., (1990), *Modelling Legal Argument*, Bradford Books, MIT Press.

[10] Bench-Capon, T., and Staniford, G., PLAID - Proactive Legal Assistance, Proceedings of the Fifth International Conference on AI and Law, Maryland, ACM Press, pp81-88