Check for updates

# SmartLaw: Adapting "Classic" Expert System Techniques for the Legal Research Domain

Carole D. Hafner
College of Computer Science
Northeastern University, Boston MA 02115 USA

Virginia J. Wise
Harvard Law School
Cambridge MA 02138 USA

## Abstract

This report describes research in progress on the development of a computer expert system (SmartLaw) for giving advice on legal research problems. Legal research exhibits many of the characteristics of a suitable domain for expert system development; however, it also poses unique challenges for knowledge-based system design. To meet these challenges, we use a four-level knowledge structure of research STRATEGIES, GOALS, RESOURCES and PLANS, with three processing components: a rule-based backward-chaining reasoning component, a database component, and a hypertext component. This paper explains our evolving model of legal research knowledge and describes the architecture and implementation of a working prototype of the SmartLaw system.

## 1. Introduction

As societies and their legal systems have become more complex, the publication of legal information has experienced explosive growth. Publishers and consumers of legal information have turned to technology to assist them in managing this growth by offering full-text databases and other computer-based systems. While such developments offer many advantages to legal researchers, they are a mixed blessing. Researchers now not only have to learn the content and arrangement of the material, they also have to deal with a variety of hardware, software, and communications options. In 1993, a researcher must master both print and electronic sources because not everything is available online. [Chester 1991, Pritchard, 1988] There has also been a revolution of rising expectations about both the quality, quantity and timeliness of information a researcher should be able to deliver, given the existence of online resources. [Pressman, 1989] Both law students and legal professionals are increasingly finding their research skills falling behind these rising expectations. Many law schools, including the top American law schools, are now offering courses in Advanced Legal Research. [Kauffman 1986] Law firms are implementing their own legal research training programs and mandating that summer clerks attend vendor training. But such courses, if done well, are extremely labor intensive and expensive. Furthermore, existing courses and textbooks in legal research (including computer-assisted instruction programs such as Don Trautman's interactive videodisk tutorial [Harvard 1990] and Paper Chase [Hardy 1990]) are organized bibliographically. Students learn about the various kinds of publications, what functions they perform, and how to use them., but this information is difficult to retain, since it is not learned in a problem-solving context.

In this paper, we describe a research effort to apply "classic" expert systems technology to address some of these difficulties. We are developing an expert system ("SmartLaw") which is intended to perform as a reference librarian or other consultant that a legal researcher may turn to for assistance. By organizing SmartLaw's knowledge in terms of goals and strategies rather than legal bibliography, we hope to achieve a new and useful technique for helping researchers which would supplement existing instructional materials and offer several advantages:

a. The advice can be customized to fit a particular research task, by skill level of the researcher, by jurisdiction and by topic.
b. The system can be available 24 hours a day, seven days a week, when no human assistance (such as a law

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1993 ACM 0-89791-606-9/93/0006/0133 $1.50

librarian) is available. This is significant since electronic media make it possible to do legal research outside of a library.

c. It can provide help for those who are hesitant to ask a human expert for assistance for fear of revealing their ignorance.
d. It can provide point-of-use instruction so that lawyers who may have learned about a specific research tool, such as the Current Index to Legal Periodicals, can be refreshed about the details of how to use that tool, warned about particular hazards and common errors, or merely reminded that it exists.
e. It can be updated to reflect changes in legal research tools more easily than a textbook. (This also applies to computer-assisted instruction programs.)
f. It can integrate the use of online, print, and CD-ROM resources by suggesting specific strategies which may be appropriate for each.

In the next section, we present a brief discussion of why we believe legal research is a good domain for a "classic" expert system. Section 3 gives a detailed description of SmartLaw's design and implementation, and Section 4 presents a sample dialogue. Section 5 discusses plans for future work.

## 2. Why legal research is a good domain for a "classic" expert system

We can consider the characteristics of a problem domain which make it a good candidate for a "classic" expert system. A "classic" expert system such as MYCIN [Buchanan 1984] acts as a consultant to a person who has a problem to solve. The system asks the user some questions in order to diagnose the problem and then offers advice. A classic expert system is intended to mimic the advice and the reasoning of a human expert. Several domain characteristics have been universally recognized as important for a successful expert system development project: [Prerau 1985].

a. The domain must be narrow, but complex enough to be worth creating a computer system for giving advice.
b. The amount of knowledge required for effective performance must not be too great (several thousand rules but not tens of thousands.)
c. There must be recognized human experts to provide the system's knowledge.
d. The role of general world knowledge or common sense in solving this class of problems must be small.
e. There must be a category of users who need the system and will be willing to use it.

Legal research appears to match these criteria well:

a. Although the domain of law is very broad, the specific knowledge of legal publications, what they contain, and how to use them is quite specific and technical.
b. Although the amount of knowledge is great for a complete legal research advisor, it can be broken down by jurisdiction, and to some extent by legal topic, and thus made manageable. (For example, there are legal topics such as tax law, labor law or transportation law for which most research is conducted using specialized resources. Knowledge of these resources can be added to Smartlaw in a modular fashion.)
c. Teachers of legal research and reference librarians are recognized experts in the domain, who currently offer advice to law students and others.
d. Although general world knowledge is essential for actually doing legal research, there is not much common sense involved in knowing how to use a digest or a citator, or knowing how to look up the text of a Congressional committee hearing. (The system must have some primitive knowledge of temporal relations, i.e. calendar dates).
e. Law students definitely need this kind of advice (as reference librarians in law firms and law schools can testify), and we hypothesize that they will not be resistant to using a computer system if the advice offered by the system is useful.

## 3. System Design of Smartlaw

This section explains our evolving model of legal research knowledge and describes the architecture of a prototype system implemented using the Knowledge Pro development environment [Knowledge Garden 1991]. Several design requirements were identified:

i. The system should offer advice on research strategy, in addition to offering advice on which legal resources to consult and how to use them.

ii. The sysem must be able to "scale up" to a large database of legal resources, and be able to customize its advice to the current research topic and jurisdiction.

iii. Advice given to the user should be "how to" information addressed to the current research need, rather than general expository information about the resources. However, more descriptive tutorial information should be easily accessed via hypertext, allowing users at varying levels of expertise to benefit.

Requirement ii. rules out a purely rule-based expert system, since it is not feasible to represent unique rules for hundreds (or thousands) of resources (including print and electronic publications) -- even if it were

possible to write such a large rule base, it would be too difficult to update. We require a more abstract approach, where the specific resources could be entered into a database, and the expert system would embody general reasoning heuristics for selecting resources from the database. We also require that the advice of Smartlaw should depend not only on the resources selected, but also on the current research goals, user characteristics and other elements of the research context.

Our answer to these challenging design requirements appears in Figure 1, which shows the architecture of the Smartlaw system. The system consists of three major components: a rule-based reasoning component, a database component, and a hypertext component. First the rule-based component selects a strategy, defined in terms of the user's research GOALS. Next, a database of RESOURCES is searched to find resources that satisfy the user's goals. Once those resources are found, PLANS for using the resources are displayed as hypertext documents.

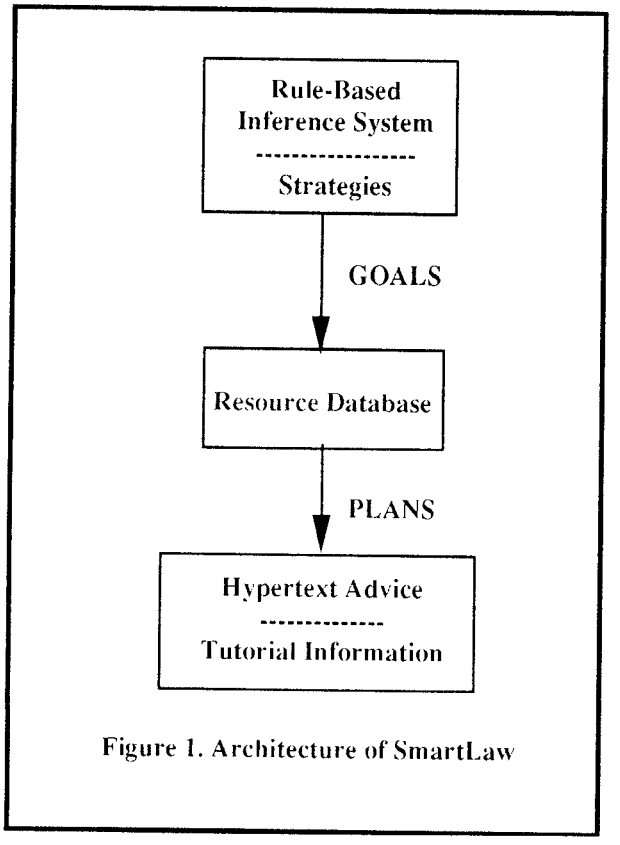


Figure 1. Architecture of SmartLaw

ID_____ CITE _____ABBREV_____

NAME _____

AUTHOR_____ KIND _____

TOPIC_____ JURIS_____

START_YR _____ STALENESS_____

SUIT_____ AUTH_____

LEVEL_____ REF_____

MEDIUM_____ GOAL_____

PLAN_____ OTHER_____

Figure 2. SmartLaw Database Schema

### 3.1 SmartLaw's Database of Legal Resources

Each record in the database represents a RESOURCE for achieving a GOAL for example, a court reporter satisfies the goal of looking up the full text of a case. Records can represent a particular resource, such as the United States Code Service, or a generic resource, such as a "state legal encyclopedia". The fields of the database (as currently implemented) are shown in Figure 2. The name, citation,author and type of the resource are followed by the legal topic it covers (or "any topic"), and the jurisdiction, which can be US, State, Non-US, a particular state, or generic. The use of the "STATE" jurisdiction allows the system to recommend resources (or types of resources, such as "a state encyclopedia") which are useful for state law research, even if the state does not yet have a module in the knowledge base.

The next fields indicate the time span covered by a resource. The Start Date field is particularly important for on-line resources, many of which only began in the 1980's and do not have the same information as their print counterparts. Users must be warned about these limitations. The "staleness" information indicates how quickly new information appears in the resource. For example, if pocket parts are published every six months for an annotated code, on any given day it would be at most six months behind. Since staleness is represented in days, the value for six months would be 180.

The next group of fields helps SmartLaw determine

which resources, among those which can satisfy a goal, are the best ones to recommend. The fields represent:
a. the suitability of a resource for achieving the goal
b. the authoritativeness of the resource
c. the level of expertise needed to use it
d. its usefulness to locate other relevant material.

The Medium field indicates whether the resource is published print, Lexis, Westlaw, CD-ROM, etc. Resources available in more than one medium have a YES value in the Other field, and are linked by ID to information about the resource in other media.

The most important database fields, for the purpose of this discussion, are the GOAL and PLAN fields. The goal represents one of a defined set of legal research goals that SmartLaw recognizes. It is the task of the rule-based inference component to identify one or more research goals which the user should pursue. These goals are used to select records from the database, and these records in turn, are incorporated in the advice offered to the user. A resource record is selected if it matches the GOAL, the jurisdiction, and the topic of the user's research, as determined by the rule-based component. Our detailed model of legal research goals is described in the next section.

While the GOAL field links the resource database to the rule-based inference system, the PLAN field links the resource database to hypertext documents containing advice for the user. The value in the PLAN field identifies a hypertext block to be displayed (which in turn may be linked to other hypertext blocks.)

One additional point should be noted about SmartLaw's database: it is a hidden component of the system, from the user's point of view. The user interacts with the rule-based component through menu selections and question/answer sequences. The user also interacts with the hypertext component through viewing hypertext screens and following hypertext links. However, the user does not interact with the database component of SmartLaw.

### 3.2 SmartLaw's Rule-Based Inference System

The "top level" of SmartLaw is the rule-based inference system, which invokes the database and hypertext components. The rule-based component asks questions of the user in order to find out the stage of the user's research, the topic and jurisdiction, the kind of legal authorities which dominate the research topic, the user's access to on-line resources, etc. Depending on the answers to these questions the system selects one or more *strategies* to recommend to the user. A strategy is represented as a goal tree, where the leaves of the tree represent discrete GOAL elements which match resources in the database.

SmartLaw applies a simple goal-reduction model to formulate what needs to be accomplished in solving a legal research problem. Figure 3 shows part of of a SmartLaw goal tree. The top-level goal "Research the Problem", is associated with three attributes:
--the legal topic,
--the jurisdiction,
--the dominant source of legal authority for the current problem (e.g., case law, statutes, or administrative law)

The top level goal has four sub-goals:
a. attain general background on the legal topic, and sufficient knowledge of its vocabulary to formulate search requests
b. identify the relevant legal theories for resolving the current problem
c. find controlling legal authority
d. evaluate the status of the authorities found

The first two levels are common to almost all legal research. At the next level, goal trees will differ depending on the user's problem situation. One strategy for identifying the relevant legal theories and primary authority is the "periodical strategy". This strategy is elaborated by using a periodical index, followed by looking up the periodical articles identified using the index, followed by looking up primary authorities mentioned in the periodical articles. At the leaves of the tree are specific GOALS which can be used to retrieve resources from the database.

In SmartLaw, a GOAL has three components: *function*, *object* and *given*. The *function* component of a SmartLaw goal is either BKG, FIND, UPDATE or FULL. The BKG function represents the acquisition of background knowledge about a legal topic. FIND and UPDATE represent the familiar cycle of locating primary authority and checking its current status. The FULL function represents a part of the research process which we take for granted in our theoretical analysis, but which can pose problems for the less experienced researcher -- even if the researcher has the citation of a relevant document, it is still necessary to know how to locate the full text in order to read it. So, in reading a law review article, the user may come up with a citation to a ruling of the Environmental Protection Agency, and may decide it is important to read the text of the ruling. How does one look up such a ruling? Are they available

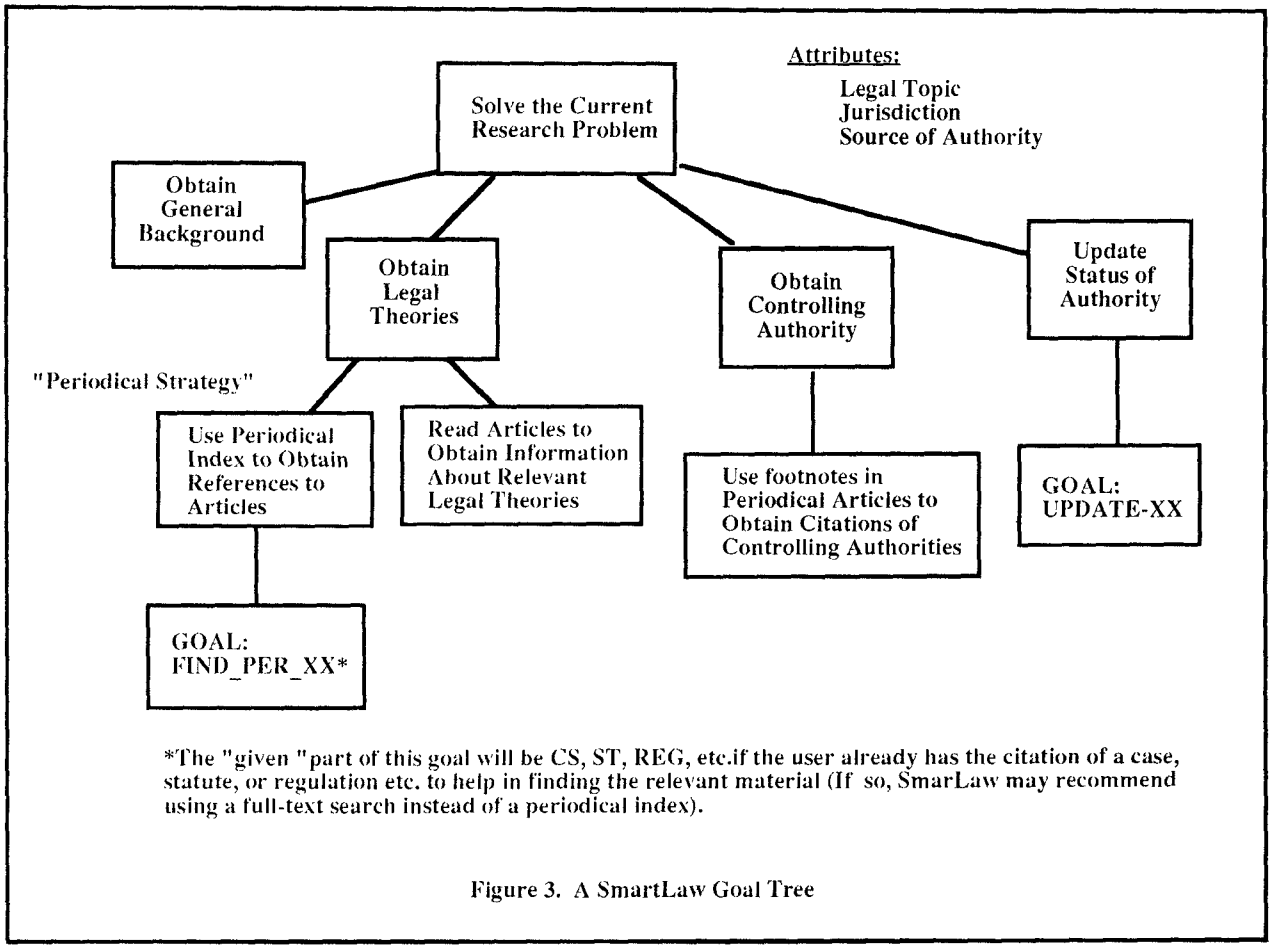


Figure 3. A SmartLaw Goal Tree

on LEXIS? For most law students, this would pose some difficulty.

The *object* component of a SmartLaw goal represents the kind of authority that the user is seeking. For example, immigration law is primarily statutory, while contract law is primarily case based. Thus, the goal UPDATE_CS indicates that the current task is updating the status of cases which have been located. If the jurisdiction is Massachusetts, this goal would retrieve database records for Shepard's Massachusetts Citations. If several kinds of law are important in the user's research, goals will be created for each of them. The current choices are: case, statute, administrative, constitutional, legislative history. Other kinds can be readily added if necessary. If the function component of a goal is BKG, then the object component indicates what level of background the user requires.

The *given* component of a SmartLaw goal describes the information that can be used as an entry into the literature: for example, a case citation, a statute citation, or subject terms to be used with an index. This allows SmartLaw to take advantage of the fact that users, even those having little or no background, often know the citation of some relevant authority which they have been told about by a teacher or senior colleague (e.g., "be sure to look at Smith v. Jones" or "you can start with Section 10b5 of the Securities Act").

### 3.3 SmartLaw's Hypertext Advising System

Once a set of resources has been located, SmartLaw displays their names in a menu and allows the user to select resources to learn more about. When a resource is selected, the associated advice text is displayed. At this time, the user can read the advice for using the selected resource to achieve the current goal. The advice text may include one or more hypertext links signified by "hot words" which are highlighted. When a hot word is activated, a new window opens up with a

further explanation of the hot word. The use of hypertext allows SmartLaw to display some tutorial information for those users who want it, while users who do not want to be tutored at a particular time can still make use of the system as a problem solving aid. .

Another knowledge-based legal research system that combines hypertext with rules is the Datalex Workstation [Greenleaf 1991]. However, unlike SmartLaw, Datalex represents substantive knowledge of a narrow legal domain (the Autralian Privacy Act). Its hypertext component is used to link sections of the Act to each other, and to the corresponding rules of a substantive legal expert system . Hypertext has also been applied directly to electronic treatises without the use of expert system techniques.

### 3.4. Further Exploration

As mentioned above, SmartLaw's inference component selects one or more goals representing what the user should do next in his or her research. For each goal, resources are retrieved from the database, and displayed in a menu. The user can select one resource at a time, and view the associated hypertext advice block which explains how to use the resource to achieve the current goal. On exiting from the hypertext module, the same menu is repeated until the user selects "Proceed" -- thus, information about one or more of the recommended resources can be displayed and repeated as many times as the user desires.

### 3.5. Implementation of SmartLaw

An initial prototype of SmartLaw has been created, using the KnowledgePro DOS development system (a product of Knowledge Garden, Inc.). KnowledgePro is a high-level language that supports both rule-based programming (with backward chaining) and conventional programming in an integrated system. KnowledgePro also has built-in hypertext capabilities which are completely programmable, and an interface to DBase compatible database files. A new implementation is under development using KnowledgePro for Windows.

Version 0.1 of SmartLaw, the first prototype, includes about 120 legal resources, and 20 inference rules. A sample interaction with this prototype is shown in the next Section.

## 4. Sample Dialogue

Figures 4 through 8 show a sample dialogue with the first prototype of SmartLaw. The "strategy" component described in Section 3.2 is not yet integrated into the advice system, thus the dialogue does not show multi-step recommendations.

Figure 4 is a "Welcome Screen" that gives the user a choice of advice on "finding the law" vs. looking up the full text of a known document. The processes involved in these two tasks (and the dialogue following the top-level screen) is sufficiently different to be treated as two "components" of SmartLaw. (Researchers working on general "library expert systems" have found the same division useful, e.g. [Parrot 1986]).

Figure 5 shows SmartLaw acquiring general information about the research problem: the legal topic, the jurisdiction, and the media to which the user has access. Figure 6 shows the acquisiting of information about the current state of the user's knowledge. The last question in this figure will determine the *given* component of the inferred research goal. Figure 7 shows SmartLaw's recommendations. The first line "In order to . . ." provides a paraphrase of the current goal. When the user selects a resource which exists in several media, such as *Index to Legal Periodicals*, another menu appears informing the user of that fact, and asking which media the user wants to see advice for. Figure 8 shows the initial hypertext display for *Index to Legal Periodicals* at the top, with the lower frame showing the expansion of the hypertext term **Example**.

## 5. Research Plans and Questions

Currently, our research is aimed at using SmartLaw experimentally with students in the Advanced Legal Research class. The knowledge encoded in SmartLaw will be coordinated with student assignments, and we can observe the student's use and acceptance of SmartLaw. Questions about the potential of SmartLaw as a useful educational tool include:

a. Naturally a computer can never compare with a reference librarian, who applies common sense, legal expertise, and the ability to interpret non-verbal cues (such as a puzzled look when the student does not understand the advice that has been offered). However, we have assumed that the computer can perform some of the functions of the reference libarian in transmitting important information to users. It remains to be seen whether, without common sense and human communication skills, the advice that is offered will be "on point".

b. Will students have the patience to work with a

computer, answering questions from menus, etc? Also, will that fact that SmartLaw, while offering an attractive, mouse activated interface, does not include "sexy" animation, keep students away?

c. Are the questions asked by SmartLaw meaningful to users, and are the categories of user needs (represented by GOALs) an effective way to structure advice?

d. Is it practical to keep a system like SmartLaw sufficiently up to date, given the rapid rate of change of legal material? (See Section 1)

Our plans for future expansion of SmartLaw will take advantage of its modular structure to add both new jurisdictions and new subjects. Other important enhancements are the ability of users to request printed summaries of the advice they have received, and the ability to link easily to on-line research services to retrieve full text of the resources identified during a SmartLaw consultation. We also are interested in exploring possibilities for customizing SmartLaw for use in different settings -- for example, allowing librarians to add local call numbers to document descriptions.

## 6. Acknowledgements

This research is supported by a grant from the National Center for Automated Information Retrieval (NCAIR). We would like to acknowledge the useful discussions of Terry Martin, Professor of Law and Law Librarian, Harvard University, Frederick Schauer, Frank Stanton Professor of the First Amendment, Kennedy School of Government, Harvard University, Donald H. Berman, Richardson Professor of Law, Northeastern University, and the students in the Computational Methods in Legal Research course offered at the Harvard Law School.

## 7. References

Buchanan, B.G.; Shortliffe, E.H., 1984. *Rule based-expert systems: the MYCIN experience of the Stanford Heuristic Programming Project.* Reading, MA: Addison-Wesley.

Chester, S., 1991. Electronic malpractice: Does reasonable competence require computer research? *Law Practice Management* 17(8): 22-27.

Greenleaf, G, A. Mowbray and A. Tyree, "The Datalex Legal Workstation - Integrating tools for lawyers." In *Proc. 3rd Intl. Conf. on Artificial Intelligence and Law,* Oxford, UK, June 1991 (ACM Press, New York), 215-224.

Hardy, I. T., 1990. Project CLEAR's paper choice: hypertext system for giving advice about legal research. *Law Library Journal* 82(2): 209-237.

Harvard Law School, *Introduction to Legal Research: A Survey of Basic Research Techniques*, 1990.

Kauffman, S.B., 1986. Advanced legal research courses: A new trend in American legal education. *Legal Reference Services Quarterly* 6 (3-4): 123-139.

Knowledge Garden, Inc., 1991. Knowledge Pro Windows User Manual. Knowledge Garden, Inc., Setauket NY.

Parrott, J.R., 1986. Expert systems for reference work, *Microcomputers for Information Management* 3(3): 155-171.

Prerau, D.S, 1985. Selection of an application domain for an expert system. *AI Magazine* (Summer): 26-30.

Pressman, S., 1989. The 24-hour-a-day lawyer. *California Lawyer* 9(10): 61.

Pritchard, T., 1988. Attorneys in the electronic information age; Is there a duty to make the transition? *Florida Bar Journal* 62(3): 17-20.

```
    Welcome to SmartLaw - the expert legal research consultant!
                (Copyright 1992, C. Hafner and V. Wise)

You should use this system if you have a research problem and would like
advice on finding or accessing the relevant law. The expert system will ask
you some questions about your problem, and then offer its recommendations.
If your problem is outside the system''s expertise, it will advise you to
seek advice from a human expert. SmartLaw can offer two kinds of advice:

  -- FIND - finding the law relevant to your research problem
  -- FULL - locating full text of a particular document whose
            citation, name or author you already know.

                What would you like to do next?

                    |-------------|
                    |   FIND      |
                --->|   FULL      |
                    |   Tutorial  |
                    |   Quit      |
                    |-------------|
```

Figure 4. Sample Dialogue with SmartLaw

```
If your research is in one of the specialized topics listed
below, please select that topic.  Otherwise, select "Any topic"
to get advice that is applicable to all legal research.

                          |----------------------|
                      --->| Any topic            |
                          | Government Benefits  |
                          | Immigration Law      |
                          | Tax Law              |
                          |----------------------|

What kinds of resources are available to you for this research?
Select one or more answers in the list below . . .
(or "All" for Print, LEXIS, and WESTLAW)

                          |------------|
                      --->| All        |
                          | Print      |
                          | LEXIS      |
                          | WESTLAW    |
                          |------------|

What jurisdiction does your research fall under?

                             |-------------|
                             | US          |
                         --->| State . . . |
                             | Non-US      |
                             | don't know  |
                             |-------------|

Please choose one of the states listed below, or select
"Any state" to get general advice on state law research.

                             |------------|
                             | Any state  |
                         --->| MA         |
                             | NY         |
                             |------------|
```

**Figure 5. Sample Dialogue (cont.)**

```
Most legal research proceeds through several stages:
 a. Acquiring general background knowledge about a legal topic about which
    you have very little knowledge (Ex: Immigration Law).
 b. Acquiring more focused background about particular legal issues
    that are relevant to your problem.
 c. Locating and reading the primary authorities relevant to your problem.
 d. Updating the current status of one of more primary authorities
    which you already have.
Which of these objectives would you like to concentrate on right now?

                                  |----------------------|
                                  | General Background   |
                              ---> | Focused Background   |
                                  | Find Authority       |
                                  | Update Authority     |
                                  | don't know           |
                                  |----------------------|

What kind(s) of law are most important to your research
problem.  Select one or more answers . . .

                                       |----------------------|
                                   --->| case                 |
                                       | statute              |
                                       | regulation           |
                                       | constitution         |
                                       | legislative history  |
                                       | don't know           |
                                       |----------------------|

  Are the areas of law relevant to your research undergoing rapid
  change, or is the law on your topic fairly stable?

                                  |---------------|
                                  | Rapid change  |
                                  | Fairly stable |
                              --->| don't know    |
                                  |---------------|
```

**Figure 6. Sample Dialogue (cont.)**

```
In order to get more background in an area you know something about:
  I recommend that you consult the resources in the menu below.
  Choose an item from the menu to learn where that item is found and how
  you can best make use of it. Choose "Proceed" when you are ready to
  continue with the consultation.

    |-------------------------------------------|
    | Legal Resource Index                      |
 -->| Index to Legal Periodicals                |
    | Current Index to Legal Periodicals        |
    | Index to Periodical Articles Related to Law |
    | Proceed                                   |
    |-------------------------------------------|

   The Index to Legal Periodicals
   which you have asked about is available in several forms.
   Please select one or more of the following:
   (or "All" to see all recommendations)

    |------------|
 -->| All        |
    | In Print   |
    | On WESTLAW |
    |------------|
```

Figure 7. Sample Dialogue (cont)

```
The Index to Legal Periodicals is an example of a
legal periodical index produced by the H.W. Wilson Co.
To find citations to articles in English language journals from
the U.S., Great Britain, Canada and Australia which are at
least 3-5 pages long, consult the subject index under your topic.
You will be given a reference to the title, author and citation
of a legal periodical.  To see an example of a typical Index
to Legal Periodicals entry, choose Example.  The subject
headings used in the Index to Legal Periodicals can be located
by using the Index to Legal Periodicals Thesaurus.

------------------------------------------------------------------------
Hypertext for Example

Tax credits
  Developing computer technology with the research credit.
   R.W. McGee.  31 Prac. Law 13-24 Je '85

Tax credits (Subject heading assigned by H.W. Wilson).
  Developing computer technology with the research credit. (Title
  of article).  R.W. McGee (author).  31 (volume number) Prac.
  Law.  (abbreviation assigned by H.W. Wilson to periodical
  title, here Practical Lawyer) 13-24 (inclusive pagination, e.g.
  article begins on page 13 and goes to page 24)  Je '85 (date of
  issue, here June, 1985).
------------------------------------------------------------------------
```

Figure 8. Sample Dialogue (cont)