3295

Computer-Assisted Diagnosis Using Expert System-Guided Hypermedia

A.J. Thomson, J.R. Sutherland, and C. Carpenter
Pacific Forestry Centre
506 West Burnside Road
Victoria, British Columbia
Canada V8Z 1M5

ABSTRACT. A hypermedia system, based on the text and graphics of an existing manual, was developed to provide information on forest seedling nursery disease and insect problems and their management. An expert system was included to guide the choice of links to follow in order to access the most suitable information for diagnosing unknown problems. The expert system used a filtering process in relation to a threshold value to drop or retain possible choices of links to follow. Successive application of the filtering process is analogous to a complex database query.

In prototyping an expert system for diagnosis of forest seedling nursery diseases and insects, Thomson et al. (1991, 1992) indicated that although the rule-based shell used allowed rapid prototyping, the expert system approach by itself resulted in a system with certain limitations. The major limitation was a restricted domain which included mainly insects and diseases and some environmental damage. The possibility of extra-domain conditions, such as nutritional deficiencies or the presence of mycorrhizal fungi, precluded development of a system with sufficiently accurate performance that it could be distributed for general use. Constraints identified in the prototyping exercise led us to examine alternatives to expert systems for computer-assisted diagnosis.

Diagnosis can be facilitated by answers to database-type questions such as, "What are the most significant foliage blights of one-year-old Douglas-fir in container nurseries of coastal British Columbia?" Although this type of knowledge was implicit in the original expert system, it was not accessible. Consequently, we wanted to incorporate database-type functions in the application. Also, the prototype expert system did not include management advice or guides to sample collection, handling, and examination, and we wanted to have a computer system that addressed these issues.

Thomson et al. (1992) proposed the use of a hypertext adjunct to the expert system to deal with the sampling and handling issues indicated above. Hypertext is based on the traditional method of storing information on 3 x 5-inch cards, and is defined as the creation and representation of links among discrete pieces of data, with the data being in text form. When the data include graphics or sound, the resulting structure is referred to as *hypermedia* (Parsaye et al. 1989, Rauscher and Host 1990). In hypermedia systems a screenful of information is the analogue of the card. Formation and activation of connections (links) between cards is described by Rauscher and Host (1990). When a link to text is activated, the text segment which is displayed is referred to as a *node* or *chunk* of text.

In the study described here, we show how hypermedia, rather than merely being an adjunct to the expert system, provides the basis of a system that overcomes many of the limitations of expert systems used alone. The expert system then becomes an adjunct to the hypermedia. The present system was developed using the hypermedia authoring system HyperWriter! (NTERGAID Inc., Fairfield, Connecticut), with some additional programs in C (Borland International, Inc., Scotts Valley, California). See Rauscher and Johnson (1991) for a brief introduction to the HyperWriter! program. The nursery system runs under MS-DOS and requires EGA graphics capability. Copies of the program are available from the senior author on request.

Knowledge Organization

The system we developed was based on a manual (Sutherland et al. 1989) in which information was given about a variety of nursery problems, with each problem being a self-contained unit that included sections of text, section headings, figures, and references, with a table summarizing the conditions under which the problem occurs. A simple diagnostic guide was provided at the beginning of the book, but the primary means of using the book for diagnostic purposes was by browsing the photographs of signs and symptoms, whereas experts use a variety of information sources in diagnosing problems (Thomson et al. 1991, 1992).

The manual on which the system was based contained information that was structured in a form readily converted from the original word processor files to hypertext. In the manual, knowledge about each insect or disease was organized separately, with a brief introduction followed by sections on hosts and damage, life history, management, selected references, and a table summarizing the host and nursery conditions with which the problem is associated.

Text from each of these sections formed separate nodes in the system, with access to the node being primarily through the name of the causal agent via a section-title screen (Fig. 1). Figure 1 indicates that an additional section (node), not present in the manual, was added to each insect or disease description: a "look-alikes" section, which listed other organisms or environmental effects that result in signs or symptoms similar to those of the currently evaluated causal agent. Other links indicated in Figure 1 are discussed below.

Knowledge Browsing

The two contrasting modes of knowledge access, *browsing* or *guided search*, are highlighted in the opening screen of the application (Fig. 2), where graphic buttons (active screen areas attached to links) permit access to the browsing mode (the **Select Problem** button) or guided search (the **Diagnosis** button). Other buttons permit access to on-line tutorials on use of the system (the **Help** button) or information on the primary source material and funding agencies for the project (the **About** button). An **Exit** button permits the application to be closed down. The buttons activate links to other screens.

When the **Select Problem** button is activated, to enter the browsing mode, the screen displayed has three buttons: **Diseases, Insects,** and **Environmental Effects**. This mimics the three major sections of the original manual. Activating the **Disease** button, for example, displays a screen listing all the diseases covered by the system (Fig. 3). Here the links are indicated by triangular markers, rather than by buttons. A few diseases, such as Phytophthora root rot, were only mentioned briefly in the manual, with no accompanying details, while Keithia blight is a new problem not included in the manual. These are included on the screen pending development of appropriate descriptions. This illustrates one major advantage of hypermedia over books: the ease of updating. Note the organization of the diseases into groups with descriptive text, such as "Blights" or "Root rots." This facilitates the browsing in a manner similar to the simple diagnostic guide in the manual.

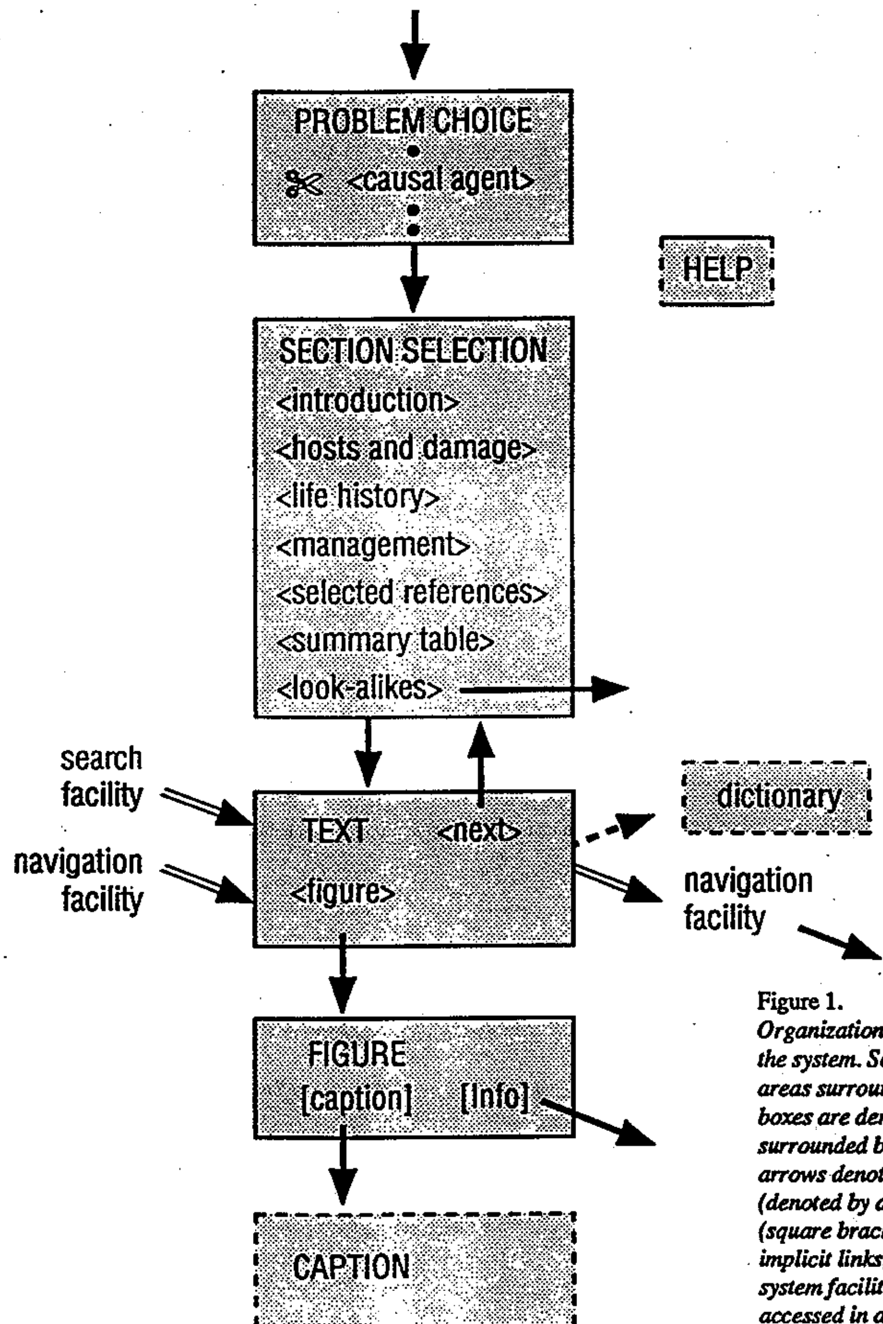

Figure 1.
Organization and linkage of knowledge in the system. Screens are denoted by shaded areas surrounded by solid lines. Pop-up boxes are denoted by shaded areas surrounded by dashed lines. Solid single arrows denote normal links from text (denoted by angle brackets) or buttons (square brackets). Dashed arrows are implicit links, while double arrows are system facilities. The HELP facility may be accessed in a number of ways and presented in a variety of formats, including on-line tutorials. The scissors symbol indicates the removal (filtering) of causal agents from the "Problem Choice" list by the expert system.

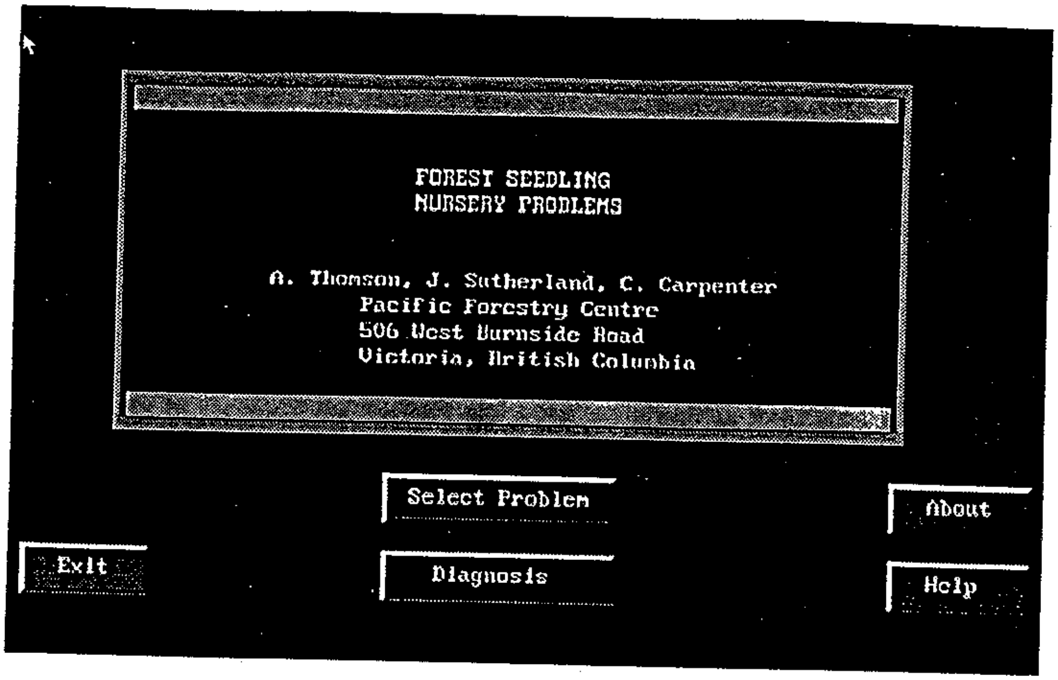

Figure 2.
The opening screen of the application. The rectangular areas titled ***Select Problem, Diagnosis, About, Help,*** *and* ***Exit*** *act as link navigation buttons accessed by use of the mouse device or, less easily, by computer keyboard commands.*

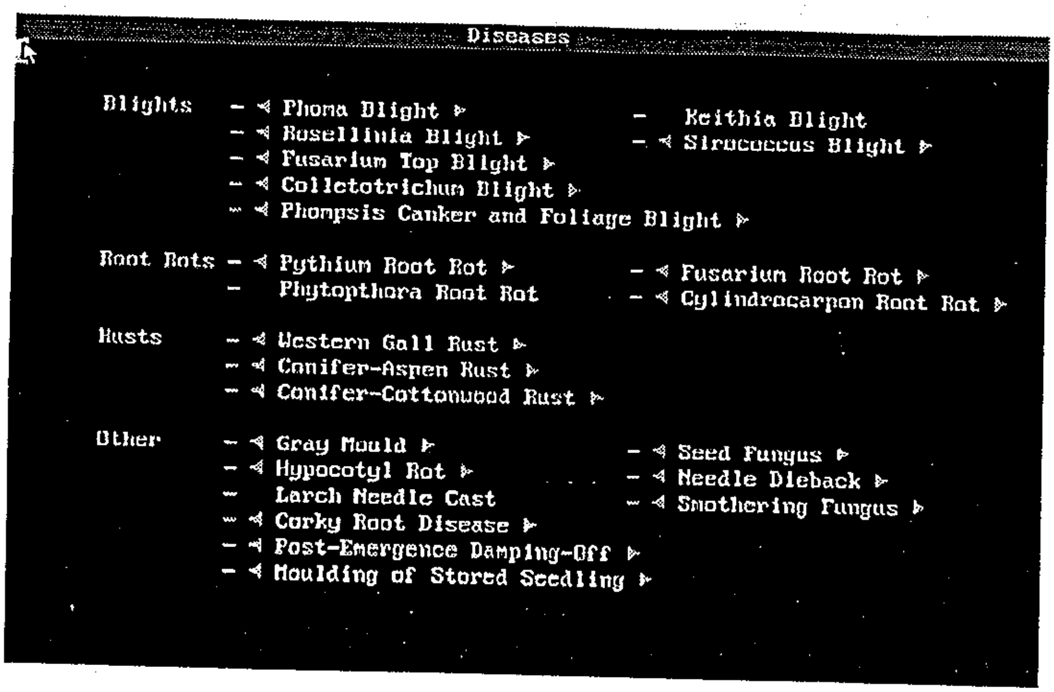

Figure 3.
A Problem Choice *screen for diseases in the browsing mode of operation. Links are denoted by the triangular symbols at the ends of text segments.*

Text Nodes

By following the sequence of problem choice (e.g., *Sirococcus* from Fig. 3) followed by section (e.g., hosts and damage), the text of that section of the manual is displayed (Fig. 4). Note that if the material to be displayed is too long to fit on the display, a scroll bar is provided. Screens can also be accessed through the search facility or navigation facility provided by the application development product (HyperWriter!) in the form of pull-down menu functions.

Each text screen is provided with a "Next topic" field, through which the next section in the problem description can be accessed. For example, the "Next topic" field of the **Hosts and Damage** screen links to the **Life History** screen.

All the links described to this point have been identifiable by markers such as buttons or triangular delimiters. Implicit links have no markers and are illustrated in Figure 4 for the word "pycnidia." When text has been identified as an implicit link, clicking on any occurrence of that text displays a section of text in a pop-up box inserted over the main text screen. This was used to provide a dictionary facility (Figs. 1, 4).

Links can also be made to graphics (Fig. 5). However, only one picture can be displayed at a time. A display of multiple figures, as in Figure 5, requires compositing outside of the HyperWriter! environment and saved as a single picture accessed by the link. Each of the figures has two associated buttons. The button marked with the figure number from the original manual presents the caption of that figure in a pop-up box similar to the dictionary function. The **Info** button links to the **Section** screen for that problem.

The hypermedia format permits a more powerful use of graphics than a book or manual. Parts of the picture can be defined as active areas. Clicking on an active area can then link easily to other information, although this feature was not used in the present version of the system. In the original manual, all figures for a particular problem were grouped together, while in the hypermedia product, figures can be grouped by similar signs or symptoms. In this way, the browse path can be from one problem through the composite figure to another problem via the **Info** button. Other ways of moving from one problem to another are through the **Look-alikes** screen or the system navigation facility.

Expert System Guidance

The expert system developed by Thomson et al. (1991, 1992) was constructed using an expert system shell, EXSYS (EXSYS Inc., P.O. Box 11247, Albuquerque, New Mexico 87192). Knowledge in this shell is expressed in the form of if-then-else rules. For example, the first rule in the system is

RULE NUMBER 1:

IF: The location of the nursery is coastal
and The nursery type is bareroot

THEN: Colletotrichum blight (*C. acutatum*)-
Probability=0/10
and Colletotrichum blight (*C. gloeosporioides*)-
Probability=0/10
and Needle dieback (*Pythium* spp.)-
Probability=0/10
and Gray mould (*Botrytis cinerea*)-
Probability=2/10
and Phoma blight (*Phoma* spp.)-
Probability=8/10
etc.

There is no ELSE section to this particular rule.

By printing the expert system to an ASCII file, a text version of the expert system was produced which could easily be incorporated into the hypermedia product (Figs. 6, 7). The "IF" sections of similar rules were collected together in a set to produce a navigation screen. For example, the seven seasonal periods and two seedling age classes resulted in 14 options with regard to this one rule set, with the germinant age class being a special case. A button format was used to provide an easily used and visually attractive user interface (Fig. 6), with one button being provided for every combination of conditions used in the rule set.

Each button links to a separate **Problem Choice** screen, as shown in Figure 7 for 1+0 seedlings in early summer. The figure shows

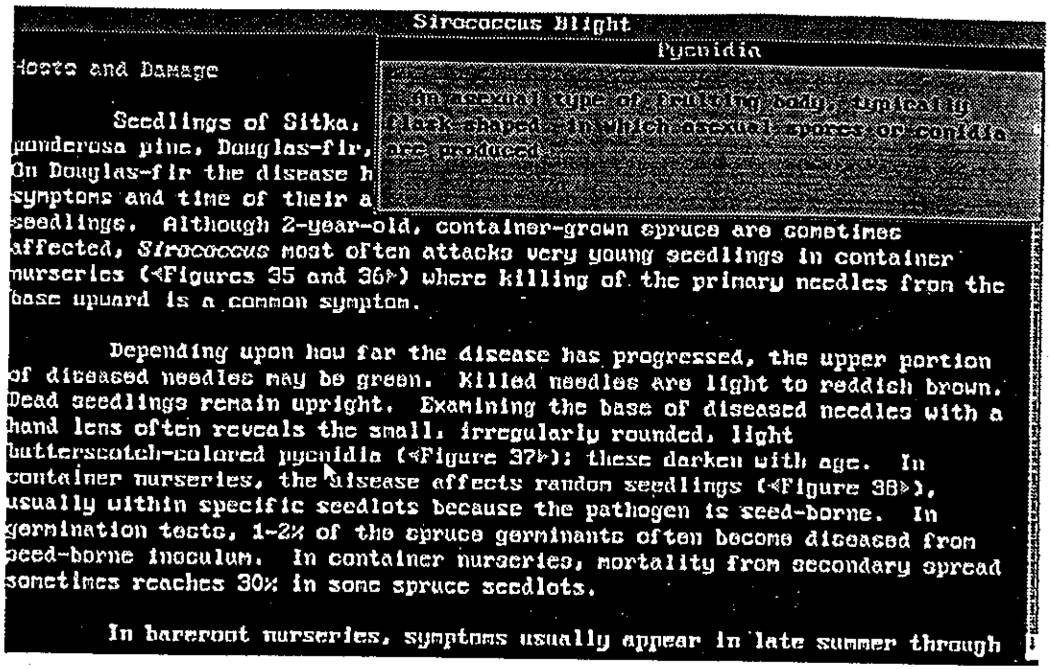

Figure 4.
A basic text screen illustrating scroll bars, a pop-up text box from an implicit link, and a link to a figure. The "Next topic" link is hidden by the pop-up text box.

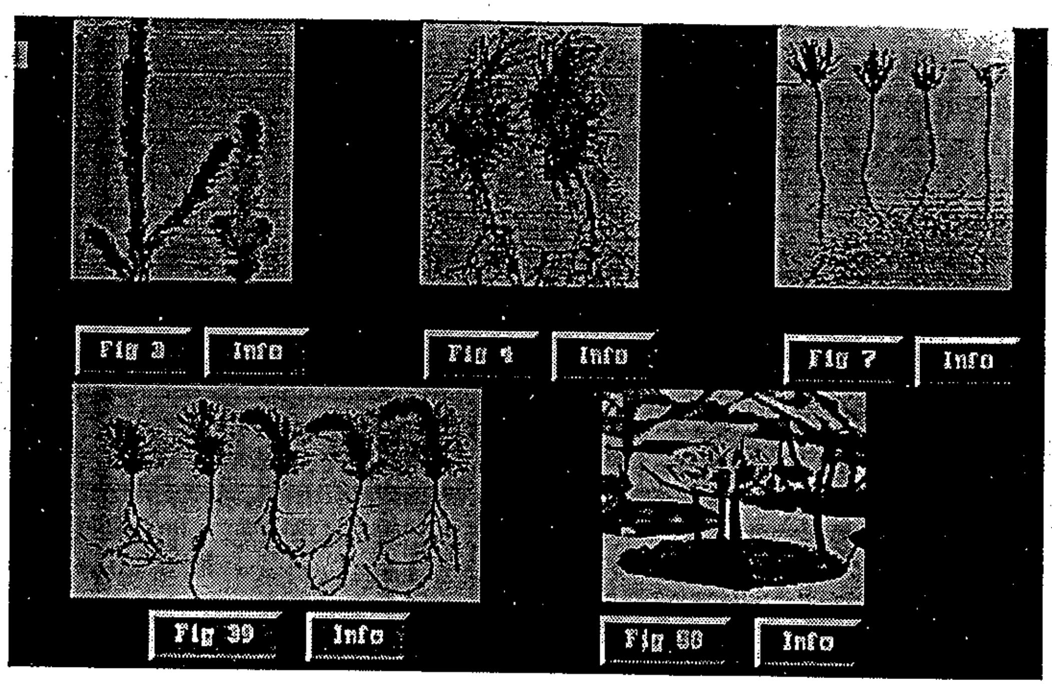

Figure 5.
A graphics screen with buttons linking to the figure captions and to the Section Selection screen.

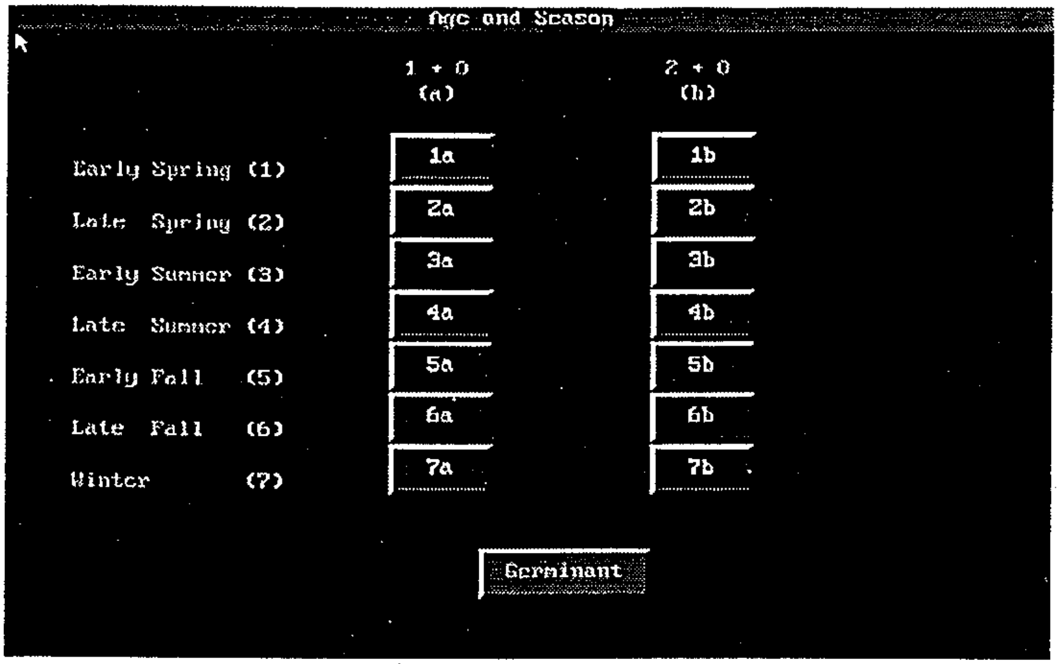

Figure 6.
A navigation screen derived from expert system rules. The combination of conditions of the rule (in this case age and season) is represented in button format.

Figure 7.
A Problem Choice screen derived from the Nursery Condition rule in the expert system. Buttons on the screen permit filtering of choices based on comparison of likelihoods with a pre-set threshold.

the screen scrolled to the end of the diseases, at the start of the insects. Each problem and its likelihood, obtained from the EXSYS listing, is converted into a link to the appropriate Section Selection screen (Fig. 1), described earlier for the browsing mode of operation.

Four buttons at the bottom of the screen permit use of the information to assist diagnosis. The first button, marked **Export**, copies the screen to a file for later viewing or printing to examine the likelihood values. The entire hypermedia screen (card equivalent), not just the part currently scrolled and visible on the monitor, is exported.

The three remaining buttons are to provide full expert system guidance. By activating the **Filter** button, the contents of the screen are passed to a program, written in C, that permits setting of a threshold likelihood value. All problems with a likelihood less than the threshold are filtered out of the list of possible problems. By applying this filter with successive screens describing nursery conditions and seedling signs and symptoms, the list of problems can be reduced to a few possibilities that can be examined by browsing.

Note that the weightings ("Probabilities") on the screens imported from the EXSYS program no longer function in any combinative manner. They are used merely in relation to the threshold values, and choices with values exceeding the threshold are not ranked in any fashion. Use of a zero-valued threshold eliminates all choices that are never associated with the characteristics defining that screen. Higher-valued thresholds relate to the concepts of "rare" or "common" in a fuzzy logic manner.

Ideally, the filtering process would remove problems from the **Problem Choice** screens, as suggested in Figure 1. However, the current version of the software development package does not permit dynamic construction of screens during run-time. To get around this, the **View** button is provided, which simply displays the contents of a file containing the current filtered list. Alternatively, the **Print** button sends the list to a printer. By referring to this list, the user can select an appropriate topic.

Rules involving signs and symptoms are handled in a slightly different manner than the nursery condition rules. Nursery conditions are easy to define and likelihoods relatively easily established, whereas evaluation of signs and symptoms is complicated by issues of differing viewpoints and severity ratings (Thomson et al. 1992). Because of difficulties in establishing likelihoods, these were dropped from screens relating to signs and symptoms (Fig. 8). Such screens showed all problems where the sign or symptom occurred, unless it was very rare. The filtering procedure in these cases dropped all problems not indicated on the screen.

When the **Diagnosis** button on the first screen (Fig. 2) is activated, a navigation screen with two buttons, **Nursery Conditions** and **Signs and Symptoms**, is presented. The **Signs and Symptoms** button links to other navigation screens, refining the conditions as defined in the various rules, analogous to the process illustrated in Fig. 6. The **Nursery Conditions** button links to a navigation screen with four choices, reflecting the main uses of this information by the expert in diagnosing problems (Thomson et al. 1992). Three of these choices, Nursery Type and Location, Host Species, and Age and Season, can be used with the filtering process to guide the diagnosis. The fourth choice, Diseases Enhanced by Nursery Conditions, is a database function and is described below.

Database Functions

Questions, such as "What diseases are enhanced by poor seedling nutrition," are database-type questions. Ideally, it should be possible to access a database and dynamically construct a screen displaying choices linked to the appropriate information, where knowledge about the links is also retained in a separate database. However, as discussed above, this is not possible with the current version of the software. Instead, choice selection screens were constructed based on the EXSYS rules relating to modification of likelihood by particular conditions.

Other database-type questions, such as "What are the most significant foliage blights of 1+0 Douglas-fir in container nurseries of coastal British Columbia?" described above, can be answered by use of appropriate filters in conjunction with particular threshold settings to denote rarity. In an ideal system, a product such as JAKE (English Knowledge Systems Inc., Scotts Valley, California), which interprets English text into database queries, would facilitate user interaction with the system.

Discussion

By combining hypermedia and expert systems, powerful software products for knowledge management and advice-giving can be developed to overcome many of the limits of expert systems used alone. Special diagnostic features can be highlighted in text or figures, or extra-domain look-alikes indicated, without being concerned with the processing of the features to make a final diagnosis. When there are only a few problems in a domain and their interactions and severity effects are well defined, it is possible to have the system comment on extra-domain possibilities or unusual behavior of a condition (Thomson and Taylor 1990). However, when there are many choices in the system and their interactions and severity categories are ill-defined, then extra-domain possibilities are more difficult to deal with.

In developing systems of the present type from existing books or manuals, some graphics may require re-evaluation. When very small features on figures are scanned into a computer format, they may be lost due to averaging of photographic image points into a single screen pixel. Similarly, features that require fine gradations of color change may be lost when the image is formed with a limited palette (e.g., 16-color EGA graphics in the present system). However, the ease with which the system may

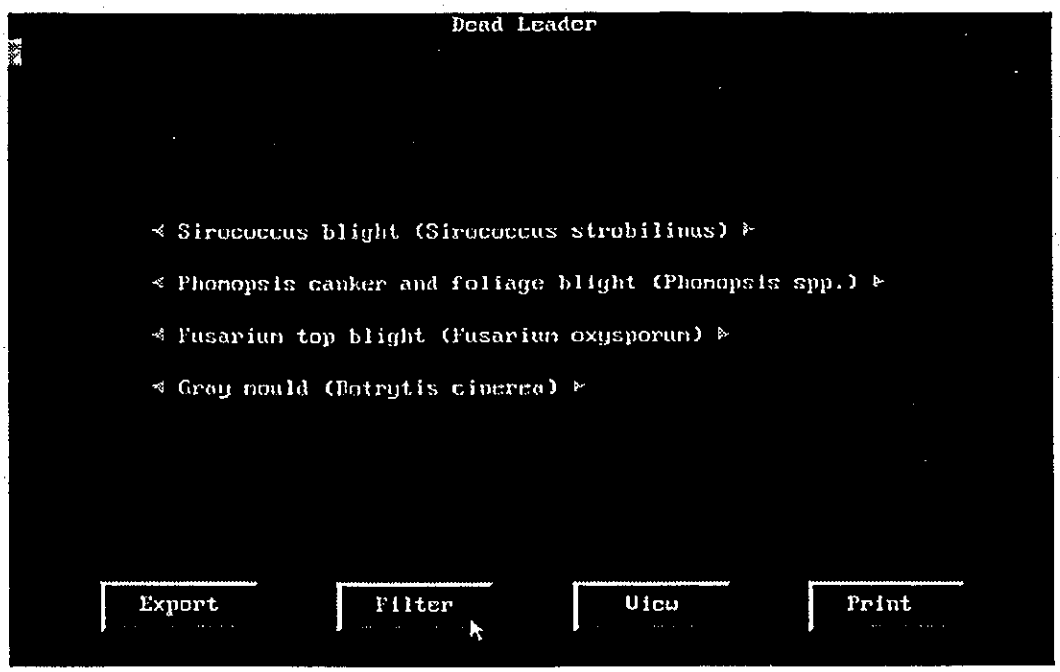

Figure 8.
A Problem Choice screen derived from a Signs and Symptoms rule in the expert system.

be modified makes inclusion of new material as it becomes available a simple task.

Although the present system is a considerable improvement over the original expert system, some desirable features were not feasible with the existing technology. In particular, the ability to create new screens dynamically during run-time would be desirable. If this were possible, the system could link to an external compositing program to display the most appropriate combinations of pictures to facilitate the diagnosis. The requirement for graphics in the user interface of diagnostic systems was emphasized by Thomson et al. (1992).

The filtering approach developed here, with its analogy to database access, offers nursery staff an alternative to the traditional forward-chaining type of expert system that starts with observations, such as "spots on leaves." Based on the expert knowledge contained in the system, nursery conditions such as nursery type (bareroot or container) and location, host species, age of stock, and season have a high diagnostic value and are easily defined by the system user. Through an initial filtering based on these factors, a short list of possibilities can generally be obtained which can then be further filtered on the basis of observed symptoms.

Kidd (1985) indicates that rigid dialogues and inadequate explanations are major shortcomings in the consultative function of many current expert systems. We believe that the filtering process of the expert system combined with the browsing capabilities of the hypermedia address many of the issues raised in Kidd's study.

Acknowledgments

This study was supported in part by the Canada/British Columbia Forest Resource Development Agreement study FC-ETT 17. We also wish to acknowledge financial support from the B.C. Ministry of Forests Extension Services.

References

Kidd, A.L. 1985. The consultative role of an expert system. Pages 248-254 in: People and Computers: Designing the Interface. P. Johnson and S. Cook, editors. Proceedings, Conference of the British Computer Society, Human Computer Specialist Group, University of East Anglia. Cambridge University Press, New York.

Parsaye, K., M. Chignell, S. Khoshafian, and H. Wong. 1989. Intelligent databases: object-oriented, deductive, hypermedia technologies. John Wiley and Sons, New York.

Rauscher, H.M., and G.E. Host. 1990. Hypertext and AI: a complementary combination for knowledge management. AI Applications in Natural Resource Management 4 (3): 58-61.

Rauscher, H.M., and S. Johnson. 1991. Authoring hypertext documents using Hyperwriter!™ AI Applications 5: 116-119.

Sutherland, J.R., G.M. Shrimpton, and R.N. Sturrock. 1989. Disease and insects in British Columbia forest seedling nurseries. Forestry Canada and B.C. Ministry of Forests, FRDA Report 065.

Thomson, A.J., J.R. Sutherland, M.E. Blache, and J. Dennis. 1991. Expert systems for diagnosing nursery insect, disease and environmental problems. Pages 223-225 in: Proceedings, IUFRO Symposium, Diseases and Insects in Forest Nurseries, J.R. Sutherland and S.Glover, editors. IUFRO Working Party S2.07-09. Victoria, B.C., August 22-30, 1990.

Thomson, A.J., and C.M.A. Taylor. 1990. An expert system for diagnosis and treatment of nutrient deficiencies of Sitka spruce in Great Britain. AI Applications in Natural Resource Management 4 (1): 44-52.

Thomson, A.J., J.R. Sutherland, M. Blache, and J. Dennis. 1992. Prototyping an expert system for diagnosis of forest seedling nursery problems. AI Applications 6(1): 21-31.

Alan J. Thomson graduated in 1968 with a B.Sc. (Hons.) in zoology (entomology) from Glasgow University. He received a Ph.D. in ecology and parasitology from McMaster University, Hamilton, Ontario, in 1972. Following four years of postdoctoral research (1972-1976) into simulation of nutritional and physiological systems at the Institute of Resource Ecology, University of British Columbia, University of British Columbia, Vancouver, B.C., he joined the Canadian Forestry Service as a biological systems analyst, involved in development of pest management and forest management systems. From 1988-1989 he was on leave at the Department of Forestry and

Natural Resources of Edinburgh University, developing expert systems for natural resource management. He is currently developing such systems in a forestry setting. He is a member of the Advanced Forest Technologies program at the Pacific Forestry Centre.

Jack R. Sutherland is principal research scientist at Forestry Canada's Pacific Forestry Centre (PFC), Victoria, British Columbia. He holds a B.Sc. in agriculture (1958) and an M.Sc. in botany (1960), both from Ohio University, and a Ph.D. in plant pathology (1964) from West Virginia University. He joined Forestry Canada in Quebec in 1964 and transferred to the PFC in 1967. Throughout his research career he has worked with diseases of forest nursery seedlings, seeds and cones, and seed orchard trees. He has published about 125 scientific papers and technology transfer publications on these topics. Presently he is doing research on gray mould and root rots of container-grown forest nursery seedlings. Early in 1991 he became program leader for forest pathology at PFC.

Carol Carpenter graduated from the University of Victoria in 1991 with a combined computer science/mathematics degree. She is currently employed by the B.C. Systems Corporation, working in the offices of the Ministry of Forests.