ABSTRACT. Hypermedia and expert system technologies were combined to develop a system for diagnosis and management of forest insect and disease problems (HFOREST). Once a diagnosis has been made, the management alternatives are discussed in the material available through hypertext links. Significant changes from the approach used in an earlier system for diagnosis of forest nursery problems include a change in the viewpoint used to assign likelihoods to the problems. Delivery of the system on CD-ROM permitted inclusion of approximately 350 high-quality color illustrations.

Forest Insect and Disease Diagnosis and Management Using Expert System-Guided Hypermedia

A.J. Thomson and G.A. Van Sickle
Pacific Forestry Centre
506 West Burnside Road
Victoria, B.C.
Canada V8Z 1M5

Expert systems by themselves have limitations in providing computer-assisted diagnosis (Thomson et al. 1991, 1992), the main problems being difficulties in dealing with extra-domain possibilities. An extra-domain possibility is a problem that might cause signs and symptoms similar to the problems covered in the system (insects and diseases in this case), but is either of a damage type not included (e.g., nutritional problems or environmental damage), or is caused by an insect or disease unknown to the system. To surmount some of these difficulties, Thomson et al. (1993) combined artificial intelligence (AI) and hypermedia (Parsaye et al. 1989, Rauscher and Host 1990) approaches in diagnosing insect and disease problems in forest seedling nurseries (HYPERNUR).

The underlying premise of the nursery system was that the AI guided system users to the most appropriate hypermedia links to obtain the information to assist them in diagnosing the problem, provided it was within the domain of the system. Extra-domain possibilities were indicated by a list of extra-domain "look-alikes." If none of the hyperlinked descriptions indicated by the AI conclusively identified the problem, then the system user could explore the extra-domain possibilities through other sources, and the system did not have to deal with these other possibilities in any computational way

The hypermedia part of the nursery system was developed using the hypermedia authoring system HyperWriter! (NTERGAID Inc., Fairfield, Connecticut) (Rauscher and Johnson 1991), while additional programs in C++ (Borland International, Inc., Scotts Valley, California) provided the expert system functionality. The nursery system ran under MS-DOS and required EGA graphics capability, the common standard in the computers used in nurseries at that time. The system was delivered in compressed form on a single 3.5" disk. Hypermedia and C++ have also been combined in a computer-based training system based on an object-oriented tree graphics program (Thomson and Van Sickle, 1995).

The nursery system was based on a manual (Sutherland et al. 1989) in which each nursery problem was a self-contained unit which included sections of text, section headings, figures, and references, with a table summarizing the conditions under which the problem occurs. Text from each of these sections formed separate hypertext nodes in the system, with access to the node being primarily through the name of the causal agent.

In the three years since development of the nursery system, multimedia systems have become widespread, with high-quality graphics and CD-ROMs common. In this study we describe the development of a forest insect and disease diagnostic system using the expert system-guided hypermedia approach pioneered in the nursery system, highlighting improvements from the original approach.

Knowledge Organization

The forest insect and disease diagnostic system (HFOREST) is based on the Forest Pest Leaflet series of the Forest Insect and Disease Survey (FIDS) unit of the Canadian Forest Service at the Pacific Forestry Centre. The material in this series differs from the nursery manual in several respects. Each leaflet has different authors and varies in style and content, requiring design of a range of navigation screens (Figure 1). The text was available in electronic form only for the most recent leaflets as they were produced or revised; earlier leaflets were scanned and the text extracted by Optical Character Recognition (OCR) software (Busch 1990).

While the text in the nursery manual is in small sections that approximate a screen in size, the leaflets often contain long sections requiring breakdown by insertion of subheadings to facilitate hypermedia paging. In some cases, the scientific names had changed since the original publication and a guide to the changing nomenclature had to be included.

The original negatives or slides for many of the illustrations in the leaflets are no longer available; thus illustrations were revised and updated (Figure 2), taking account of the limitations of digital representation (Thomson et al. 1993). The color slides from which the images were obtained were initially digitized on photo CDs and converted to PCX files for inclusion in the system. PCX files were provided in three formats to accommodate different hardware graphics configurations: VGA images of 16, 256, or 16 million colors.

Knowledge Browsing

As in the nursery system (Thomson et al. 1993), the two contrasting modes of knowledge access in the forest pest system are browsing or guided search (Diagnosis). For browsing, pest species are divided into insects and diseases, and can be accessed through either their scientific or common name.

Nomenclature was an issue that had to be addressed in the system. Leaflets that describe pests that have continued to be important are periodically revised, using up-to-date nomenclature. However, some problems were important in the early years of the series but thereafter declined in significance. Leaflets describing such problems were not updated although, in the interim, the scientific nomenclature of these pest species had sometimes changed. Changes in nomenclature from the original publication are indicated in the insect or disease selection screens, but the nomenclature used in the original publications was not changed.

The forest system contains information from 41 insect-related pest leaflets (Table 1) and 29 disease-related leaflets (Table 2). Twenty-five percent of the leaflets had been produced or revised within the last five years, while 33% were within the last 10 years. Some publications, such as the leaflet on dwarf mistletoes (Unger 1992), which covered five species of *Arceuthobium* (Viscaceae), dealt with multiple pest species.

The pest leaflet *Introduction to Forest Diseases* (Callan and Funk 1994) was a special case in that it did not deal with a specific problem, but rather gave an overview of forest diseases and the principles of forest pathology and disease control. This publication was well illustrated and already in elec-

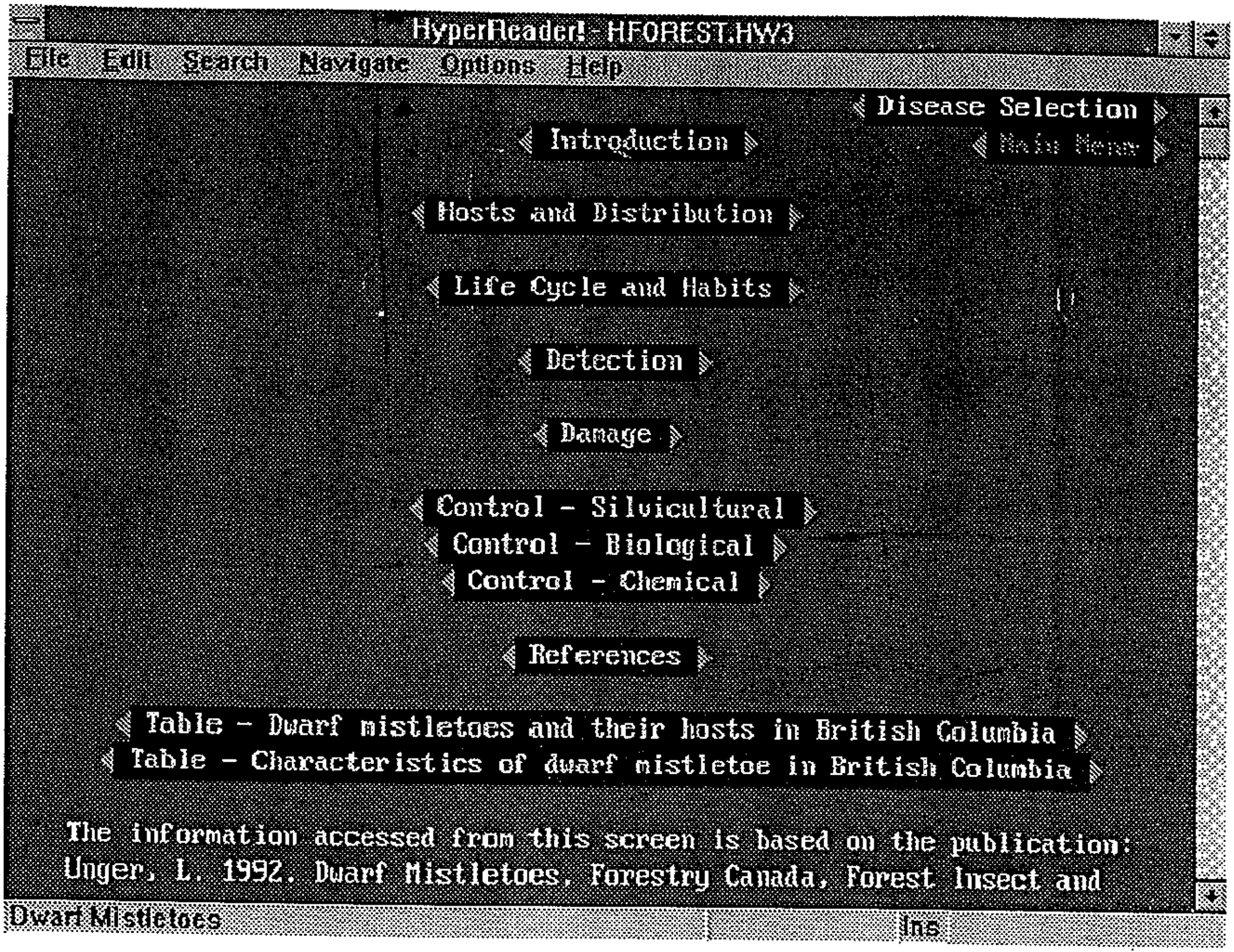

Figure 1. *Custom-designed navigation screen for the dwarf mistletoe Forest Pest Leaflet (Unger 1992), showing links to specific control options and tables.*

tronic form in a desktop publishing system. These illustrations were included in the system (Figure 3).

When navigating among the topics describing a problem (Figure 1), it is a common practice to go initially in sequence using the "Next" link available on each topic screen (Figure 2). The authoring tool keeps track of the browse history. This history trail can become a cumbersome impedance when moving back and forth among topics in different problems. Figure 3 illustrates two links back to the topic selection screen (Figure 1): the "Contents" link retains the browse history within that pest section, while the "Return" link drops the low-level browse history.

The section dealing with the mountain pine beetle (*Dendroctonus ponderosae* Hopkins) includes a unique section describing the population dynamics, impact, and management of this pest using a computer graphics presentation developed in C. The linkage of hypermedia and custom graphics software, especially object-oriented graphics, has been previously demonstrated to be very effective in computer-based training (Thomson and Van Sickle 1995).

Expert System Guidance (Diagnosis)

The expert system used in the nursery system had its origins in an EXSYS-based prototype (EXSYS Inc., P.O. Box 11247, Albuquerque, New Mexico 87192) (Thomson et al. 1991, 1992) in which knowledge was expressed in the form of if-then-else rules. For a given condition, such as host species or nursery location, each problem was assigned a likelihood

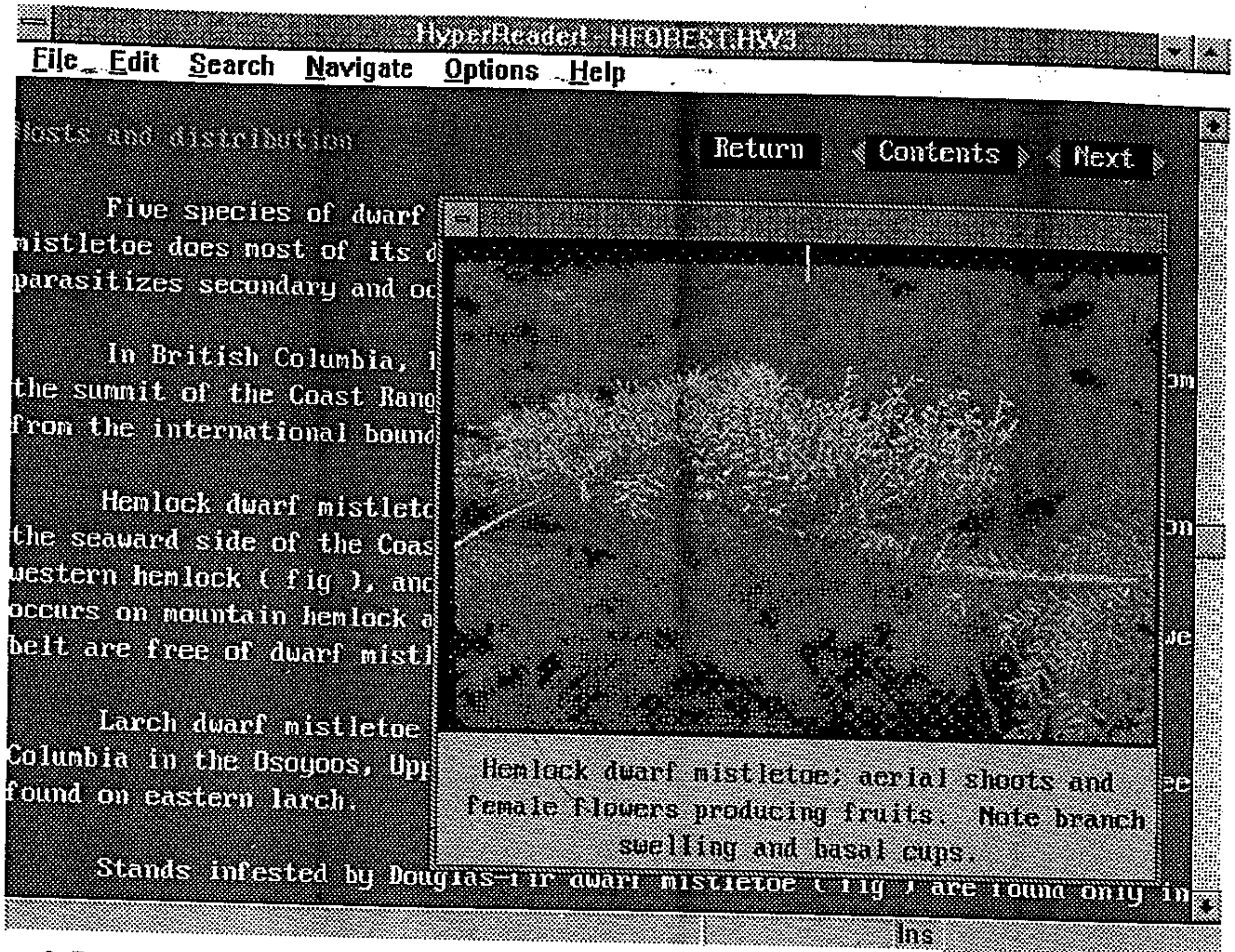

Figure 2. *Example of a link to a figure, from within the "Hosts and distribution" topic of the "Dwarf Mistletoes" section.*

Table 1. *Insect topics included in the system. The spruce spider mite is also included in this list. Each topic is a separate Forest Pest Leaflet.*

Adelgids	Larch Sawfly
Ambrosia Beetles	Lodgepole Terminal Weevil
Balsam Twig Aphid	Northern Pitch Twig Moth
Balsam Wooly Aphid	Northern Tent Caterpillar
Boxelder Bug	Poplar-and-Willow Borer
Cooley Spruce Gall Aphid	Redwood Bark Beetle
Cypress Leaf Tier	Satin Moth
Cypress Tip Moth	Sequoia Pitch Moth
Douglas-fir Beetle	Silver-spotted Tiger Moth
Douglas-fir Tussock Moth	Spruce Aphid
Engraver Beetles	Spruce Beetle
European Pine Shoot Moth	Spruce Budworms
Fall Webworm	Spruce Spider Mite
Forest Tent Caterpillar	Spruce Weevil
Golden Buprestid	Warren's Root Collar Weevil
Gray Spruce Looper	Western Balsam Bark Beetle
Green-striped Forest Looper	Western Blackheaded Budworm
Gypsy Moth	Western Cedar Borer
Juniper Scale	Western False Hemlock Looper
Juniper Webworm	Western Hemlock Looper
Larch Casebearer	

Table 2. *Disease topics included in the system. Each topic is a separate Forest Pest Leaflet.*

Annosus Root Rot	Needle Rust of Lodgepole Pine
Armillaria Root Disease	Needle Rusts of True Firs
Atropellis Canker	Phellinus weirii Root Rot
Black Stain Root Disease	Phomopsis Canker of Douglas-fir
Broom Rusts	Pine Needle Casts and Blights
Canker Diseases of Spruce	Poplar Shoot Blight
Common Needle Diseases of Spruce	Rhabdocline Needle Cast
Corky Root Disease	Rhizina Root Rot
Dwarf Mistletoes	Spruce Cone Rusts
Elytroderma Disease of Pines	Stem Rusts of Pine
Foliage Diseases of Western Larch	Tomentosus Root Disease
Fusarium Root Rot	True Fir Blights
Lophodermium Needle Cast of Pines	True Heartrots
Melampsora Foliage Rusts	White Pine Blister Rust
	Wound Parasites

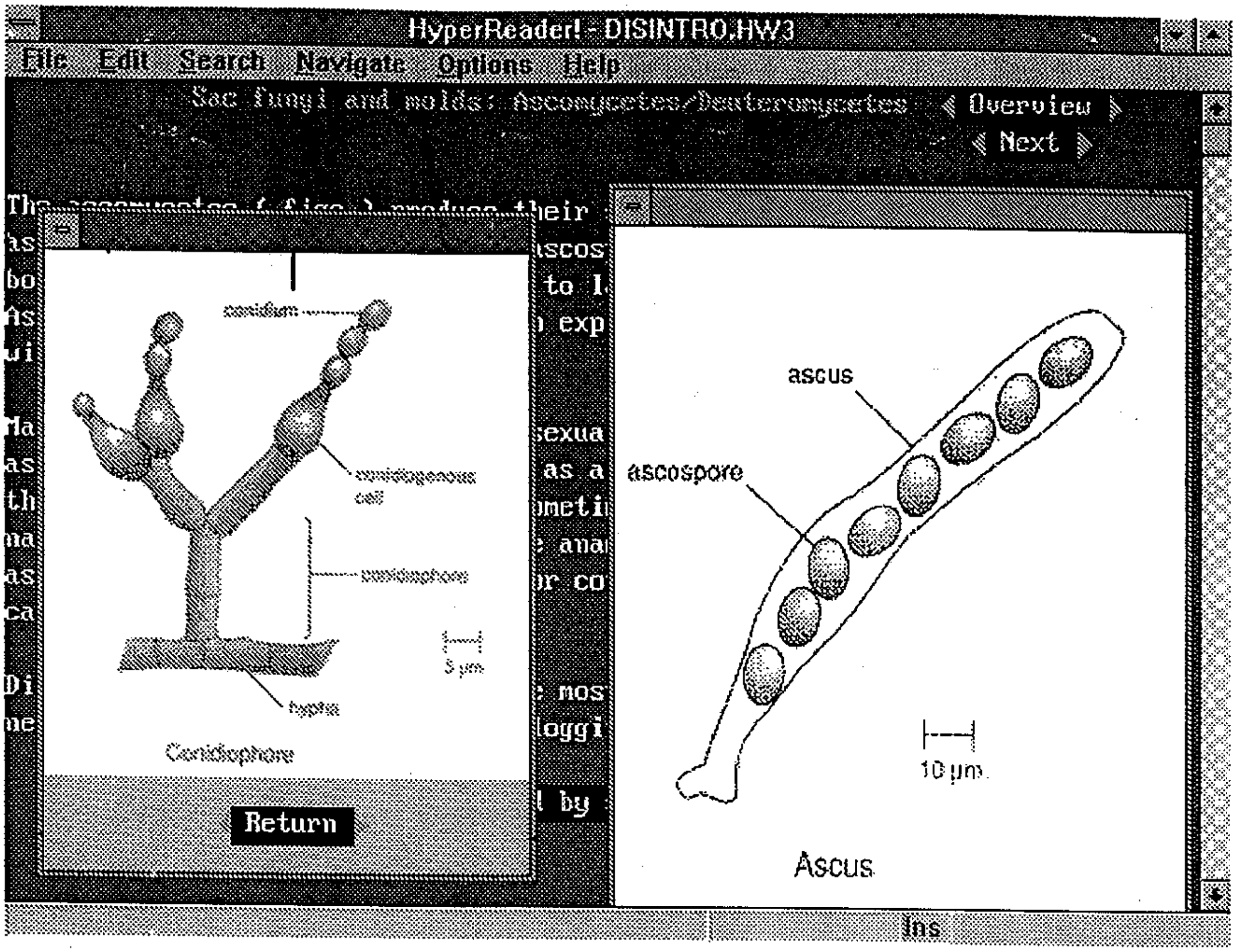

Figure 3. *Line drawings from the publication* Introduction to Forest Diseases *(Callan and Funk 1994), included in the system.*

value in the range 0-9, with zero indicating that the problem never occurred under the specified condition. The likelihoods in HYPERNUR were assigned based on a problem-oriented viewpoint compared to a topic-oriented viewpoint in HFOREST.

Thomson et al. (1992) describe the difference between these two viewpoints as follows. Consider a rare disease (problem) A which can occur on hosts B and C. From the problem's viewpoint, if it occurs more often on host B than on host C, its likelihood is high on B and low on C. However, from the point of view of the current topic (host in this case), because the problem is rare compared to other problems on these hosts, the likelihood would be assigned a low value on both hosts B and C.

A threshold (0-8) was selected by the user to define the sensitivity of the diagnosis. In the present system (HFOREST), threshold-setting was in relation to problem rarity (Figure 4). Filtering of possibilities, in a manner similar to that in the nursery system, could be set to consider only the common problems, all except rare problems, or all problems. To facilitate this approach, likelihoods were assigned values of 0 (never occurs for this condition), 2 (rare), 5 (often occurs) or 9 (common with this condition) (Figure 5). It was much easier for the pest specialists to assign values in this way than to use the continuous 0-9 scale of the nursery system. The likelihoods therefore capture both the expertise of the entomology and pathology specialists at the Pacific Forestry Centre, and the many years of experience of the Forest Insect and Disease Survey Rangers, who evaluate the health of British Columbia's forests each year. There was no preliminary development of an expert system as in the nursery system; rather, each navigation screen with the appropriate likelihoods (Fig-

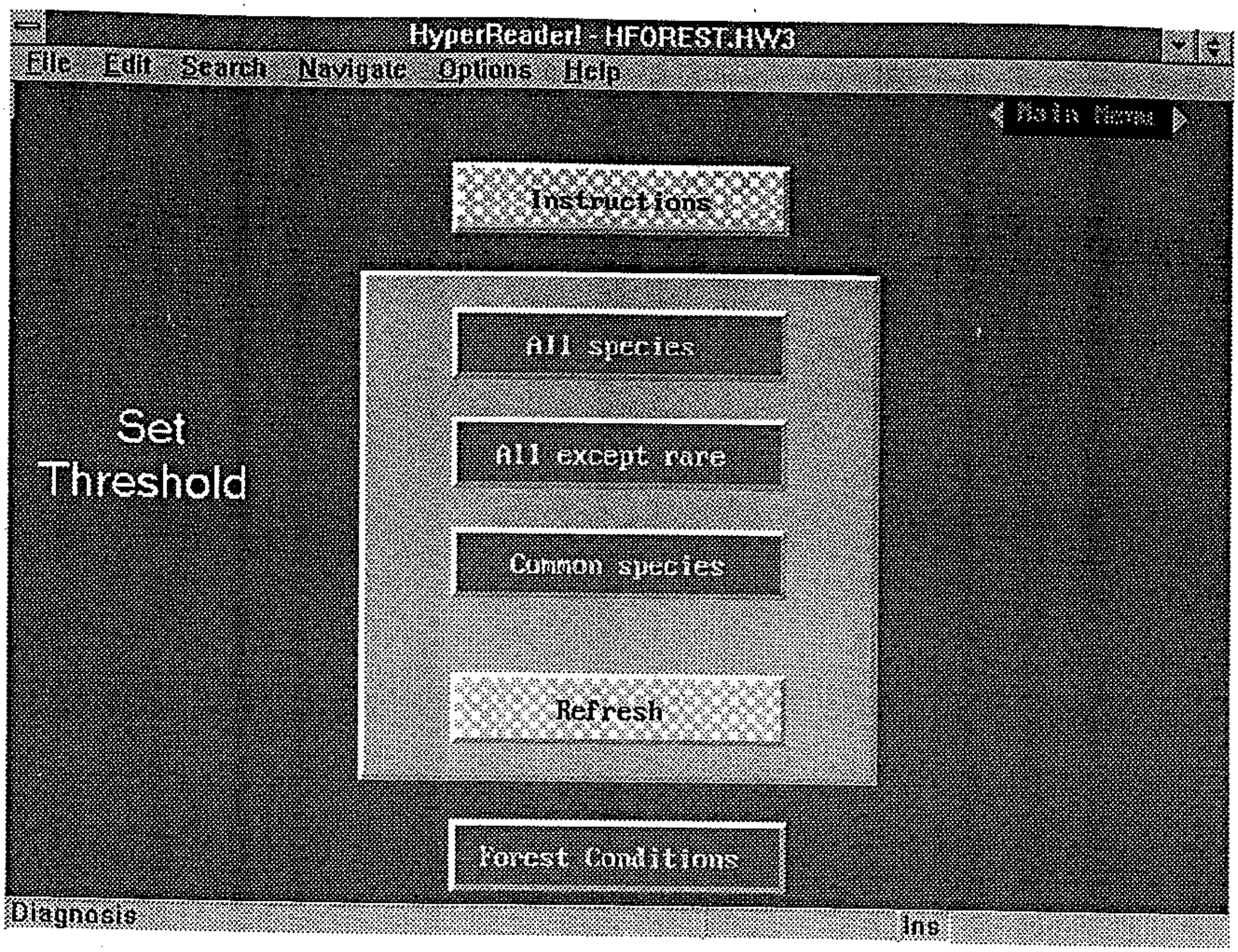

Figure 4. *Threshold-setting screen. For further description see text.*

ure 5) was edited directly into the hypermedia environment, using the categories of the earlier system as a guide.

To use the rare versus common approach required use of a host- or region-oriented viewpoint compared to the problem-oriented viewpoint of the nursery system. Thus, a rare disease for which the major host was Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) would have a likelihood of 2 for all hosts on which it occurred. In the nursery system, on the other hand, where likelihoods were based purely on the disease from its own perspective, a likelihood of 9 would be assigned to reflect the principal host status of Douglas-fir (Thomson et al. 1992), and lower likelihoods on its other hosts.

The buttons used for exporting the screen information for use in the filtering process (Figure 5) are similar to those in the nursery system, with the addition of an "Unfilter" button. If the list of possible diagnoses becomes empty or the suggested possibilities are not borne out by checking of details in the text material, then the "Unfilter" button can be used to restore possibilities previously removed by use of that topic.

The approach taken with signs and symptoms was similar to the nursery system, with a pest being indicated if it exhibited the sign or symptom under any conditions. Signs and symptoms were grouped under a range of categories, including foliage discoloration, crown damage, crown deformity, root problems, fruiting bodies, wood stains, wood decay, mortality, and other signs and symptoms.

As an example of applying filters successively, consider the following sequence, with the threshold set at zero so that all species are considered, so long as they occur with the specified condition. First,

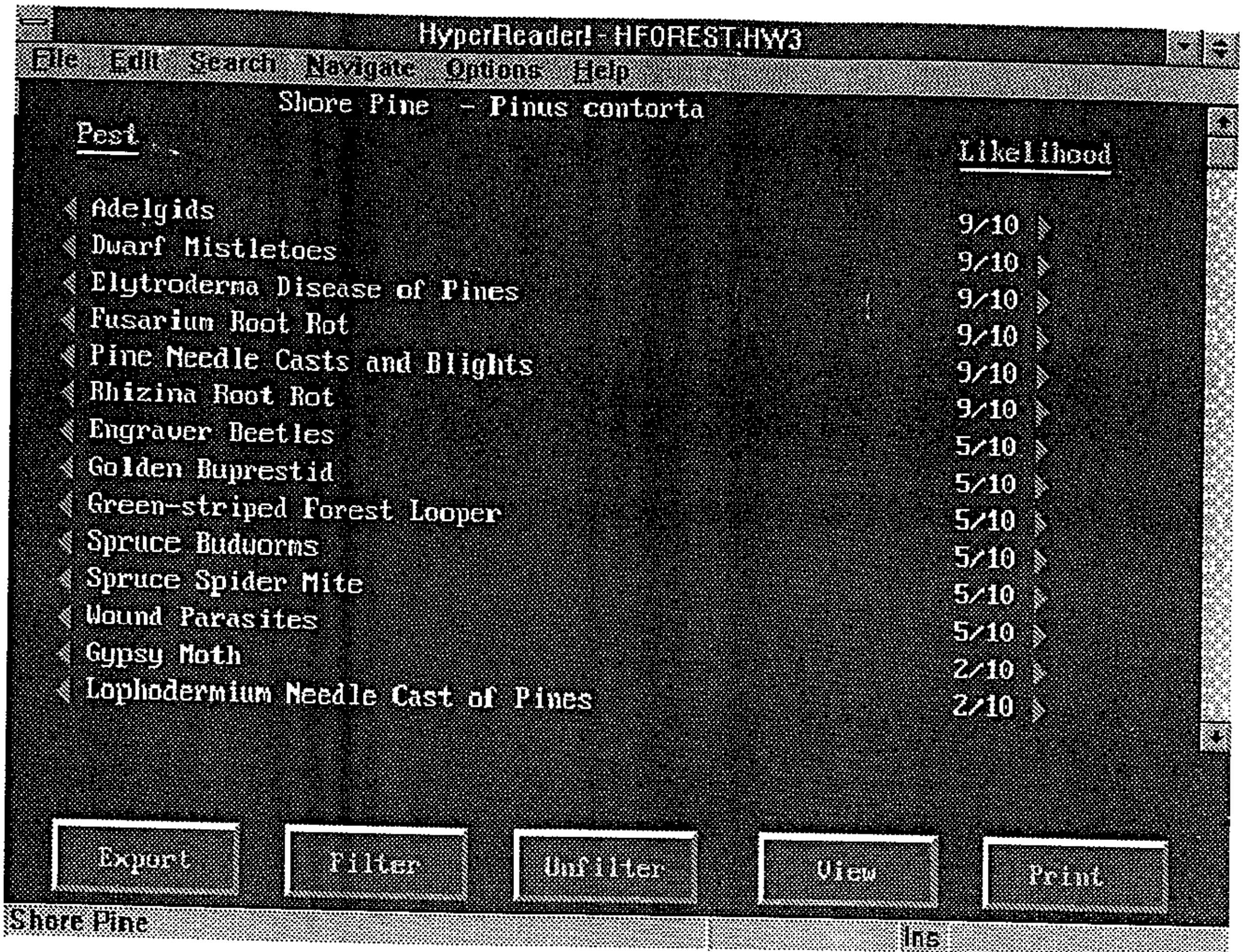

Figure 5. *Likelihood list for the host species "Shore Pine." Note that insects or diseases with zero likelihood for the specified condition are not listed.*

filter on the host species lodgepole pine (*Pinus contorta* var. *latifolia* Engelm.), and the Nelson Region. Next select foliage discoloration under signs and symptoms and filter based on yellow foliage. Finally select fruiting bodies, and filter on fruiting bodies on needles. This sequence results in two possibilities (filtered out zero times) (Figure 6). A further eight problems had been removed by a single filter, while many of the problems (not visible on the screen in Figure 6) had been filtered three or four times.

Discussion

Diagnosis is a knowledge-based process. Traditionally, human knowledge has been recorded as text on paper, with the reader extracting relevant facts and organizing them mentally to shed light on some question. During this process, the reader weights the recorded facts for relevance, error, bias, and accuracy. Knowledge can be recorded in a similar form in computers and accessed through hypertext systems (Rauscher et al. 1993a, b).

Knowledge can also be expressed in the computer as collections of specific facts, and processed by expert systems that perform relevance-checking, bias and accuracy evaluation, organization and synthesis activities in a fashion similar to human knowledge processing (Thomson and Taylor 1990, Thomson and Williamson 1992). Hypertext and expert systems together form a very powerful combination for knowledge representation and processing (Rauscher and Host 1990).

In certain domains, it is possible for a diagnostic expert system to determine if a particular set of ob-

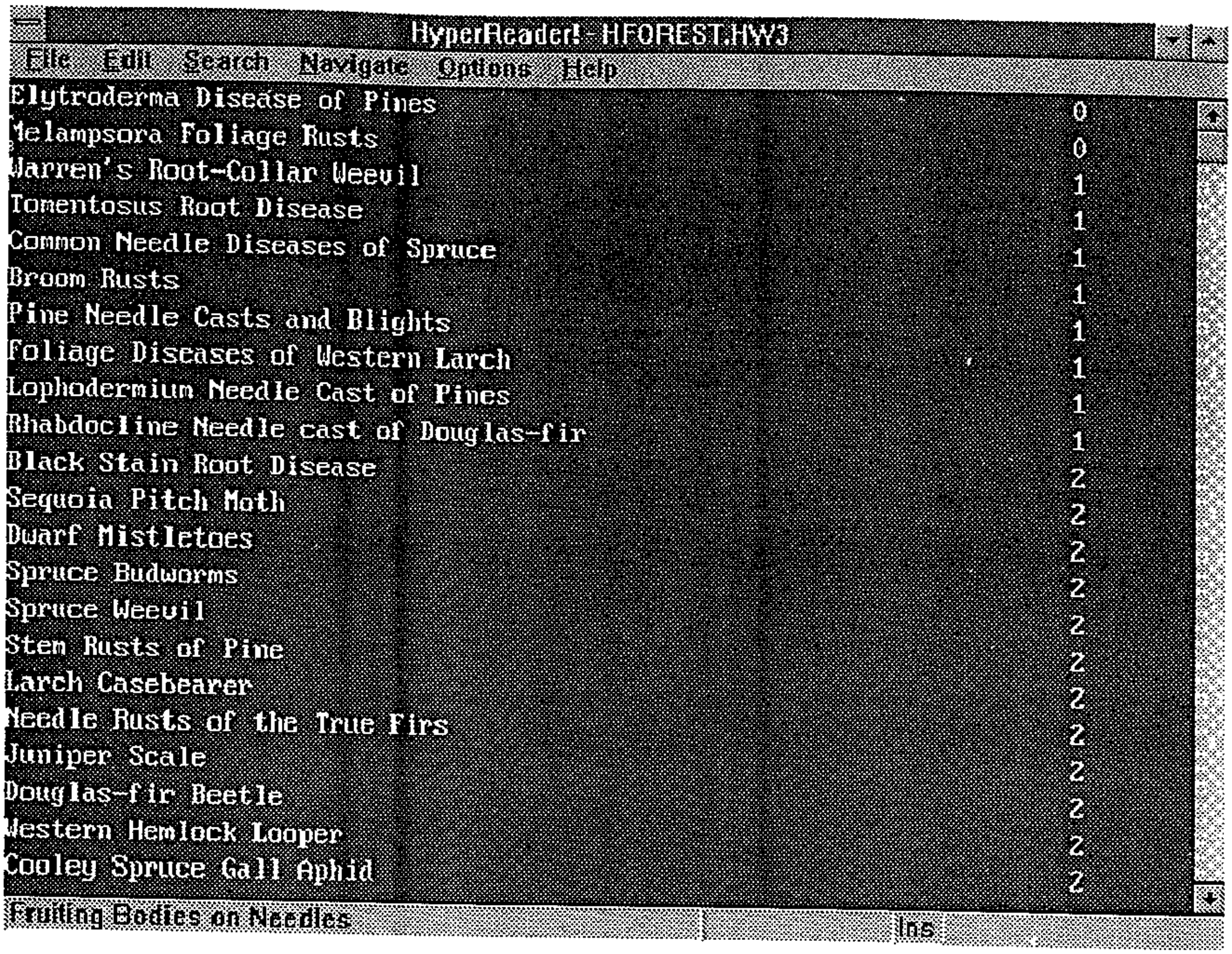

Figure 6. *Example of the results of successive filtering, displayed by selection of the View button of Figure 5.*

servations suggests an extra-domain problem (Thomson and Taylor 1990), but in many cases, computer-assisted diagnostic systems must contend with the difficulty of dealing with extra-domain possibilities. In an earlier system for diagnosing problems in forest seedling nurseries (Thomson et al. 1993), we showed that by combining hypermedia and expert systems, we can overcome many of the limitations of expert systems used alone by using the expert system to guide the system user to appropriate knowledge, then leaving some of the relevance-checking, bias and accuracy evaluation, organization and synthesis activities to the system user.

In developing the HFOREST system, we changed some of the approaches used in the nursery system. A fundamental difference between the two systems was in the viewpoints used to assign likelihoods. The nursery system's problem-oriented viewpoint was replaced in the forest system by a host- or region-oriented viewpoint. The latter viewpoint was more appropriate for use with the concept of "common" versus "rare" on which the filtering was based.

A second major difference was in the presentation of the results of the diagnosis. In the nursery system, as problems were removed by filtering, only the list of possible solutions was retained. In the forest system, however, all problems are listed, with the number of times each has been removed by the filtering process provided as an indicator of the confidence which can be placed in that result. This feature, combined with the ability to "Unfilter" a particular set of likelihoods, permits more diagnostic power.

The last major difference was in the provision of 331 high-quality digital color images and 30 black and white drawings, which required that the forest system be delivered on CD rather than disk. The rapid increase in the general use of multimedia systems since the production of the nursery system makes this approach valid for the proposed audience.

While the main emphasis of the system is on diagnosis and pest information, once a solution is reached, the material included in the system gives a discussion of the management options. Selection of particular management options can be aided by separate expert systems (Thomson and Taylor 1990, Thomson and Williamson 1992).

Acknowledgments

This study was supported in part by the Canada/British Columbia Forest Resource Development Agreement study FC-ETT 17.

References

Busch, D.D. 1990. The complete scanner handbook for desktop publishing, PC edition. Dow Jones-Irwin, Homewood, Illinois.

Callan, B., and A. Funk. 1994. Introduction to Forest Diseases. Forest Insect and Disease Survey Pest Leaflet No. 54. Pacific Forestry Centre, Canadian Forest Service, Victoria, British Columbia, Canada. 16p.

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 4 (3): 58-61.

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

Rauscher, H.M., D.H. Alban, and D.W. Johnson. 1993a. Managing the global climate change scientific knowledge base. AI Applications 7 (4): 17-40.

Rauscher, H.M., D.H. Alban, D.W. Johnson, C.V. Worth, and K.A. Berg. 1993b. The encyclopedia of hypertext. AI Applications 7(4): 41 (with disc: 1,990KB; 460 chunks, 950 links).

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. 85p.

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 4 (1): 44-52.

Thomson, A.J., and G.A. Van Sickle. 1995. Computer-based training in tree damage assessment. AI Applications. In Press (with disks).

Thomson, A.J., and D.R. Williamson. 1992. Formation and use of intermediate inferences in advisory systems: A herbicide example. AI Applications 6(4): 29-37.

Thomson, A.J., J.R. Sutherland, and C. Carpenter. 1993. Computer-assisted diagnosis using expert system-guided hypermedia. AI Applications 7(1): 17-27 (with disk).

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.

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. Working Party S2.07-09. Victoria, B.C., August 1990.

Unger, L. 1992. Dwarf mistletoes. Forest Insect and Disease Survey Pest Leaflet No. 44. Pacific Forestry Centre, Victoria, British Columbia, Canada. 7p.

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, 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 and is currently developing such systems in a forestry setting. He is presently a member of the Advanced Forest Technologies program at the Pacific Forestry Centre.

Allan Van Sickle is a research scientist at the Pacific Forestry Centre where he has been head of the Forest Insect and Disease Survey (FIDS) since 1979. As survey head, Dr. Van Sickle's focus is the development and application of techniques to detect, monitor, and assess the impact of forest pests. He also ensures that annual regional and national reports of pest conditions are accurate, concise, and timely.

Ordering HFOREST

HFOREST, an interactive multimedia compact disk diagnostic system, was created to supplement the Forest Pest Leaflet series produced by the Canadian Forest Service (CFS). This series describes the major forest insects and diseases in British Columbia.

System Requirements:

80386-based (or greater)
DOS-compatible computer
2X CD-ROM drive
3.5" 1.44MB FDD
Microsoft Windows 3.1 or greater
4MB of RAM and 4MB Hard Drive space
VGA monitor or better
16,256 and 16 million color graphics options are provided for use with the appropriate hardware.

To order a copy of the HFOREST CD, please send your name and address and a check for $10.00 (in Canadian dollars) made payable to the Receiver General of Canada to:

Distribution Clerk
Pacific Forestry Centre
506 West Burnside Road
Victoria, B.C., Canada
V8Z 1M5