AN EXPERT SYSTEM TO ASSIST TURFGRASS MANAGERS IN WEED IDENTIFICATION

by

Thomas F. Fermanian
Ryszard S. Michalski
Bruce Katz

Presented at the Proceedings of the 1985 Summer Computer Simulation Conference, Chicago, Illinois, July, 1985.

SCSC '85

The Proceedings of the

1985 Summer Computer Simulation Conference

July 22-24, 1985
Chicago, Illinois

SCS

Sponsored by The Society for Computer Simulation

Distributed outside the USA/Canada by
North-Holland (A division of Elsevier Science Publishers B.V.)
Amsterdam, The Netherlands

An Expert System to Assist Turfgrass* Managers in Weed Identification

Thomas W. Fermanian
Department of Horticulture

Ryszard S. Michalski and Bruce Katz
Department of Computer Science

University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

Expert systems technology offers a new approach to the identification of weeds found in turf. The dichotomous key, a traditional tool used for weed identification, is generally available in a printed form. To successfully identify an unknown weed using a key, appropriate identifying characteristics must be present on the weed and recognizable. Partial or erroneous information results in non-identification of the weed with no suggestion of a possible solution.

An expert system provides the ability to make a knowledgeable decision on the identification of an unknown weed when only fragmented or partial information is available. The system can also justify its conclusions providing an indication of the degree of certainty of the results. This paper describes an expert system PLANT/tm for the identification of weeds in turf. PLANT/tm was developed using micro/ADVISE, a set of tools for designing expert systems on microcomputers. Micro/ADVISE has been implemented on an IBM PC. It is a subset of the ADVISE Meta-Expert System, developed at the Artificial Intelligence laboratory at the University of Illinois at Urbana-Champaign (UIUC) and implemented on a VAX 780.

The Knowledge Base contains rules for the identification of 39 grassy weeds commonly found in turfs throughout the United States. Rules can be prepared with any editor and converted to the appropriate form through a program within the Knowledge Acquisition Block. The ability to develop and edit the Knowledge Base within the PC environment is a significant improvement over previous systems (PLANT/ds) which required a mainframe environment for knowledge acquisition. Preliminary testing of the performance of PLANT/tm has shown a high degree of reliability in recongizing unknown weeds.

INTRODUCTION

An expert system is a computer program that contains formally encoded knowledge of experts in a given problem area or domain and is able to use this knowledge to provide help to a non-specialist in problem solving in that domain. Expert systems differ from conventional computer data base management systems in several important aspects:

1. They can formally represent not only specific facts, data, statistics, etc. but also inference rules, decision rules, general information about the subject, "rules of thumb", uncertain information, etc. All of this information is contained in the Knowledge Base of the expert system.

2. They can conduct formal inferences on such information. The process is performed by the Query Block.

3. They can explain their inference process to a non-specialist by pointing out the decision rules, sources of information, and lines of reasoning used in arriving at the given advice.

The essence of expert systems is that they can manipulate and reason using complex symbolic structures rather than collections of numbers as do ordinary computer programs. Key concepts in the design of rule-based expert systems are summarized below (1):

1. Nature of Problem: The problem should be chosen with a narrow enough scope to be of reasonable size, but complex enough to require expertise.

2. Representation: Task-specific knowledge is separated from the method of using the control information to maintain flexibility and understandability of the Knowledge Base. Inclusion of very specific peices of knowledge, as well as general rules is required.

3. Inference: Since much of an expert's knowledge is heuristic in nature both deductive and plausible reasoning should be utilized. Problem solving strategies should remain independent of both the knowledge base and inference methods. This separation is critical for efficient debugging and maintenance of the program. An interactive system is important to facilitate a thorough examination of the problem.

4. Explanation: An expert system must handle not only static queries of the Knowledge Base but also dynamic queries requiring the system to document and explain the line of reasoning utilized to arrive at the final conclusion.

5. Knowledge Acquisition: Encoding domain knowledge into the knowledge base is best expedited through the communication between the "expert", who provides the expertise on the subject, and a "knowledge

*This work is supported in part by grants from ONR no. N00014-82-K-0186, National Science Foundation no. NSF DCR 84-06801, project No. 65-0267 of the Agric. Exp. Stn., College of Agric., Univ. of Illinois at Urbana-Champaign, and the Univ. of Illinois Research Board.

engineer" who understands the formalism for knowledge representation and the process of constructing an expert system.

6. Validation: Both static evaluation and empirical measures of adequacy should be utilized to evaluate the accuracy and completeness of the system.

There appears to be three main motivations for continuing research in the area of expert systems. First, is the ultimate replication of expertise. For the case at hand, this will provide a weed expert at any site for immediate access to the expertise of weed identification in turf. Secondly, an expert system allows the union of several experts in one system. Finally, expert systems provide a clear documentation of the decision-making process. This documentation can be used to expose areas in the domain where expertise is lacking or inadequate. Further research could then be conducted in these areas to broaden the Knowledge Base.

DESCRIPTION OF THE PROBLEM

The control of weeds in a turf landscape represents a significant portion of the total turf care budget. Management of weed species found within the turf environment is only possible through their correct identification. Many methods are normally used for this identifying process. Turf managers often rely on experience to recognize common visual clues for a rough or quick identification. Often times inexperienced turf managers misidentify weed species or have a hard time in determining the correct identification. Even highly experienced and knowledgeable turf managers can experience problems in identifying a rare or uncommon weed. When this situation arises the turf manager may either seek the help of a weed control specialist or consult a written key for plant identification.

Many keys have been written to assist in the appropriate identification of all flora found in a specific area. These keys generally take the form of dichotomously branching statements. The reader is offered the choice between two or more statements and must determine which statement most correctly identifies a structure on the unknown plant. After the appropriate identification of the structure, the statement will direct the reader to the next appropriate group of statements where an additional clue will need to be recognized. If within this process, the appropriate statement cannot be identified due to an inability to correctly identify a structure, the final identification of the species will not be possible.

In order to facilitate correct identification of unknown species with partial or incomplete information, an expert system (PLANT/tm) was developed to allow a non-weed scientist to identify a weed by using only those characteristics which were readily visible on the unknown plant sample. The system will then make a decision on the appropriate identification of the weed based on available inputs.

In general, there are 100 to 150 weeds considered to be pests in turf. In this group the grass family is represented by 39 species which were selected for the initial system. The grass weeds were sub-divided into two groups: those with rolled vernation or those with folded vernation (Figure 1). Each group was placed in a separate Knowledge Base to achieve higher efficiency in identification. Vernation is the configuration of expanding leaf blades in the bud shoot. Shoot vernation is an example of the type of characteristic the user will be required to identify.

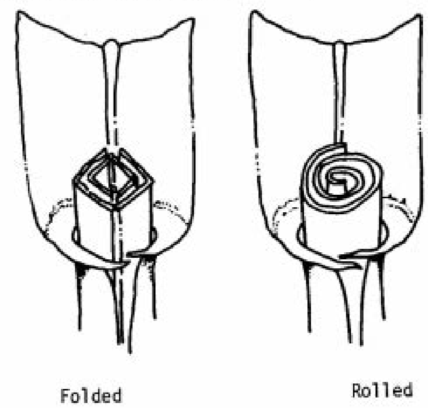

Figure 1. Vernation of the bud shoot

Due to frequent defoliation through mowing, most plant species in a turf remain in a vegetative state or lose their flower parts. It was felt that the use of vegetative characteristics for identification would be more appropriate and would be used most often for identification. Floral characteristics were also added to each rule in the event they were present on the unknown plant. When flowers are present the identification process is simpler and more reliable.

Initially a large matrix was developed with each potential grass weed representing a row with 11 columns of identifying characteristics, vegetative or floral, for each species. After information was collected from appropriate sources (text books, weed identification manuals, botanical handbooks, and expert experience), a rule was developed for each weed.

DESCRIPTION OF PLANT/TM

PLANT/tm is a unique program which attempts to capture all the capabilities of a full-fledged expert system on a microcomputer. It is based upon the highly successful ADVISE Meta-Expert System (2). Among its advanced features are the ability to easily add and modify rules in its knowledge base, the ability to create rules from examples, and the ability to test the performance of user-created rules, as well as their consistency and completeness. Figure 2 shows a conceptual outline of PLANT/tm.

The Control Block is a set of menus that allow the user to easily move within the system. From the menu of the program, the user has two choices. First, he can choose to receive advice from the system (described below). The advanced user can also choose to alter or examine the Knowledge Base. The Knowledge Base consists of two parts, the rule-base and knowledge tables. The rule-base consists of groups of rules organized by their function in the system.

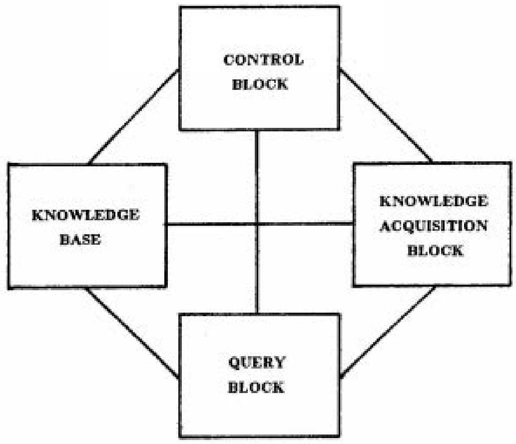

Figure 2. A block diagram of PLANT/tm

An example of a rule is:

Bermudagrass

If	1. Habit is rhizome_&_stoln,	(85%)
	2. Blade_width is fine to medium,	(60%)
	3. Ligule is ciliate,	(50%)
	4. Sheath is compressed,	(45%)
	5. Collar is narrow,	(25%)
	6. Auricle is absent,	(10%)
	OR	
	1. Flower is spike,	(80%)
	2. Florets = 1,	(75%)
	3. Glumes are shorter,	(35%)
	4. Disarticulate is above,	(35%)
	5. Awns are absent,	(15%)

Then Weed is Bermudagrass.

This is a rule which states that either of two sets of conditions leads to the conclusion. If the six vegetative or the five floral conditions are met, then the weed is identified as bermudagrass. The numbers to the right are degrees of certainty which indicate the relative importance of the elementary condition towards the conclusion. This degree of certainty represents the weed specialists degree of confidence or certainty that the condition supports the conclusion by itself. While the rule is complex, it takes the general form:

If CONDITION *then* DECISION

The CONDITION denotes or represents one or more elementary condition with an associated weight or degree or certainty. The CONDITION can be either a single conjunction of a complex of elementary conditions or a disjunction of two or more complexes. According to formal logic, the certainty of a conjunction of elementary conditions is the minimum of the certainties of the elementary conditions. Previous experiments (4) however, have shown higher performance of the system is achieved when the certainty of the conjunction is interpreted as the average of the certainties of the elementary conditions in the conjunction. The conjunction of a complex in PLANT/tm is the average of the degree of certainty of the elementary conditions.

Often times there is more than one way of describing a weed. In the above example, bermudagrass can be described by either its vegetative or floral characteristics. The satisfaction of these two complexes can be expressed as a logic sum or disjunction of the two complexes. For the disjunction of complexes, the certainty of such a disjunction is evaluated with the probabilistic sum (Psum) of the degree of certainty of the complexes. The average degree of certainty of each complex is used to determine the Psum. For two complexes, the degree of certainty (A) for the conjunction of the first complex and the degree of certainty (B) for the conjunction of the second complex:

$$\text{Psum} = A + B - AB$$

If there are multiple complexes, the formula can be repeated several times using the degree of certainty of the first disjunction with the degree of certainty of the next complex to determine the degree of certainty of their disjunction (Psum).

The Knowledge Base also contains two kinds of tables. First there is the variables table which allows the user to add or change different types of information about the variables in the system such as the variable type (nominal, linear, or structure) and the variable domain. In the above rule, the variables are habit, blade_width, ligule, sheath, collar and auricle for the first complex. The Knowledge Base also contains tables of examples in which variable values are associated with decision classes (bermudagrass is the class in the above rule).

The user is also provided with a program that makes it easy to enter these examples. The purpose of the examples is twofold. First, rules can be created from them using the GEM (generalization of examples by machine) (3) program which is a component of the Knowledge Acquisition Block. It is interesting to note that in a previous agricultural system developed at the UIUC, The GEM induced rules outperformed the rules supplied by an expert (4). Since this is not always the case, however, the user is also permitted to add rules to the system using an intelligent rule editor provided within PLANT/tm. The other purpose of the examples is to provide a set of events to test the user-provided rules. Thus the user may choose to use examples to test the performance of the rules even if it is not desirable to use rules that are created by induction from the examples. The ATEST module (5) that performs this test also allows the testing of the consistency and completeness of the rule-base.

The Query Block may also be chosen from the top-level menu so that the user can get advice from the rule base, which may contain induced rules, user supplied rules, or a combination of both. This module asks the user a series of questions to determine the confidence levels of the various rules in the Knowledge Base until all rules are either confirmed or unconfirmed, as determined by a preset confirmation level. The Query module works with a utility based evaluation scheme that asks questions in such an order as to receive the most amount of information to discriminate among the various rules.

Thus the user can alter the Knowledge Base by adding rules or modifying existing ones. The rule-base can be tested for performance, consistency, and completeness by utilizing a large number of examples of weed identifications. Finally, the user can receive advice from PLANT/tm based on responses to the system's queries.

FUTURE CONSIDERATIONS

Because the system was implemented on a relatively small computer which has limitations in memory and resources, it was not possible to develop a domain which was well defined. It is important that the system is defined very carefully involving sufficient detail and limiting the problem size to achieve efficiency on a microcomputer which is severely limited for this kind of a complex task. Expert systems are very complex and sophisticated systems. In building an expert system on a microcomputer, to make it efficient, the amount of knowledge placed in the Knowledge Base must be limited.

Secondly, in order for PLANT/tm to be useable in a practical field situation, the user will need assistance in identifying structures asked in the questioning process. Graphic support materials to show samples of various identifying structures will be necessary to aid in selecting the appropriate answers.

The ability of PLANT/tm to offer a possible solution when presented only partial information clearly exceeds the performance of any present dichotomous key. The success of this ability is due to the assignments of degrees of certainty to the elementary conditions and to the control schemes which direct the combining of more than one elementary condition. In addition, the system's ability to display these combinations of elementary conditions as an explanation of the decision making process, offers a component not available with a key. This provides for the best possible separation between suggested weeds with the least amount of information.

The Atest program will be utilized to test the performance of the rules, their consistency, and their completeness. This mode of evaluation, while not available for a dichotomous key, can help in improving the performance of keys by pointing out knowledge which is most important to the identification of a weed or knowledge which is currently missing from the key.

Initial evaluation of PLANT/tm has shown a high level of performance in the identification of an unknown weed. Depending on the character chosen, it is possible to identify a weed using a single characteristic. Further work is being planned for enlarging the rule base to incorporate broadleaf weeds or other pests.

REFERENCES

1. Buchanan, B. G. and R. O. Duda, "Principles of rule-based expert systems," Heuristic Programming Project Report No. HPP-82-14, Dept. of Computer Science, Stanford University, August, 1982.

2. Michalski, R. S. and A. B. Baskin, "Integrating Multiple Knowledge Representations and Learning Capabilities in an Expert System: The ADVISE System," Eighth International Joint Conference on Artificial Intelligence, pp. 256-258, 1983.

3. Michalski, R. S. and J. B. Larson, "Selection of Most Representative Training Examples and Incremental Generation of VL1 Hypotheses: The Underlying Methodology and Description of Programs ESEL and AQ11," Report No. 867, Dept. of Computer Science, University of Illinois, 1978.

4. Michalski, R. S. and R. L. Chilausky, "Learning by Being Told and Learning from Examples: An Experimental Comparison of Two Methods of Knowledge Acquisition in the Context of Developing an Expert System for Soybean Disease Diagnosis," International Journal of Policy Analysis and Information Systems, Vol. 4, No. 2, pp. 125-160, 1980.

5. Reinke, R. E., "Knowledge Acquisition and Refinement Tools for the ADVISE Meta-Expert System," Report No. 921, Dept. of Computer Science, University of Illinois, 1984.