Special Report

A Computer-Based Advisory System for Diagnosing Soybean Diseases in Illinois

R. S. MICHALSKI, Professor, and J. H. DAVIS, Graduate Research Assistant, Department of Computer Sciences; V. S. BISHT, Graduate Research Assistant, and J. B. SINCLAIR, Professor, Department of Plant Pathology, University of Illinois at Urbana-Champaign

Computer data bases that store facts and numerical data on a given subject have been developed in many fields. These systems allow a user to easily retrieve information but leave all the decisions about its use and interpretation to the user. They are unable to answer questions lacking stored answers or to relate the stored information to new information in order to provide advice. New types of computer-based information sytems called knowledge-based or expert systems are now being developed to overcome these limitations.

At the University of Illinois at Urbana-Champaign, an experimental expert system, PLANT/ds, was developed to provide consultation on the diagnosis of soybean diseases. This system is a component of a general system, PLANT, designed to advise users about the diagnosis and decision making regarding both crop diseases and insect damage.

An expert system consists of a "knowledge base" and an inference mechanism that conducts formalized reasoning involving information in the knowledge base and the user's answers to questions formulated by the system. The knowledge base contains general decision rules that represent the knowledge of experts on a given topic, eg, the diagnostic rules linking symptoms with diseases.

A typical form of a decision rule is "if CONDITION, then DECISION," written formally as

$$\text{CONDITION} :: > \text{DECISION},$$

where CONDITION is a list of conditions characterizing a situation or an object to which the rule is applied (eg, a diseased plant) and DECISION is a specific advice or action to be performed when the CONDITION is satisfied.

In general, the CONDITION part of the rule may be only partially satisfied. For example, the condition: *precipitation* = *normal* is partially true if precipitation was just above average. In this case, the system computes the "evidence degree," which is a numerical measure of the match between the CONDITION and a situation. Since the CONDITION may consist of several elementary conditions, and since each of the conditions may be satisfied to a different degree, the final evidence degree is a combined measure of the evidence degrees of the elementary conditions. This final evidence degree is taken as a measure of confidence in the correctness of the DECISION. The derivation of the combined measure is explained in the next section.

The DECISION part of a rule may be the assignment of the status "TRUE" to some conditions that are in the CONDITION part of another rule. Consequently, the satisfaction of a rule may cause the satisfaction of one or more other rules, and in this fashion, the system can perform a chain of inferences.

Most expert systems developed to date are in the experimental phase and typically address some relatively narrow but important practical problem (1–3,5,8–12). Some aspects of formalized agricultural decision making have been described (4).

In the following sections we describe the methodology for developing the knowledge base of the PLANT/ds, the method applied for using the rules in the knowledge base for providing diagnostic advice, and, finally, the results of experimental testing of the system. For a more technical discussion, see Michalski and Chilausky (5) and Michalski et al (7). A portion of this material is summarized in an abstract (6).

Representing Diagnostic Knowledge

Diagnostic knowledge is represented as decision rules, which specify all conditions indicating each disease. Such rule representation makes it easy to comprehend the conditions leading to a given diagnosis and to correct or refine knowledge in this form. It also facilitates an incremental extension of the knowledge base and an explanation of the inference process to a user.

Specification of descriptors. The first step in building the knowledge base was to determine the "descriptors" or "variables" most appropriate for the diagnosis of the diseases. The choice of descriptors depended both on the relevancy of descriptors to diagnosing soybean diseases and on the ease of reliably determining their values in the field. We chose those descriptors which would be easy to specify by a grower with no special training in plant pathology and with no special tools (eg, a microscope). Forty-one descriptors were selected to characterize diseased soybean plants and the environment affecting their development. The study included 17 common soybean diseases in Illinois.

For each descriptor, a "value set" was specified containing all possible values, ie, individual diagnostic symptoms, the descriptor may take for any diseased plant. In determining such value sets, it is important to avoid excessive precision, ie, to

This study was part of Project No. 1-6-55552 of the Agricultural Experiment Station, College of Agriculture, University of Illinois at Urbana-Champaign. It was supported in part by grant 321512344 from the U.S. Department of Agriculture.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. § 1734 solely to indicate this fact.

©1983 American Phytopathological Society

distinguish between values only if such distinction may make a difference in diagnostic decisions. For example, the values of the descriptor: "condition of leaves" were chosen simply as "normal" and "abnormal." When more specific information is needed on this subject, other descriptors are used. For example, we used the descriptor "leaf spots," which has the values "absent," "present," "with yellow halos," or "without yellow halos."

If "condition of leaves" is "normal," then all the descriptors describing leaf abnormalities have the value "does not apply," and the system never asks for their values. Such relationships among values of particular descriptors impose restrictions on the "problem description space." Such a space is defined as the set of all possible combinations of the values of all the descriptors. If the system detects a combination that is not allowed (owing to restrictions as above), an error signal is produced.

Also taken into consideration are relationships among the values of the same descriptor. Depending on the type of the relationship, a descriptor is "nominal," "linear," or "structured." For *nominal descriptors* it is assumed there is no relationship among their values. "Condition of leaves" (normal or abnormal), "presence of fruiting structure" (yes or no), and "seed shriveling" (absent or present) are examples of nominal descriptors. *Linear descriptors* have values that can be arranged in a linear order, like numbers; given any two distinct values, one is smaller and one is larger. "Time of occurrence" and "number of years crop repeated" are examples of linear descriptors. When more complicated relationships are needed, structured descriptors are used. Whereas values of a linear descriptor can be placed along a line, *structured descriptors* have values that can be placed as nodes of a treelike structure, ie, a hierarchy. "Leaf spots" is such a descriptor. At the first level of the hierarchy its values are simply "absent" or "present." At the second level the value "present" is divided into the cases with and without yellow halos. Thus, if a rule stated that a disease would have leaf spots present and the user of the system said his field had leaf spots with yellow halos, that part of the rule would be satisfied.

Basic Building Blocks of Decision Rules

Relational statements: selectors. The most elementary component of decision rules is a relational statement, called a selector, that formally expresses a single condition. In its simplest form, a selector specifies that a given descriptor should take only one of its possible values, as in [condition of leaves = abnormal]; square brackets always surround a selector. If the user indicates the specimen has "abnormal" leaves, this selector is "satisfied" and assigned the evidence degree "1." If the user indicates the leaves are "normal," the selector is "not satisfied" and the assigned evidence degree is "0." In a more general case, a selector may allow a variable to take more than one value, eg, [canker lesion color = brown, tan]. In this case, the selector is satisfied only if the user indicates either brown or tan as the canker lesion color. With linear variables, a range of values can be specified instead of listing the values individually, eg, [time of occurrence = June . . . September] or [precipitation ⩾normal].

Suppose an expert wants to indicate that bacterial pustule is most likely to occur in August, less likely to occur in July or September, even less likely to occur in June, and quite unlikely to occur in any other month. To express such information, a more general form, a "weighted selector," is used that allows one to specify the degree of evidence associated with each value of the descriptor. The *evidence degree* may range between "0" and "1," where "0" indicates no evidence and "1" indicates the maximum evidence. Thus, if 0.8 is the evidence value characterized as "less likely," 0.6 as "maybe," and 0.2 as "quite unlikely," a weighted selector expressing the above information about bacterial pustule would be: [time of occurrence = August: 1; July, September: 0.8; June: 0.6; else: 0.2]. This way of representing a dependence of the evidence degree on the different values of a descriptor is closely related to the so-called fuzzy algorithmic definition of imprecise concepts (13).

Conjunctive conditions: complexes. A very common form of a decision rule is one stating that several conditions must be simultaneously satisfied in order to support a DECISION. These conditions are formally represented as the logical product (conjunction) of selectors, called a *complex*. The complex is expressed either by concatenating (placing side by side) all the selectors involved (for implicit conjunction) or by joining them by the symbol Λ (read as "and") when more than one line is needed to write them down. For example, a complex describing "purple seed stain" is:

[time of occurrence = September, October] Λ
[condition of seed = abnormal] [seed discoloration = purple].

When all the selectors above are satisfied, the evidence degree has the maximum value. Suppose, however, that only two selectors are satisfied and the first one is not because the time of occurrence was August rather than September. Using the classical logical interpretation of conjunction, if one condition is not satisified, the whole complex is not satisfied. Since August is "close to" September, such an interpretation would be too rigid. A more flexible interpretation is needed. A way to handle this is to change the selector into a weighted selector specifying weights for each value. In this case, one would add August to the selector above with a weight slightly smaller than "1."

Another solution is to specify a *default weighting function* for selectors involving linear descriptors. In this case, an unweighted selector is evaluated as if it were a weighted selector, with a standard weight function of a "bell" form, which assumes the maximum value for values specified in the selector and continuously decreases on both sides. Since August is "close to" September, the weight computed from such a bell function would be slightly less than "1."

A problem now arises of how to combine the degrees of evidence provided by each selector into the degree of evidence provided by the whole complex. We investigated three evaluation schemes for the conjunction:

- PROD: the *product* function, which computes the evidence degree of a complex as the arithmetic product of the evidence degrees of its selectors;
- MIN: the *minimum* function, which computes the evidence degree of a complex as the minimum of the evidence degree of its selectors; and
- AVE: the *average* function, which computes the evidence degrees of a complex as the average of all the evidence degrees of its selectors.

These functions satisfy the relation PROD⩽MIN⩽AVE, ie, given any two degrees of evidence, their product (PROD) will always be less than or equal to their minimum (MIN) and their minimum will always be less than or equal to their average (AVE). The best scheme to use depends on the particular problem under consideration, and the most reliable method for determining the best is experimentation.

Disjunctive statements: logical unions. A conjunction of selectors may not be sufficient to express complicated relationships, such as when two or more different conditions (complexes) indicate the same disease. For example, "brown spot" is indicated when either small leaf spots without yellow halos or large spots with yellow halos are present. This is expressed using the logical union (disjunction) of two complexes, each describing one of the alternatives:

[leaf spot size > 1/16″] [leaf spot = with yellow halos]
V
[leaf spot size < 1/16″] [leaf spot = without yellow halos].

The symbol V denotes logical "or." (It is assumed the conjunction of selectors is always evaluated first, before the disjunction.) The condition part of a rule that is a disjunction of complexes is said to be in *disjunctive normal form* (DNF).

As in the case of conjunction, there is more than one way of combining the evidence degrees of complexes into the evidence degree provided by their disjunction. There are two basic evaluation schemes for disjunction:

- MAX: the maximum function, which computes the evidence degree of a disjunction as the maximum of the degrees of the component complexes; and

- PSUM: the probabilistic sum, which computes the evidence degree of a disjunction of two complexes as

$$a + b - ab,$$

where "a" and "b" are evidence degrees of the component complexes. When there are more than two complexes in a disjunction, this rule is repeated the appropriate number of times.

These schemes satisfy the relation MAX $\leqslant$ PSUM. Again, the most reliable method for determining the best scheme is experimentation.

One of the two sets of diagnostic rules used in PLANT/ds is in DNF. This set of rules was obtained by applying a computer program capable of inductive inference (5,7). The program was given a few hundred examples of diagnostic decisions (diagnoses made by plant pathologists) and from them created the general DNF diagnostic rules (7). Here is an example of an inductively derived rule for diagnosing "purple seed stain":

[plant stand = n] [precipitation > n] [severity = minor] Λ
[plant height = n] [leaf spot halos =
no yellow halos] [seed = abn] Λ
[seed discoloration = p] [seed size = n]
V
[leaves = n] [seed = abn] [seed size = n] :: >
[diagnosis = purple seed stain].

The second set of rules used in PLANT/ds was obtained by directly expressing the plant pathologists' descriptions of how they make diagnoses. To express these rules, more advanced formal constructs than DNF were used, namely, implicative statements and linear modules.

Implicative statements. From the formal viewpoint, any logical condition can be expressed in the DNF form. However, when one wants to express in this form the diagnostic knowledge of plant pathologists, the DNF rules may be very long and have no direct relationship to the human descriptions. An important additional construct that facilitates expressing experts' descriptions is the "implicative statement." The implicative statement is used when one wants to state that if a condition(s) is present, some other condition(s) must also be present. Thus, if the first condition is not present, the other implied conditions are irrelevant. There are many instances of such relationships in the rules. For example, in the rule for "downy mildew" we have:

[time of occurrence = September, October] = >
[condition of seeds = abnormal] [seed mold growth = present],

where = > stands for implication. This condition states that if the disease is occurring in September or October, the seeds should be abnormal and appear moldy; during other months, these properties of the seeds are irrelevant.

The general form of the implicative statement is:

DNF expression = > complex.

The evidence degree of the implicative statement A = > B is computed by transforming it into a logically equivalent statement ~ A V B, where ~ A denotes negation of A and is evaluated as 1 − evidence degree A.

Linear modules. In describing a disease, it is sometimes important to express the idea that some symptoms are relatively more important than others for the diagnosis. For example, when diagnosing "downy mildew," the important symptoms are abnormal leaves, leaf spots without yellow halos, mildew growth on the lower leaf surface, and abnormal seeds with mildew growth if time of occurrence is September or October. If any of these conditions is not present, "downy mildew" is probably not the problem. On the other hand, presence of premature defoliation and presence of leaf malformation are confirmatory but not crucial conditions for the diagnosis. To express such relations, a construct called the "linear module" is used (5).

The linear module has the form:

$$q_1 \cdot C_1 + q_2 \cdot C_2 + q_3 \cdot C_3 + \ldots,$$

where $C_1, C_2, C_3 \ldots$ stand for conditions of the forms considered so far (ie, complexes, DNFs, or implicative statements) and $q_1, q_2, q_3 \ldots$ are coefficients indicating the relative significance of these conditions (q_i's vary between 0 and 1). A linear module is the most general form for expressing descriptions of diseases. To compute the evidence degree of a linear module, the C_i's are substituted by the evidence degrees of the corresponding conditions, the $\cdot$ is integrated as the arithmetic multiplications, and the + is integrated as the arithmetic addition.

In expressing the "expert-derived" diagnostic rules, two-part linear modules with coefficients denoted q_s and q_c for significant and confirmatory evidence, respectively, were used:

$$q_s \cdot C_s + q_c \cdot C_c.$$

Since coefficients q_s and q_c express the relative importance of the conditions C_s and C_c in the total description, they must sum up to 1: $q_s + q_c = 1$. In experiments with the rules, we assumed that $q_s = 0.8$ and $q_c = 0.2$. Thus, 80% of the evidence for a disease comes from the conditions in C_s and 20% comes from the conditions in C_c. If any condition in C_s is not satisfied, the evidence degree of the whole linear module will be greatly reduced. But if the same condition were in C_c and not satisfied, the evidence of the whole linear module would be reduced only slightly.

Complete rules. Diagnostic rules are expressed in the form:

CONDITION : : > DECISION,

where CONDITION can vary from a simple selector to a linear module and DECISION is a complex (conjunction of selectors) that describes the decisions assigned to the situation. In all the rules developed so far, the DECISION is a single selector specifying which disease the rule implies. The confidence in the DECISION is computed as the evidence degree of the CONDITION.

An example of a diagnostic rule with CONDITION as a linear module is given later.

Evaluating Decision Rules

To diagnose the disease of a plant, rules in the knowledge base are evaluated using values of descriptors obtained by questioning the system's user. The system asks questions in a dynamically changing order, trying to achieve the diagnosis with the minimum number of questions in each use. At each stage the question asked is selected to maximally reduce the uncertainty of the diagnosis. To simplify the explanation of the system, we will assume that values of all descriptors for a diseased plant are already known. In this case, the system determines diagnosis by finding the rule that best matches the conditions of the plant, ie, the rule whose evidence degree will reach the maximum for the given values of descriptors.

In general, there may be more than one rule with the maximum degree of evidence or there may be rules whose evidence degrees differ only slightly. The system resolves this problem by giving alternative diagnoses when the evidence degrees vary from the maximum within an experimentally determined interval Δ ($\Delta = 0.2$, ie, 20% from the maximum). Also, to avoid erroneous advice, PLANT/ds suggests an advice only if the evidence degree is above a certain threshold. This threshold was determined experimentally as 0.65 for expert-derived rules and as 0.8 for inductively derived rules. Also, we found that the best evaluation scheme for the expert-derived rules was:

- AVE: the average function for conjunction, and
- MAX: the maximum function for disjunction.

For inductively derived rules, the best evaluation scheme was:

- AVE: the average function for conjunction, and
- PSUM: the probability sum for disjunction.

The evidence degree computed for the rules should not be taken as an expression of the statistical probability of the correctness of the diagnosis. The evidence degree expresses the degree to which a rule matches the description of the diseased plant (using a given evaluation scheme) and serves as an approximate indicator of the confidence in the advice.

To illustrate the rule evaluation process, we will evaluate expert-derived rules characterizing "downy mildew" and "powdery mildew" for some exemplary case. The values of the descriptors are:

Time of occurrence	=	August
Precipitation	=	normal
Temperature	=	normal
Damaged area	=	whole fields
Condition of leaves	=	abnormal
Leaf spots	=	without yellow halos
Leaf mildew growth	=	on upper leaf surface
Premature defoliation	=	present
Seed mold growth	=	present
Leaf malformation	=	absent
Condition of seed	=	normal
Condition of stem	=	normal.

First, we evaluate the diagnostic rule for "powdery mildew":

q_s · [time of occurrence = August . . . September]Λ
[condition of leaves = abnormal]Λ
[leaf mildew growth = upper leaf surface]
+
q_c · [precipitation < normal] [temperature > normal].

This rule states that significant indicators for "powdery mildew" (prefixed by coefficient q_s) are: late occurrence in the season (August or September), abnormal leaves, and mildew on the upper leaf surface. The confirmatory indicators (prefixed by q_c) are: less than normal precipitation and temperature at least normal, if not above normal.

Since the time is August, the first selector [time of occurrence = August . . . September] is satisfied and is assigned the evidence degree "1." The other two signficant selectors also are satisfied and so have the evidence degree "1." The combined degree of evidence is the average of individual degrees, ie, "1." The first selector in the confirmatory part is not satisfied, as the precipitation was normal. The selector [precipitation < normal] is false, so the degree of evidence is "0." The last selector, which involves temperature, is satisfied, so it has the evidence degree "1." The average of "1" and "0" is "0.5." Now the degree of evidence of the whole rule can be evaluated and is $(0.8 \cdot 1) + (0.2 \cdot 0.5) = 0.9$.

A similar calculation is done for the rule describing "downy mildew":

q_s · [time of occurrence: / T7] [precipitation > normal] Λ
[damaged area = whole fields] [condition of leaves = abnormal] Λ
[leaf spots = no yellow halos] [condition of stem = normal] Λ
[leaf mildew growth = on lower leaf surface]
([time of occurrence = September, October] = >
[condition of seed = abnormal] [seed mold growth = present])
+
q_s · [premature defoliation = present] [leaf malformation = present]
:: > [soybean disease = "downy mildew"],

Table 1. Summary of results from PLANT/ds diagnosing 340 experimental cases of soybean diseases

	Type of rules[a]	
	Expert derived	**Inductively derived**
Correct[b] (%)	96.90	100.00
First choice correct (%)	71.80	97.60
Not diagnosed (%)	2.10	0
Indecision ratio[c]	2.90	2.64
Threshold[d]	0.65	0.80

[a] Expert derived = performance of diagnostic rules obtained by formalizing plant pathologists' decision process; inductively derived = performance of diagnostic rules obtained by a computer learning process from several hundred cases of diseases diagnosed by plant pathologists.
[b] Percentage of decisions in which correct diagnosis was among alternatives indicated by the system.
[c] Average number of alternatives.
[d] Minimum degree of match between a diseased plant description and a diagnostic rule required for a diagnostic decision.

where T7 is defined:

$$T7 = \begin{cases} 1.0 - \text{July or August} \\ 0.8 - \text{June, September} \\ 0.7 - \text{October} \\ 0.1 - \text{other.} \end{cases}$$

This rule is more complicated than that for "powdery mildew." The significant indicators are that the disease is most likely to occur in the middle of the season, precipitation is normal to above normal, the whole field will probably be damaged, the leaves are abnormal, leaf spots are present but without yellow halos, the stem is normal, mildew is present on the underside of the leaves, and, if it is September or October, the seeds are abnormal, with mold growing on them. The corroborative indicators are that there may also be premature defoliation and the leaves may be malformed.

The value of the first selector is "1," which is determined by the function "T7." The evidence degrees of the next five selectors are "1," "1," "1," "1," and "0," respectively. The next three selectors are surrounded by parentheses and are a part of the implicative statement, asserting that if the disease occurred in either September or October, there should be abnormalities in the seed and there should be mold growing on the seeds. Since it is not September or October, this implicative statement is ignored by evaluating it as "1." The evidence degree provided by the "significant" part of the rule is then: $(1 + 1 + 1 + 1 + 1 + 0 + 1) = 0.86$. The evidence degree for "downy mildew" provided by the whole rule is: $0.8 \cdot 0.86 + 0.2 \cdot 0.5 \simeq 0.79$.

When we compare the two diseases, "powdery mildew" has a degree of evidence of 0.9, while "downy mildew" has 0.79. Both values are above the expert-derived threshold of 0.65, so neither is eliminated. "Powdery mildew" has a higher degree of evidence so it would be the primary diagnosis (assuming no other rule had a higher degree of evidence). Since downy mildew's value is within 0.2 of powdery mildew's, it would be an alternative diagnosis.

Suppose that mildew growth was on the lower instead of the upper leaf surface. This variable is found in the "significant" portion of both rules, so a change in the diagnosis can be expected. The degree of evidence becomes 0.64 for powdery mildew and 0.9 for downy mildew. The value 0.64 is below the expert-derived rule threshold, so powdery mildew is eliminated as a possible diagnosis, leaving downy mildew as the only diagnosis (unless there were some other diseases with higher degrees of evidence).

An Experiment

The rule evaluation process is embedded in two computer programs, a batch program and an interactive program. The batch program is used only for experimental purposes, to diagnose many diseases at once in order to evaluate the program's performance. The collected statistics about rule performance can help decide which rules need further refinement and which evaluation method is best. The output from the batch program consists of the diagnosis for each example grouped by disease and a summary section.

A summary of the results from the batch version using 340 cases of soybean diseases is given in Table 1. "Inductively derived" and "expert derived" refer to the two rule sets used, those obtained by inductive inference from examples and those obtained by formalizing plant pathologists' decision rules, respectively. "First choice correct" tallies how often advice of the PLANT/ds was correct, while "correct" tallies how often the correct diagnosis was either the first diagnosis or an alternative diagnosis. "Not diagnosed" gives an indication of how often the system could not identify a case of soybean disease. The "indecision ratio" is a measure of how unique decisions were, ie, the more alternative diagnoses, the higher the indecision ratio. It is defined as the ratio of the number of alternative diagnoses per case of disease. Hence, a low indecision ratio is desirable but does not imply correctness. "Threshold" refers to the minimum degree of evidence a rule must have not to be eliminated.

These statistics can be collected only when many known cases of disease are available. When a grower wants to use the system on a single (or a few) unknown cases, the interactive version of the program is used. This program is designed to treat cases one at a time. The system asks multiple-choice questions to which the user either gives an answer or uses one of several commands for help or explanation.

The order of the questions is not fixed but varies according to the user's responses. At any given point in the session, the question that will affect the most rules still under consideration is the next one selected. This is accomplished by keeping track of each use of a descriptor. When a rule is eliminated because its degree of evidence is too low, the usage counts of all the descriptors in that rule are decreased accordingly. This means that if a descriptor is not used in all the rules and if all the rules it is used in are eliminated, the system does not need to ask for its value. So in a typical session, the system needs to ask only about half the 41 questions.

This system advises users about soybean diseases on the basis of symptoms communicated to it by a user. The system is easy to use, as all the symptomatic information needed for the diagnosis consists of answers to multiple-choice questions provided by pushing on an appropriate key or, in terminals with touch panels, by touching an appropriate place on the screen. The system has been implemented in PASCAL language on the computer CYBER 175 and requires approximately 128,000 words of memory. Experimental results show a high level of correctness of the system's advice.

Since this paper is intended to be an introduction to PLANT/ds, many details and some less important features are not discussed. For example, we omit here a description of explanatory capability of the system, ie, various functions that permit a user to understand steps of the inferential process conducted by the system. A reader interested in this and other details about the system is referred to other publications (5,7).

ACKNOWLEDGMENT

The authors thank A. Boulanger, Department of Computer Sciences, University of Illinois at Urbana-Champaign, for editorial suggestions.

LITERATURE CITED

1. Buchanan, G. B., and Feigenbaum, E. A. 1978. Dendral and Meta-Dendral, their applications dimension. Artif. Intell. J. 11:5-24.
2. Davis, R. 1978. Knowledge acquisition in a rule-based system-knowledge about representation on a basis for systems construction and maintenance. Pages 99-135 in: Pattern Directed Inference Systems. D. A. Waterman and F. Hayes-Roth, eds. Academic Press, New York.
3. Duda, R. O., Hart, P. E., Knoligdie, K., and Rebou, R. 1979. A computer-based consultant for mineral exploration. Final Report, SRI International, Menlo Park, CA. 185 pp.
4. Gladwin, C. H. 1980. A theory of real-life choice: Applications to agricultural decisions. Pages 45-85 in: Agricultural Decision Making. P. Bartlett, ed. Academic Press, New York.
5. Michalski, R. S., and Chilausky, R. L. 1980. Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. Int. J. Policy Anal. Inf. Syst. 4:125-161.
6. Michalski, R. S., Davis, J. H., Bisht, V. S., and Sinclair, J. B. 1981. PLANT/ds: An experimental computer consulting system for the diagnosis of soybean diseases. (Abstr.) Phytopathology 71:242.
7. Michalski, R. S., Uhrik, C. T., Bisht, V. S., and Sinclair, J. B. 1983. A description of PLANT/ds—An expert system for soybean disease diagnosis. Dep. Computer Sci., University of Illinois at Urbana-Champaign. In press.
8. Michie, D., ed. 1979. Expert Systems in the Micro Electronic Age. Edinburgh University Press. 287 pp.
9. Michie, D. 1980. Knowledge-based systems. Dep. Computer Sci., Rep. 1001, University of Illinois at Urbana-Champaign. 33 pp.
10. Pople, M. E. 1977. The formation of composite hypotheses in diagnostic problem solving: An exercise in synthetic reasoning. Pages 1030-1037 in: Proc. Int. Jt. Conf. Artif. Intell. 5th. Vol. II. Massachusetts Institute of Technology, Cambridge.
11. Shortliffe, E. H. 1976. Computer Based Medical Consultation, MYCIN. Elsevier, New York. 264 pp.
12. Weiss, S. M., Kulikowski, C. A., Amarel, S., and Safir, A. 1978. A model-based method for computer-aided medical decision-making. Artif. Intell. J. 11:145-172.
13. Zodeh, L. A. 1976. A fuzzy algorithmic approach to the definition of complex or imprecise concepts. Int. J. Man-Mach. Stud. 8:249-291.