

Explanation and Acquisition
in Expert Systems
Using Support Knowledge

Alexander Pasik
Jens Christensen
Douglas Gordin
Agata Stancato-Pasik
Salvatore Stolfo

CUCS-164-85

Department of Computer Science
Columbia University
New York City, New York 10027

Explanation and Acquisition in Expert Systems Using Support Knowledge

1. Introduction

There are many criteria that an expert system must meet in order to be considered successful in a domain. An important criterion is that it be able to solve problems in its domain with a satisfactory level of expertise. In addition to this an expert system should also be able to communicate well with its users. This means not only asking for relevant information when needed but also providing explanations of its reasoning process that are acceptable to a user. Furthermore, an expert system should be easily expandable to incorporate new knowledge or correct outdated or erroneous knowledge.

DTEX is an expert system building tool for diagnostic and therapeutic problem solving designed to provide an environment for the building and using of expert systems, without involving a programmer or a knowledge engineer. Many generic expert system tools have been built, but in general their design was geared towards finding better methods for knowledge engineers to encode expertise (Balzer *et al.*, 1980, van Melle, 1979, Weiss and Kulikowski, 1979). These efforts led to the development of better environments in which to build expert systems, but ones which still required some degree of programming proficiency. The goals of DTEX are similar to the interactive transfer of expertise found in Teiresias (Davis, 1977). However, DTEX differs in several important areas. DTEX provides a problem-solving paradigm that appears to be well suited for diagnostic and therapeutic domains. DTEX differs from earlier work by maintaining a support knowledge structure that is used for explanation and knowledge acquisition.

Support knowledge consists of information which provides the rationale for an inference step. For example, an inference rule such as "if the age of the patient is less than eight then do not administer tetracycline" is supported by the knowledge that tetracycline causes tooth discoloration due to the chelation of the drug in growing bones (Clancey *et al.*, 1981, Swartout, 1981). For a given problem domain, a set of characteristics is acquired along with relationships between these and possible sets of input data. A conclusion is reached by DTEX forming hypotheses about the states of the problem characteristics. Each inference step has connotations which change the current hypotheses.

DTEX uses support knowledge to assist the expert in acquiring the knowledge base. The following have been identified as patterns in the support structure that can be used as acquisition aids by alerting the expert to possible mistakes:

- Matching hypotheses leading to contradictory conclusions
- Differing hypotheses leading to the same conclusion
- Reaching a conclusion without considering relevant problem characteristics.

Support knowledge is used to bolster the explanation of the recommendation, in addition to inference based explanations similar to those used in MYCIN (Shortliffe, 1976). Justification of inference steps is extremely important in expert system explanation. During the normal use of an expert system the ability of the system to give detailed support for its decisions increases the confidence that a user will have in its recommendations. In addition, support-knowledge-based explanations are adequate for educational use of the expert system.

2. The DTEX system

It is important to extract what is generic about diagnostic and therapeutic problem solving. The existence of a problem usually indicates that there is an anomaly in an object that must be identified (i.e. diagnosis) or if an anomaly has already been identified an action which must be taken (i.e. therapy). Problem solving is the process by which the diagnosis or the therapy is determined. Intelligent solutions can be arrived at if the expert has an internal model of three factors involved: the object, the kinds of anomalies, and the kinds of actions to be taken.

Objects can be represented by a set of properties with associated values. Properties may be test results, observed findings, or other information which can affect the decision making process.

DTEX allows domain experts in diagnostic and therapeutic fields to interactively transfer their knowledge into a knowledge base. DTEX acquires three types of knowledge from the expert:

- the different properties
- how properties affect the problem characteristics
- how conclusions can be arrived at in terms of hypothesized states of the problem characteristics.

Properties of a problem domain may fall into a hierarchy of relationships. Also, the value of one property may divide the remainder of the inference process into several distinct paths. These, along with other characteristics of problem domains allow the knowledge for the decision making process to be organized in a decision tree. This approach has been found to be appropriate for a variety of diagnostic and therapeutic domains. In DTEX, the tree formalism is adopted; a node represents a property to consider and each edge is labeled with a value or range of values of the property of its source node. Inferences are encoded as transitions between nodes in the tree. This representation is equivalent to a constrained production rule system such that no inference loops exist (Konolige, 1979).

A large quantity of knowledge is represented in addition to the tree. Each property has an associated record which includes its description, its definition, and its valid values. A property can have an associated formula. If so, when the value of the property is needed, it is calculated rather than expected as direct user input.

An example of a property in an expert system to recommend possible treatments for adolescent idiopathic scoliosis¹ would be *risser scale value*². It is a measurement of *growth potential* and its valid values are *integers between 0 and 5*.

When enough information is present to direct the inference to a leaf node, a solution is presented. The user can then request explanations of the inference process which will be supported by additional justification.

3. Support knowledge for explanation

Two types of knowledge are used by DTEX in the generation of explanations: the decisions made by the system during the inference process and support knowledge for these inference steps (Clancey, 1983). A summary of the inference process is given which includes all the information used to arrive at the solution. This does not include the support knowledge associated with the decision tree.

¹Idiopathic scoliosis is an abnormal lateral curvature of the spine without known cause.

²Risser scale value is determined from radiologic examination of the pelvis.

Once the recapitulation is available, the user has the option of asking a variety of detailed questions about any one inference step, in addition to why certain inferences were not performed. DTEX then generates English responses using a phrasal lexicon. The information in the response is extracted from the support knowledge structure associated with the corresponding inference step. Recalling the example of the scoliosis domain, if the user were to ask why the balance measurement was considered, DTEX's response would be³:

It has been determined that the risser scale value is 4. This indicates that the growth potential is minimal, therefore since the possibility of curve progression is unlikely, the amount of deviation between the midline occipito and the gluteal crease should be established in order to determine the amount of deformity present. So the balance measurement is considered.

If the risser scale value [RISSE] had been:
between 0 and 3, non-surgical techniques [NON-SURGICAL] would have been suggested.

Inferences not selected are mentioned to inform the expert of other possible solutions or considerations (see overriding the system, below). At this point the user could ask such questions as why non-surgical techniques were not considered or what would have been different if the risser scale value had been 1, both of which, in this case, would cause DTEX to respond:

If it had been determined that the risser scale value was between 0 and 3 this would have indicated that the growth potential is significant. Therefore, since the possibility of curve progression is likely non-surgical techniques would be recommended to arrest the curve from further progression.

When and if the user is satisfied the explanation module terminates. If the user disagrees with any inference step it can be overridden as long as adequate additional support knowledge is supplied.

³All examples are actual generated explanations of the DTEX system.

Additional knowledge is associated with both the properties and with inference steps. Inference support is represented by linking transitions with indications of the problem characteristics being addressed. The problem characteristics are English phrases associated with inferences by unique identifiers. This information is not used for determining the inference itself but rather it is accessed when justification is needed during consultations or educational sessions. Each node is associated with one or more general characteristics of the problem. An example from the scoliosis system is depicted in figure 1.

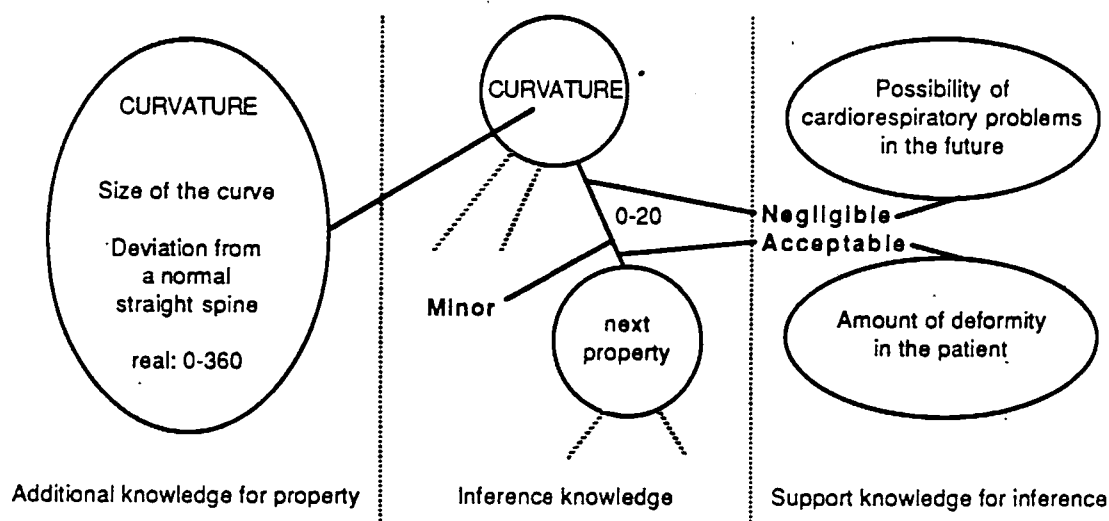


Figure 1: Knowledge structure in DTEX

A node considering the *size of the curve* would be linked to the characteristics of *the amount of deformity* and *the possibility of cardiorespiratory problems in the future*. Along each edge in the tree are modifiers on each of the characteristics being addressed at the source node as well as a modifier on the property being considered. In the above example, these are: amount of deformity is *acceptable* and possibility of cardiorespiratory problems in the future is *negligible* because the deviation from a normal straight spine is *minor*.

This support knowledge is not used during problem solving because it encodes knowledge which is often implicitly understood by a human expert. A human expert might access this knowledge for the purpose of teaching or explaining the concepts involved but not during the actual problem solving process (Clancey, 1983).

4. Intelligent assistance during knowledge acquisition

An additional advantage of using support knowledge for explanations is for detection of errors in the knowledge base. An expert running a system may not notice an incorrect inference if explanations are limited to English translations of the inference steps themselves. A rule may be taken as correct with the assumption that the underlying justification is unknown to the user analyzing the system. On the other hand, if support is given for an inference, the error will be more apparent (Weiner, 1980). DTEX takes this one step further. The process of automatic error detection in the knowledge base is extensive in the system. DTEX uses the support knowledge to detect inconsistencies in a system as it is being constructed.

DTEX acquires the support knowledge in terms of relationships between problem characteristics and properties, and those same characteristics and conclusions interactively. Inconsistencies are detected by comparing different instances of these relationships. If the support knowledge of two different inference steps is the same but the conclusions derived from these steps are different the system notifies the expert. There are several possible causes for such inconsistencies. The support knowledge could be inaccurate, or, more importantly, one of the inference steps could be in error. Upon notification the expert can choose to change the system or specify that the suggested discrepancy was incorrect. Conversely, if two inference steps lead to the same conclusion but due to differing support, DTEX suggests that the expert should verify the support given.

For example, in the scoliosis system the risser scale value is related to the possibility of curve progression. When the possibility of curve progression is likely, some method of arresting the progression is needed. Hence, if there is such an inference and the conclusion does not provide a method of arresting the progression a discrepancy has been detected. The above linking of curve progression and risser scale value is provided during the acquisition of relationships between properties and problem characteristics.

5. Implementation and current uses

DTEX is implemented entirely in OPS5 (Forgy, 1981). An expert system to recommend preliminary treatments for many cases of adolescent idiopathic scoliosis has been implemented in the DTEX system. Physicians at the Mount Sinai School of Medicine reacted very favorably to the initial demonstrations of the system. Other systems are being developed concurrently

including a student advisor. Development of several domains during the ongoing construction of the DTEX system helps to guide the research so as to create a system which is both powerful enough and easy enough to be used by domain experts in a variety of fields.

6. Conclusion

The structure of the DTEX system is ideal for incremental expansion. Features that will be added to DTEX include database access for developing expert systems for data analysis. Database analysis systems represent an important new class of expert systems as demonstrated by the ACE system produced at AT&T Bell Laboratories (Vesonder *et al.*, 1983). A graphics interface will improve all facets of the interaction between DTEX and its users from problem solving to knowledge acquisition.

Support knowledge is necessary for adequate explanations. It provides solid justification in response to a user's questioning of a given inference. In addition, it enriches a system's power as a teaching tool. The encoded knowledge reflects a human model in that problem solving ability is compiled into concise inference rules whereas support knowledge is accessed only when needed for further justification.

The acquisition of knowledge bases has been the major bottleneck in the development of expert systems. Automating this process requires that the system guide the expert, offering as much assistance as possible. DTEX allows for this by the incorporation of the support knowledge structure. As the system grows, all the previous knowledge is analyzed for consistency in its support. In the future, it will be possible to transfer the support information from one system to a new one in a similar domain thus helping to guide its construction.

References

- Balzer R., L. Erman, P. London, and C. Williams. *Hearsay-III: A Domain-independent Framework for Expert Systems*, pages 108-110. AAAI, Stanford University, August, 1980.
- Clancey W. J. The Epistemology of a Rule-based Expert System - a Framework for Explanation. *Artificial Intelligence*, 1983, 20, 215-51.
- Clancey W. J., and R. Letsinger. *NEOMYCIN: Reconfiguring a Rule-based Expert System for Application to Teaching*, pages 829-36. IJCAI, Vancouver, B. C., August, 1981.
- Davis R. *Interactive Transfer of Expertise: Acquisition of New Inference Rules*, pages 321-8. IJCAI, August, 1977.
- Forgy C. L. *OPSS User's Manual*. Technical Report CMU-CS-81-135, Department of Computer Science, Carnegie-Mellon University, July 1981.
- Konolige K. *An Inference Net Compiler for the PROSPECTOR Rule-based Consultation System*, pages 487-489. IJCAI, Tokyo, Japan, August, 1979.
- Shortliffe E. H. *Computer-based Medical Consultations: MYCIN*. New York: American Elsevier 1976.
- Swartout W. R. *Explaining and Justifying Expert Consulting Programs*, pages 815-823. IJCAI, Vancouver, B. C., August, 1981.
- van Melle W. *A Domain-independent Production-rule System for Consultation Programs*, pages 923-925. IJCAI, Tokyo, Japan, August, 1979.
- Vesonder, G. T., S. J. Stolfo, J. Zalinski, F. Miller, and D. Copp. *ACE: An Expert System for Telephone Cable Maintenance*. IJCAI, Karlsruhe, West Germany, August, 1983.
- Weiner J. L. BLAH, A System Which Explains its Reasoning. *Artificial Intelligence*, 1980, 15, 19-48.
- Weiss S. M., and C. A. Kulikowski. *EXPERT: A System for Developing Consultation Models*, pages 942-947. IJCAI, Tokyo, Japan, August, 1979.