Developing Expert Systems From
Examples and Explanations

Joseph C. Ross

AI TR87-63 August, 1987
DEVELOPING EXPERT SYSTEMS FROM EXAMPLES AND EXPLANATIONS

by

JOSEPH CLINTON ROSS, B.S.

THESIS
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCES

THE UNIVERSITY OF TEXAS AT AUSTIN
August, 1987
Acknowledgments

I wish to thank my advisor, Dr. Bruce Porter for his careful guidance and continuous support, throughout the course of this research. I am also grateful to Dr. Craig Wier for his willingness to serve as domain expert and for his constructive comments as a member of my supervising committee. This thesis was also greatly influenced by many thought-provoking discussions with Ray Bareiss and Ken Murray.

Joseph Clinton Ross

The University of Texas at Austin
August, 1987
# Table of Contents

Acknowledgments ................................. ii

Table of Contents ................................ iii

List of Figures ................................... vi

1. Introduction .................................... 1

2. Previous work .................................. 4
   2.1 Three representative systems ............. 5
       2.1.1 MYCIN .................................. 5
       2.1.2 CENTAUR ................................. 8
       2.1.3 ABEL .................................... 11
   2.2 Lessons learned from this work ........... 13
   2.3 Difficulties with current techniques .... 15
   2.4 Weaknesses of current systems .......... 15

3. Discarding knowledge causes problems ...... 17
   3.1 The ways in which knowledge is discarded 17
   3.2 Knowledge not represented in typical expert systems 19
   3.3 Why knowledge is not retained .......... 20
   3.4 How discarding knowledge makes construction more difficult 21
   3.5 How discarding knowledge results in weaknesses 22
4. The proposed approach

4.1 Description of the representation .................................. 25
  4.1.1 The elements of this representation .......................... 26
  4.1.2 Why this representation is justified ........................ 28
4.2 How this representation supports classification ............... 29
  4.2.1 Examples are good models for classification ............... 29
  4.2.2 Background knowledge permits sound indexing and
        generalization ............................................... 29
  4.2.3 A "hypothesize and test" control strategy ................. 31
4.3 Why knowledge-base construction is easier .................... 31
  4.3.1 Focused knowledge acquisition ............................. 32
  4.3.2 Integration of new knowledge .............................. 33
  4.3.3 Knowledge-base refinement ................................ 34
4.4 How this knowledge reduces the typical weaknesses .......... 35
  4.4.1 Explanation capability ................................... 36
  4.4.2 Example-generation capability ............................. 36
  4.4.3 Robustness in classification ............................... 36

5. An implementation .................................................. 38
5.1 Knowledge acquisition ............................................. 39
  5.1.1 Covering the domain uniformly and adequately ........... 39
  5.1.2 Knowledge acquired for each case .......................... 40
  5.1.3 Two interviewing modes .................................... 42
5.2 Knowledge representation ........................................... 49
  5.2.1 Concept representation .................................... 49
  5.2.2 Constructing example structures ............................ 53
5.2.3 Structuring concepts into a network of relationships 55
5.3 Knowledge entry ........................................... 60
  5.3.1 Entering background knowledge ...................... 61
  5.3.2 Adding examples ...................................... 63
5.4 Knowledge refinement ..................................... 67
  5.4.1 Correcting problems .................................. 67
  5.4.2 Improving performance ................................ 68
5.5 Knowledge use .............................................. 71
  5.5.1 Explanation and inference ............................. 71
  5.5.2 Classification .......................................... 74

6. Conclusion .................................................. 81
  6.1 The general theme ........................................ 82
  6.2 The representation and techniques ....................... 83
  6.3 The implementation ....................................... 84

BIBLIOGRAPHY .................................................. 89
List of Figures

4.1 Two examples of mugs ........................................ 26
4.2 Some knowledge about mugs ................................. 27

5.1 List of diagnostic categories ............................... 39
5.2 Data presented to the expert for Case 13-1 ............ 43
5.3 Transcript of the “thinking aloud” phase for Case 13-1 . 44
5.4 Transcript of the active questioning phase for Case 13-1 . 47
5.5 Extraction of concept names from text ................... 50
5.6 Extraction of concept names from text ................... 51
5.7 Concept names identified for Case 13-1 .................. 52
5.8 Level of abstraction for features .......................... 53
5.9 Example structure derived for Case 13-1 ............... 54
5.10 Relationships derived for Case 13-1 ..................... 58
5.11 Relationships from Case 13-1 after addition of qualifiers .. 62
5.12 Some explanations relevant to the example from Case 13-1 . 65
5.13 Reminders and importances for the example from Case 13-1 66
5.14 Specializations of fixation(ossicular-chain) ............ 69
5.15 Effects of moving Example 13-1 to fixation(malleus) .... 69
5.16 Computation of inference strength from explanations .... 73
5.17 Reminding strengths and specialization links for Example 13-1

5.18 Matching a new example to Example 13-1
Chapter 1

Introduction

During the past two decades many expert systems have been developed which perform classification very well in narrow domains. This thesis examines this work, and describes a technique for developing such systems based on the following ideas:

1. Although expert-systems research has been successful, (a) the construction of these systems has been a very difficult process, and (b) even when good classification performance has been achieved, the resulting knowledge bases have not been broadly applicable.

2. An important reason for these problems has been the failure to retain all information available from an expert during knowledge acquisition.

3. An expert-system development technique which represents all available knowledge explicitly (a) makes construction much less difficult, and (b) results in a knowledge base which can support good levels of performance at a variety of expert tasks.

Chapter 2 examines previous expert-systems research in the context of three systems whose primary task is medical diagnosis. It summarizes some useful ideas which have emerged from this work, but also identifies the problems mentioned in the first of the above points.
Chapter 3 categorizes the knowledge available from an expert as consisting primarily of (1) examples of how the expert performs classification and (2) explanations of the relationships which exist between domain concepts. This chapter then describes how a large portion of the knowledge is usually discarded, and demonstrates how this knowledge loss contributes to the problems identified in the previous chapter.

Chapter 4 describes a technique for representing classification examples and explanations explicitly. This approach greatly simplifies knowledge-base construction, since very little transformation must be performed by the knowledge engineer. Furthermore, examples provide a way to systematically acquire and structure the knowledge.

This representation is effective, because the two types of knowledge complement each other very well in supporting a variety of expert tasks. In classification, examples serve as good models, while the background knowledge from explanations permits them to be used efficiently, and provides coverage in situations not explicitly covered by examples. In the explanation task, the background knowledge permits good justification for inferences, while the examples provide a concrete way to focus the explanation of a classification.

Chapter 5 describes a specific application of this technique to the domain of clinical audiology. The problems, decisions and assumptions involved are discussed in the context of the acquisition, representation, integration and refinement phases of expert system construction. Also, some algorithms implemented to make use of the knowledge for classification and explanation are covered.
In the conclusion, this research is discussed at three levels of abstraction: (1) an underlying theme, (2) the proposed representation and techniques, and (3) the way this proposal was implemented. Limitations to the approach and directions for future research are also identified.

The research described in this thesis was done as a part of the Protos project [PB86,BPW87]. Protos is a learning apprentice system which automatically develops into an expert system by interacting directly with a human expert. Many of the ideas on knowledge acquisition, representation and utilization described in this thesis originated with Protos and were implemented and justified in the context of manual expert system construction.
Chapter 2

Previous work

The history of expert systems has been one of the bright spots of artificial-intelligence research. It has been one of the few areas in which the high expectations placed on the field were matched by successful and useful systems. These systems have matched the performance of human experts in domains such as interpretation of mass spectra [BF78], interpretation of geological data [DGH79], computer system configuration [McD82], internal medicine [Pop82] and pulmonary-disease diagnosis [KFM*78].

This chapter examines this work and summarizes some useful ideas which have emerged from it. However, the following problems are also identified:

1. The construction of these systems has typically been far from straightforward. It has involved much trial and error, and has consisted of a painstaking process of handcrafting a knowledge base tailored to a specific task.

2. Even when good classification performance has been achieved, the knowledge bases, which were so laboriously built, could not support other common expert tasks. Moreover, when presented with unusual, unforeseen classification problems, they have often failed miserably.
These problems are a motivation for the approach proposed in this thesis, which addresses them directly, while retaining the useful ideas from previous work.

2.1 Three representative systems

This section reviews three expert systems whose primary task is medical diagnosis. Although there are systems like R1 [McD82], which perform other types of tasks, most systems are designed to primarily perform classification. Furthermore, although there are many noteworthy systems which perform in other domains, [BF78, DGH79, EHLR80], medical expert systems are representative of most of the approaches which have been used. The following examples are discussed because their behavior and structure have been well described, and because they span the spectrum of typical medical expert-system techniques.

2.1.1 MYCIN

The MYCIN [Sho76] system was designed as a consultation program for the diagnosis and treatment of infectious blood diseases. Although it performed both diagnosis and treatment specification, these tasks were performed by essentially separate programs, so that it is possible to discuss only the portion which dealt with diagnosis.

The domain knowledge in MYCIN is represented strictly as a data base of rules, which is interpreted by a general inference procedure. Thus MYCIN was one of the first systems to demonstrate the usefulness of representing domain-specific knowledge differently and separately from general knowledge.
In MYCIN, the modularity and uniformity of rules were seen as advantages. Modularity permitted rules to be treated as independent chunks of knowledge, without the problem of significant unforeseen interactions between rules. Uniformity permitted a simple inference mechanism to be used. Both characteristics were extremely important in permitting the incremental construction of the knowledge base.

The rules permitted the knowledge base builders to deal with small incremental units. It was thought that trying to represent the knowledge using larger, more complex knowledge structures would have required a more complete view of the knowledge than was initially available [BS84].

The primary inference mechanism is a very simple and general, goal-directed, backward-chaining procedure. The unstructured behavior of the system was seen as a way of dealing opportunistically with the ill-structured nature of the diagnosis problem. It was later found that the lack of ordering of rules and lack of structure in the rule base resulted in some bad characteristics, especially in the ordering of questions. This was partially corrected by always asking certain general questions at the start of a consultation and by the addition of meta-rules which contained knowledge of control strategies [BS84].

The concept of certainty factors was incorporated into MYCIN, because the data available to MYCIN did not permit inference with complete certainty. Certainty factors were combined in a way which treated them approximately as conditional probabilities. The certainty of the conclusions was affected by the certainty of the antecedent and the certainty of the rule.

To summarize, the MYCIN research demonstrated the usefulness of the following ideas:
• Separating the general knowledge for performing inference from the domain-specific knowledge, facilitates modification and inspection of the knowledge base.

• A modular representation for knowledge in which the units could be thought of as being approximately independent, facilitates incremental addition and modification.

• A uniform representation for knowledge permits it to be used and manipulated by simple procedures.

• An unstructured representation and loose control mechanism permits relevant knowledge to be applied to ill-defined problems opportunistically.

However, some problems with the MYCIN approach are:

• Different types of knowledge are represented in the same way and it is impossible to distinguish among them.

• Rules represent only estimates of degree of co-occurrence between concepts without the relationships on which these estimates are based [Cla83,CDD*85].

• The knowledge of overall patterns of data consistent with real cases is not represented explicitly.

• The inference process remains unconstrained even when there is sufficient information to form a good hypothesis.
2.1.2 CENTAUR

CENTAUR \cite{Aik83} is a system which interprets pulmonary function tests. It was constructed by converting the knowledge contained in the rule-based system, PUFF \cite{KFM*78}, into a more structured representation.

In CENTAUR, frames are used to represent prototypical patterns of data associated with particular categories. These prototypes are linked to prototypes for related concepts in a concept hierarchy. Rules are used in conjunction with the frames to select prototypes relevant to the current data, and to evaluate to what degree individual features of the prototype are matched by the data.

Frames make knowledge more understandable by making expected patterns of data explicit. This structure is actually present in the knowledge, and must also be captured by rule-based systems through interactions between rules. However, these connections between rules are not visible, making the knowledge difficult to understand and modify.

In CENTAUR, rules are attached to slots of frames, so that it is possible to differentiate between different types of rules. In particular, the distinction between inference and control rules is made explicit. This makes knowledge easier to inspect and manipulate in CENTAUR, than it is in systems which are strictly rule-based.

The explicit representation of patterns of data is useful in handling the multiple-diagnosis problem. If data are unaccounted for after a successful match to a category, then overlapping categories are indicated. On the other hand, if two categories each independently account for most of the data, then this suggests uncertainty between plausible diagnoses. In rule-based systems
like MYCIN, there is no indication whether the set of diagnoses suggested by
the system is conjunctive, disjunctive or a combination of the two.

CENTAUR's representation permits a *hypothesize-and-test* control
strategy, in which a data-directed phase is used to select a prototype, which
then acts as a context for the consultation. This strategy is useful for con-
straining search in ill-structured problems like medical diagnosis [Pop82]. Fur-
thermore, after a prototype has been selected, the order in which questions
are asked is focused by the prototype context, and is more efficient than the
continuously undirected questioning of MYCIN or PUFF.

Prototypical category descriptions in CENTAUR make it possible
to distinguish between the *importance* and *inference strength* components of
the certainty factors. Importance is associated with the necessity that a
particular slot of a frame be matched, while inference strength is used for
selecting the relevant frame or evaluating the degree to which particular slots
are matched by the data. The notion of importance of features to prototypes
permits a more reasonable representation of "fuzzy" concepts than is possible
with rules.

One drawback of the CENTAUR approach is that directly generat-
ing complete descriptions of categories is probably considerably harder, for
the expert, than generating rules. Thus, the structured representation makes
it more difficult to incrementally acquire knowledge. However, CENTAUR
did not address the knowledge-acquisition issue, since its knowledge was ob-
tained by converting the knowledge from another expert system.

Another problem with the CENTAUR approach is that certain
concepts are difficult or impossible to represent using single prototypes.
There is psychological evidence that natural categories are often represented
using multiple exemplars, rather than a single characteristic description [SM81,BP87]. This is especially true during initial concept learning, which corresponds to the incremental knowledge base construction of expert systems. This realization was an important motivation behind the research described in this thesis.

To summarize, some useful ideas demonstrated by the CENTAUR approach are:

- Representing prototypical patterns explicitly gives the knowledge more structure and makes it more understandable.
- Using different representations for different types of knowledge, makes knowledge of the differences visible and usable.
- The hypothesize-and-test control strategy is useful in constraining the inference process and providing structure to the consultation.

Some problems with the CENTAUR approach are:

- For some concepts, it may not be possible to identify a single prototype which is the generalization of all its examples.
- The large grain size of prototypes and overall structure imposed by their hierarchical arrangement makes incremental acquisition of knowledge difficult.
- As in rule-based systems, the relationships between concepts which make inferences valid are not represented explicitly.

The PIP [PGKS84] and INTERNIST [Pop82] systems are similar to CENTAUR in that they have explicitly represented knowledge of prototypical
disease patterns, and that they use an explicit hypothesize-and-test control strategy. Although there is variation in how they represent and use knowledge, the depth of the knowledge is approximately the same. In PIP and INTERNIST, relationships between concepts are somewhat more explicit, but this additional knowledge is not really used.

2.1.3 ABEL

The ABEL system [PSS81] was designed to provide expert consultation for electrolyte and acid-base disturbances. In this system, knowledge of causal relationships, temporal patterns and aggregate disease categories was explicitly represented and utilized. The knowledge in ABEL is represented at various levels of detail, from low-level physical knowledge about diseases to high-level syndromic knowledge.

Each level of the knowledge base is a semantic network of relations between nodes which represent findings and disease states. The following different types of relationships are represented explicitly:

- causal
- constituent of
- abstraction/elaboration

Nodes and links are represented using frame-like structures which contain much more information than just the name. This extra information represents even deeper knowledge, which can be useful to permit more intelligent performance by the procedures which use the knowledge base.

Sets of nodes at lower levels are aggregated into single, more abstract nodes at higher levels, or conversely, abstract composite nodes can be elabo-
rated into networks of nodes at lower levels. Causal and constituent-of links connect nodes at the same levels. Aggregation and elaboration also apply to links, with chains of links being aggregated into single links at more abstract levels, and with single, composite links being elaborated into multi-link paths at more detailed levels.

ABEL performs diagnosis by constructing a patient-specific model for each case. This is done by initially matching a consistent pattern of concepts and links at the lowest level to the data. The complete model at all levels of abstraction is then filled out by performing aggregation and elaboration operations on concepts and causal pathways. Knowledge at each level of abstraction is used to propagate the information to complete the model and to enforce consistency of the model at all levels.

In ABEL, patterns of data are not represented explicitly, but the useful constraints imposed by them have been built into the knowledge base, so that consistency is enforced. Although this technique might be effective in this case, building such a knowledge base incrementally is far from straightforward. Furthermore, development of this knowledge base was possible largely because of the unusually thorough understanding of causal and structural relationships which exists in this domain.

Some useful ideas demonstrated by the ABEL approach are:

- Explicit representation of different types of relationships makes the knowledge more understandable, and permits more intelligent performance by the inference procedure.
- Ability to perform component decomposition and summation for composite quantities (e.g., concentrations) permits more intelligent inference and decomposition of overlapping data.
• Representation of knowledge at various levels of abstraction permits knowledge at the different levels to be used to soundly constrain the inference process.

Some disadvantages of the approach used in ABEL are:

• The constraints necessary to define a consistent patient model from the general information in the knowledge base would be difficult to learn. This knowledge must be a generalization of case data, but the process of handcrafting such knowledge is difficult. As in CENTAUR, it is not clear how the prototypical patterns of data which must be encoded can be recognized except in the context of many examples.

• The representation and its procedures rely heavily on a rather complete understanding of the disease and patient mechanisms and processes. In many expert domains, such knowledge does not yet exist, and reasoning must be more empirical in nature, relying on probabilistic knowledge and knowledge of previous examples, as well as any categorical knowledge which may be available.

The CASNET system [WCA78] is similar to ABEL in that it is a handcrafted knowledge base containing explicit knowledge of causal links and multiple levels of abstraction. However, the knowledge in ABEL is somewhat more detailed, and the causality in ABEL is primarily mechanistic, while that of CASNET deals primarily with disease etiology.

2.2 Lessons learned from this work

Several important lessons can be learned from the work described in this section:
• Large quantities of knowledge are necessary to achieve expert performance. Weak search is useless for exploring the large search spaces associated with many human-reasoning tasks. The intractability of these problems can only be controlled by useful heuristics, which make use of natural constraints. Expert domains are typically very complicated, and the constraints are numerous and powerful, but very specific. Thus, capturing the heuristics in these domains requires that large amounts of specific information be explicitly retained.

• Knowledge base construction is inherently an incremental process. Expert knowledge is indexed in ways to optimize task performance, but not necessarily in ways that facilitate top-down, exhaustive retrieval. Knowledge acquisition in the context of a specific case [Dav79] is a useful way to subdivide this task. The modularity and small grain size of rules permits them to be incrementally acquired and integrated into the knowledge base.

• Knowledge must be represented in a way which is understandable and manipulable to be useful. Even with a representation well suited to incremental knowledge acquisition, it is necessary to be able to inspect and refine the information in the knowledge base. Modularity is important so that a piece of knowledge can be understood and modified without unforeseen side effects. Uniformity of representation for knowledge of the same type permits it to be easily understood and simply manipulated. Explicit differentiation of knowledge of different types permits it to be understood and processed more specifically.

• A "hypothesize and test" control strategy can be useful. An opportunistic data-driven phase is initially necessary to use all available
information in ill-structured problems. However, focusing on a hypothesis as soon as possible is effective in making detailed low-level reasoning tractable, and directing additional information gathering.

These ideas are fundamental to expert system construction, and are retained in the techniques and knowledge representation schemes which are the focus of this thesis.

2.3 Difficulties with current techniques

Although systems which achieve expert performance levels in narrow areas have been built, constructing these systems has been difficult. Specifically:

- Knowledge acquisition is laborious and unstructured, involving much time and effort on the part of both experts and knowledge engineers.

- Knowledge integration is not always straightforward, since the relationship and interaction between new knowledge and old knowledge is often not clear.

- Knowledge refinement has been a necessary longterm process requiring backtracking. This has consequently increased the amount of knowledge acquisition and integration effort required.

2.4 Weaknesses of current systems

The present research is concerned with systems whose primary task is identification or classification. This has been the primary task of most expert systems, and is an important activity performed by all expert systems
[Cla85]. Current expert-system knowledge bases have been developed and refined so as to support a high level of performance in the classification task in the target domain. However, when these systems are compared with human experts, whose primary task is also classification, some weaknesses become evident:

- These systems have typically not been able to effectively justify their behavior or explain their knowledge.
- These systems have typically not been able to perform example generation.
- These systems have not been robust, failing miserably when trying to classify unusual objects which they had not been explicitly built to handle.

The following chapter will describe how these construction difficulties and knowledge-base weaknesses are related problems. They are largely the result of a failure to retain information available during knowledge acquisition.
Chapter 3

Discarding knowledge causes problems

A significant reason for the construction difficulties and knowledge base weaknesses of expert systems identified in the previous chapter, is the failure to retain certain types of relevant knowledge during construction of the knowledge base. This chapter will describe this information loss, why it occurs, and how it contributes to the above problems.

3.1 The ways in which knowledge is discarded

To better understand and categorize the sorts of knowledge not retained during knowledge acquisition, it is useful to first identify and categorize the types of knowledge available from the expert. Knowledge can be characterized along many dimensions, but two dimensions which are especially relevant to expert system construction are (1) the ways concepts are described, and (2) the tasks the expert performs when generating the information.

A concept can be viewed as a single, significant idea, called the intension of the concept, or as the set of objects to which the concept applies, called its extension [Woo75]. The intension of a concept can only be described in terms of the relationships which exist between it and other concepts. These relationships, together, describe the idea represented by the concept and the reason for its existence. The extension of a concept consists of a description of
the examples of the concept. These examples are represented in the instance language [Mic83], which consists of low-level, easily observed or computed features.

For the types of expertise on which this research focuses, the primary tasks performed by the expert are classification and explanation. During classification, the expert categorizes examples described in the instance language. Typically, this requires shifting the representation of the example to the generalization language [Mic83, FD86]. The generalization language consists of criterial, functional features, which can be inferred from features in the instance language.

Expertise is essentially the ability to perform this representational shift efficiently and accurately. However, in most domains, a large part of expertise also involves being able to explain why an example was classified in a particular way. Thus, the expert can usually generate explanations for the inferences he makes in performing the mapping. These explanations are made by justifying the inferences in terms of underlying relationships between concepts.

Therefore, the information available from an expert can be characterized as consisting of (1) examples of categories, and (2) relationships between concepts. From this perspective, the knowledge loss which takes place in typical expert systems is as follows:

- The data associated with individual examples is generalized so that only a small part of the information is retained as a unit.
- The recognized relationships between concepts are reduced to simple implications, annotated only with some guess at the general validity of the implication.
3.2 Knowledge not represented in typical expert systems

In all three expert systems described in Chapter 2 one or both of these types of knowledge loss occur during knowledge acquisition. In MYCIN, each rule can be viewed as a generalized classification example, in which only a small portion of data is retained in the antecedent clauses of the rule. A rule can also be viewed as a relationship between the antecedent concepts and the conclusion concept, in which only the certainty of the inference associated with the relationship is retained. Thus, in MYCIN, both types of knowledge loss occur, to the point that the distinction between the two types of knowledge is completely blurred.

In CENTAUR, frames represent prototypical data patterns which describe concepts, with only features associated with a large number of examples of the concept being represented as slots. As such, they are more general than individual examples, with the non-generalizable information being absent. CENTAUR retains some knowledge of relationships between concepts in slots which link frames together. However, the language for these relationships is not very rich, and could not capture much of the knowledge available in explanations.

ABEL retains knowledge of several types of relationships, including causal ones, but does not represent examples explicitly. It contains information of general prototypical-category models, similar to that in CENTAUR’s frames, but this knowledge is not significantly differentiated from the knowledge of relationships between concepts. Moreover, only a small subset of the different types of relationships existing between concepts is represented.
3.3 Why knowledge is not retained

Knowledge is discarded during knowledge acquisition for several reasons. Some of these reasons are sound, but in most cases the resulting problems outweigh the possible advantages, and in all cases, postponing generalization is the more prudent and cost-effective alternative.

The most common and significant reason for knowledge loss during expert-system construction, is that it is performed with the goal of supporting only the classification task. Any knowledge which is not seen as necessary or useful to this task is discarded, resulting in a knowledge base lacking the information needed to support other expert abilities. Even within the context of classification, many potential features of an example are discarded, because they don't appear to be generalizable to other examples. The goal has been to make knowledge as widely applicable as possible, so that less knowledge need be acquired and retained.

A related, but slightly different source of knowledge loss occurs when the chosen representation is not well suited to capture all forms of available knowledge. Thus, much information may be lost in forcing it into an overly restrictive mold. This situation occurs with rules, which provide no natural place for the relationships which justify their existence [CDD*85]. Contradiction, subsumption and duplication of rules are considered errors which should be resolved [SSS84]. However, if the relationships which underlie the rules are retained, conflicting rules may be seen as representing different relationships between the same concepts.

Retaining only useful knowledge can be economical in allowing classification of unseen instances with only small amounts of information. However, there is evidence that, although an expert is able to focus on the relevant
portions of an example in performing classification, the generalization implicit in this process cannot be reliably communicated. Much of the knowledge which the expert uses in performing classification is "compiled" and indexed in ways that optimize performance, but not available for explanation. An expert who can perform classification well, cannot necessarily explain exactly what knowledge he uses to perform this task [NW77,ES80]. Thus, discarding knowledge during knowledge acquisition is an example of "false economy", since it is always possible to perform the generalization later, but is is not easy to recover from errors due to permanent overgeneralization.

3.4 How discarding knowledge makes construction more difficult

The process of developing a knowledge base for an expert system is lengthy and difficult. The lengthiness is a consequence of the enormous volume of information to be acquired and represented. Much of the difficulty is a consequence of the attempt to discard irrelevant information during knowledge transfer. Attempting to perform this generalization is a difficult task in itself, but the problem is compounded, because this difficulty makes the process prone to over-generalization errors.

If there is a large difference between the way knowledge is available from an expert and the way it is represented in the system, then a great deal of transformation is required. This is a difficult task, and cannot be performed reliably by either the expert or the knowledge engineer. It would obviously be much easier for the expert to directly generate classified examples and explanations while performing his task and for the knowledge engineer to simply record it. Although in practice some transformation will always be
needed, knowledge acquisition will be much less difficult if it is minimized.

When transformation is so difficult that it results in frequent or serious problems in the knowledge base, the knowledge base must be modified to correct these errors. This is an inherently difficult task, because the error may be hard to locate and correct reliably. Thus, the process of knowledge-base refinement becomes lengthy, and involves much backtracking.

Problems in a knowledge base are typically due to overgeneralized knowledge. This situation arises when knowledge is misapplied and used to incorrectly classify an object in an unforeseen situation. To correct these errors, it is necessary to restrict the applicability of the knowledge. However, there may be many ways to this, and only one way would restore the information which could have been simply retained during knowledge acquisition.

Overly specialized knowledge simply results in its non-use, and can be corrected in the same way as missing knowledge, by the natural process of adding information. Furthermore, if enough background knowledge is present in the knowledge base, it becomes possible for the inference mechanism to generalize “dynamically” during classification. In situations where permanent generalization is unavoidable, there is no harm in waiting until it is necessary.

3.5 How discarding knowledge results in weaknesses

The weaknesses of the systems described in Section 2.4 are directly related to corresponding deficiencies in the knowledge base:

- The inability to form good explanations is largely caused by insufficient knowledge of the relationships which underlie and justify inferences.
• The inability to generate actual examples of concepts is a direct consequence of missing knowledge of realizable data patterns and consistent concept groupings.

• The inability to perform classification in unusual, unforeseen situations is a symptom of the absence of deep knowledge, which the system could resort to in these situations.

These weaknesses may not at first seem to be too serious. It may appear that so long as a system performs the primary task of diagnosis well enough, any other abilities are superfluous. However, since these abilities are commonly demonstrated by an expert, they represent deficiencies in the knowledge base with respect to human performance. These abilities are directly related to the expert’s ability to reason about and evaluate his knowledge, and are very useful during learning. Since learning in the expert is analogous to the construction of the knowledge base in an expert system, these weaknesses are directly related to the previously mentioned difficulties in this task.

Typically, the knowledge base which was so laboriously prepared is suited for only one use. Most noticeably, it cannot be used to facilitate further acquisition, integration and refinement of knowledge. Much of the knowledge available during knowledge acquisition was almost immediately discarded, because its usefulness to the classification task was not apparent. Knowledge is a powerful resource, as is evident by the cost of acquiring and representing it, and its usefulness to the tasks to which it has been applied. This power should be retained as much as possible, and the focus of the expert system building task should be shifted from that of building a system which performs well, to that of building a complete, sound, and useful knowledge base.
The next chapter will propose a representation aimed specifically at retaining the types of knowledge available from an expert, in their original form. It will then demonstrate how this representation is useful, and addresses the types of problems identified in Chapter 2.
Chapter 4

The proposed approach

A major theme of this thesis is that expert-system construction is less difficult, and the resulting knowledge base is more useful, if all available knowledge is retained during knowledge acquisition. This chapter describes a representation scheme in which classified examples and explanations of the classifications are retained explicitly. This representation is proposed as an effective way to capture the knowledge available from a human expert without requiring significant transformation. Furthermore, the two types of knowledge in this representation are shown to both be necessary and complementary in supporting a variety of expert tasks.

4.1 Description of the representation

The proposed representation for expert knowledge consists of classified examples and explanations of the classifications. Although these are distinct types of knowledge, which are represented differently, they are tightly interconnected so that it is easy to apply both types of knowledge simultaneously.
4.1.1 The elements of this representation

In this scheme, the example knowledge is represented by linking features and categories to example nodes, while the general background knowledge is represented using concepts and the relationships between them. The two bodies of knowledge overlap, because the features and categories of the examples are concepts in the knowledge base of relationships. Figure 4.1 shows two examples of the category COFFEE-MUG, while Figure 4.2 shows some explanations relating the features of these examples to the category.

The elements in this representation scheme are defined as follows:

Concepts are ideas which can be viewed as assertions about objects in the current context. All the rectangular boxes in Figures 4.1 and 4.2 represent concepts.
Figure 4.2: Some knowledge about mugs

**Relationships** are ideas which relate concepts with one another. The intension of a concept is defined by its relationships to other concepts. The labeled arrows in Figure 4.2 represent relationships.

**Categories** represent the extension of concepts. They are sets of objects which are equivalent with respect to the associated concepts. The examples of a category are members of this set. The concept COFFEE-MUG is a category.

**Features** are the “observable” concepts associated with an example. The concepts which describe MUG-1 and MUG-2 in Figure 4.1 are features.
Examples group concepts which are consistent in the current context. They can be viewed as instances of input/output behavior, with observable features being the input, and a category as the output. In Figure 4.1, MUG-1 and MUG-2 are examples of coffee mugs.

4.1.2 Why this representation is justified

This knowledge representation scheme may seem rather arbitrary, but it is relevant to expert system construction. The classified examples are the most reliable task-specific information concerning the expert’s ability to perform classification. On the other hand, the relationships represent the expert’s general knowledge about his domain, which is not task-specific.

There is a great deal of evidence that an expert’s classification procedures are not verbalizable [NW77]. They are effectively compiled and indexed in ways that optimize performance, but cannot be communicated directly to the knowledge engineer. While the expert has the ability to reason analytically, this is done only when compiled knowledge is insufficient, such as when expectations are violated, or explanations must be verbalized.

If an expert system is to truly model the abilities of an expert, then both the general domain knowledge and the compiled classification knowledge should be captured. The verbalizable, general knowledge can be acquired directly, by asking the expert for explanations. However, the compiled knowledge can be acquired only indirectly, through examples of task performance.

As shown in Figure 4.2, a loose network of relationships between concepts provides a way to flexibly and easily represent background knowledge. On the other hand, classified examples such as those in Figure 4.1 provide a way to accurately (but incompletely) describe a classification procedure.
4.2 How this representation supports classification

4.2.1 Examples are good models for classification

Using this representation, examples serve as models for identification, since they can be viewed as completely accurate rules. One can be certain that in the situation captured by the example, the expert would (and did) provide the classification which is the example’s category.

Examples are good models even if they cannot be matched exactly, because it is easier to match a new object to a similar example of a concept, than to some general description of the concept itself. For instance, it is easier to explain the three differences between MUG-1 and MUG-2, than it is to explain why all six features of MUG-2 can be used to classify it as a COFFEE-MUG.

There are, however, some obvious problems involved with simply storing examples and attempting to match them. Although they are a sound form of knowledge, they can never in practice provide complete coverage. Furthermore, searching through all examples would be too time consuming. Therefore, it is necessary to be able to reliably generalize examples to cover unseen instances and to index the knowledge to permit rapid retrieval of relevant examples.

4.2.2 Background knowledge permits sound indexing and generalization

Much of the knowledge required to soundly index and generalize an example is present in the background knowledge, which surrounds its features and category.
The initial process by which a set of relevant examples is chosen for matching to a new object is analogous to the reminding process in humans. To be quickly reminded of relevant sets of examples, the program must know the degree to which features of the object should remind it of various example-containing categories. This information can be derived from background knowledge. For instance, in Figure 4.2, HAS-HANDLE is connected to COFFEE-MUG by a suggests relation and thus has reminding power for it and its examples. However, BROWN-COLORED is connected to COFFEE-MUG only through a spurious relation, and has no reminding power.

To match a new object to an incomplete data base of stored examples, the system must have knowledge of the importance of features of the example. It must know which differences between the new object and an example are significant. As with remindings, the knowledge necessary to generalize a model can be derived from the background knowledge. For instance, HAS-CLOSED-BOTTOM is important to examples of COFFEE-MUG, because it is required by HOLD-LIQUID, which is a generalization of HOLD-HOT-LIQUID, which is a function of COFFEE-MUG. In constructing a match between MUG-1 and MUG-2, the system must explain why the three features which differ between them are either equivalent or unimportant. It can explain that HAS-HANDLE matches INSULATED, because they both are sufficient for INSULATE-HAND. It can similarly match MADE-OF- PORCELAIN to MADE-OF-PLASTIC, and it can know that since color is spurious, it is not necessary to match WHITE-COLORED and BROWN-COLORED.

Thus, background knowledge of the domain is necessary to use the examples effectively. The background knowledge is linked to the example data base to permit sound retrieval and generalization of examples so that
they can be used as models.

4.2.3 A "hypothesize and test" control strategy

The *hypothesize-and-test* control strategy has been demonstrated to be useful method for handling ill-structured problems [Pop82,BF78,Aik83]. This technique consists of a necessarily unconstrained, data-driven phase which results in hypothesis selection. The selected hypothesis then imposes constraints on the behavior of the system for the subsequent model-driven phase.

The identification procedure in our approach is an example of this control strategy, and consists of a *reminding phase* and a *confirmation phase*. These two phases correspond to the processes of model selection, and model verification. The unstructured background knowledge of concepts and relationships permits an opportunistic control strategy for the ill-defined reminding phase. This process is data-driven and combines positive evidence to select a model. Examples, on the other hand, provide a mechanism for structuring the general classification problem in a way that the unstructured background knowledge cannot. They focus problem solving by serving as models for the confirmation phase, which is model-driven and combines negative information to invalidate an incorrect model.

4.3 Why knowledge-base construction is easier

The explicit representation of classified examples and explanations facilitates all of the difficult expert system construction tasks described previously. Less transformation is involved in simply adding examples and relationships to the system, than if this knowledge must be converted to rules,
frames or causal models. Since this task is easier and can be performed more reliably, knowledge-base construction can proceed somewhat monotonically, without the need to revise or delete previously added knowledge. This section describes how each of the difficult knowledge-base construction tasks is simplified by this technique.

4.3.1 Focused knowledge acquisition

In addition to providing structure to the knowledge, examples can focus the acquisition of the background knowledge of concepts and relationships. They provide a way to divide the general knowledge acquisition task into better defined and more clearly achievable subtasks. Instead of trying to acquire all of an expert’s knowledge about his field, the goal at any given time becomes simply “acquire knowledge relevant to classifying the current example.” This subtask can in turn be subdivided into the smaller and very manageable subtasks of explaining the relevance of each of the features of the example. Thus, knowledge acquisition is methodical and consists of a series of small, well-defined tasks.

Knowledge acquisition simply involves acquiring all knowledge available for each example. This knowledge consists of (1) the names of domain concepts associated with the it, (2) the example structure derived from the expert’s identification procedure for it, and (3) the relationships which exist between the concepts. The concept names and example structure will be naturally generated by the expert in discussing the example. The relationships can be obtained by asking the expert for explanations of the relevance of features of the example.

For instance, all of the knowledge shown in Figures 4.1 and 4.2 could
be acquired in the context of two coffee mug examples. The features of MUG-1 and MUG-2 in Figure 4.1 might be naturally generated in describing the mugs, while the knowledge in Figure 4.2 might arise in explaining the relevance of those features.

A somewhat broad and uniform coverage for initial performance testing can be achieved by simply acquiring the knowledge relevant to an example of each major domain category. At this point, most of the necessary background knowledge should be in place, and adding all the additional examples to achieve truly good coverage of categories will be limited to correcting minor defects or deficiencies. Thus, although the expert cannot generate an exhaustive listing of his knowledge, examples provide a fairly straightforward way to systematically acquire a relatively complete knowledge base.

4.3.2 Integration of new knowledge

Knowledge integration simply involves adding examples of categories and the necessary relations between concepts to explain those examples. These are the small chunks necessary for incremental knowledge-base construction. For instance, after only MUG-1 has been added to the mugs knowledge base, integrating knowledge about MUG-2 consists simply of adding (1) its description as an example of COFFEE-MUG, (2) the relationships required to connect its three new features to the background knowledge entered for MUG-1. Thus, to a large extent, the knowledge base can grow monotonically, without the need for significant backtracking.
4.3.3 Knowledge-base refinement

In this approach even the task of knowledge-base refinement is straightforward. As in traditional systems, it revolves around performance failures, but in this case, the failure is generally rectified by simply adding new examples and their associated background knowledge to the knowledge base, as described in Section 4.3.2.

Errors which require knowledge editing, will usually involve the background knowledge. These errors are not difficult to correct, since the relationships permit understandable explanations to be generated, and unsound ones can be easily identified. These problems should be rare, since the relationships which are present are those which the expert actually supplied, and which can reasonably be assumed to be "true" in a probabilistic sense.

Another type of problem can arise if the number of examples required to cover the space of a category becomes too large to be practical. In this case, it may be desirable to perform some refinement of the example knowledge base to alleviate the problem. Two techniques which can be used to do this are (1) dividing a category into subcategories and (2) merging or generalizing examples.

Dividing a category reduces the number of examples per category by increasing the number of categories. This is a sound practice, since it is a situation in which additional background knowledge has been used to assist in the task of indexing the examples within a category. For instance, in Figure 4.2 it might eventually be desirable to add INSULATED-COFFEE-MUG and PORCELAIN-COFFEE-MUG as subcategories of COFFEE-MUG. The examples MUG-1 and MUG-2, could then be moved into these more specific categories.
Generalizing examples increases the space covered by each example, but introduces the possibility of overgeneralization. However, performing this minimal generalization only as needed, is preferable to always doing it during knowledge acquisition. Moreover, as it is only carried out in the presence of supporting background knowledge, it can be viewed as knowledge compilation similar to the schema learning postulated for human experts [LMSH80, CFG81]. For instance, given a MUG-3 identical to MUG-1 except for the feature GREEN-COLORED, it might be desirable to merge the two into a slightly generalized example. This would be formed by simply dropping the color feature, or replacing it with ANY-COLORED.

4.4 How this knowledge reduces the typical weaknesses

The two types of knowledge present in the proposed representation complement each other very well. This results in a knowledge base which can support a variety of expert tasks, unlike those of typical expert systems.

Examples serve as consistency and problem-solving constraints which restrict the applicability of the general knowledge of concepts and relationships. The examples are accurate snapshots of realizable patterns, but they are so specific that in practice there must always be situations which no stored example covers. The background knowledge is much more widely applicable, but its excessive generality permits it to apply even when it should not. Retaining both types of knowledge permits the weaknesses of one to be complemented by the strengths of the other.

There may exist some optimum level of generality which mirrors an expert's compiled knowledge. However, even if there is, finding this level is far from straightforward, and if found it would be optimal only for one particular
task. The alternative is to retain both specific and general knowledge, and to be able to use the combination of the two to handle a variety of tasks.

4.4.1 Explanation capability

The ability to generate good explanations is provided by the explicit representation of knowledge in a structured form. This capability is in fact an integral and necessary part of knowledge base-construction and use. Explanations of inferences are in terms of the underlying relationships which justify them, and are much more understandable than than the inference traces generated by rule-based systems. Furthermore, since new objects are matched to stored examples, explanations of classifications can be in terms of concrete examples, rather than abstract concepts.

4.4.2 Example-generation capability

Example generation is also trivial to perform using this representation, since the examples are explicitly represented. Even when examples have been generalized through knowledge base refinement, they can be used in a reverse direction to generate specific examples. In fact, example generation is a useful way to test the validity of these generalizations, since if they are sound all the examples covered by them should be realizable instances.

4.4.3 Robustness in classification

The focus of this research is on a representation, and techniques which facilitate the construction and use of an expert knowledge base. Its scope does not currently attempt to construct a knowledge base complete enough to support truly robust behavior. However, since the proposed tech-
nique is designed to facilitate incremental construction of an arbitrarily large knowledge base, it should be useful for this task.

Brittleness in expert systems is a general symptom of missing knowledge. The proposed knowledge representation, in which background knowledge is generally applicable, would seem to encourage entry of extra, seemingly nonessential facts. In a rule-based system, in which all knowledge must be stated in terms of an inference which can be performed, the knowledge enterer would be less likely to enter such "extra", seemingly unimportant knowledge.

Furthermore, the dual representation in which classification models are represented differently from general background knowledge, permits deep knowledge to be present without interfering with the primary task of classification. Most of this task is performed using only the examples, with the background knowledge being accessed only when analytical reasoning ability is needed due to expectation failures. In such cases, the background knowledge is searched only in the region immediately surrounding the expectation failure so that even in this case, the attention of the system is restricted to a small portion of the knowledge base.
Chapter 5

An implementation

The previous chapter described an approach for developing expert systems from examples and explanations. To test this technique, an expert system was developed for the domain of clinical audiology. This was done with the assistance of Professor Craig Wier of the University of Texas Department of Speech Communication, who is an expert in clinical audiology. This chapter describes the ways in which knowledge was acquired, represented, entered, refined and used in this implementation.

The practice of clinical audiology involves the administration and interpretation of a battery of tests, which are useful in detecting and classifying hearing disorders. Classification consists of (1) measurement of the degree of functional loss, (2) identification of the problem site, and (3) if possible, an etiological explanation for the problem. The degree of functional loss is measured directly by the tests and does not really involve any inference beyond test selection. An approximate problem site can usually be inferred from the test results, and this is the primary component of the diagnostic task. Information is often not sufficient for accurate etiological explanations to be formed, although where possible, they are a part of the final diagnosis. This system, focused on the problem of classification and did not address the question of test selection.
I. Conductive hearing loss problems
   A. Fixation of the ossicular chain
   B. Discontinuity of the ossicular chain
   C. Perforation of the tympanic membrane
   D. Middle-ear effusion

II. Cochlear hearing loss problems
   A. Due to noise exposure
   B. Due to exposure to ototoxic substance
   C. Age related
   D. Meniere's disease
   E. Of unknown etiology

III. Retrocochlear hearing loss problems
   A. Eighth-nerve problems
   B. Facial-nerve problems
   C. Brain-stem problems

Figure 5.1: List of diagnostic categories

5.1 Knowledge acquisition

5.1.1 Covering the domain uniformly and adequately

In this system, the knowledge-acquisition task was decomposed by first obtaining the list of general diagnostic categories in Figure 5.1.

This list was easily generated by the expert, and represents the way in which the domain is typically decomposed in textbooks on the subject. It was understood to include only major categories, and to perhaps be colored by the expert’s area of interest. However, it served the purpose of decomposing the domain knowledge in a useful and sound way. The categories in this list thus became the first concept names generated by the expert.

The three major categories refer to the three major types of problem sites. (1) Conductive problems have to do with the outer and middle
ear up to the cochlea, (2) cochlear problems deal with the cochlea, and (3) retrocochlear problems are those which involve the neural pathways up to and including the brain. The conductive and retrocochlear categories are then further subdivided according to site and nature of problem, while the cochlear problems are subdivided primarily by etiology or disease type.

Actual cases became the manageable chunks of knowledge around which knowledge-acquisition sessions could be focused. Thus, the list of categories provided a way to achieve uniform coverage of the domain, while the cases provided a way to incrementally acquire the knowledge to a desired depth.

The initial set of cases used for this project was obtained from a published compilation of actual cases [JJ81]. To build up an initial knowledge base, one or two examples of each major diagnostic category were selected. These were usually prototypical cases, so that they most nearly covered the category space. Choosing such cases also meant that the background knowledge obtained was more generally applicable and better characterized the category concept.

5.1.2 Knowledge acquired for each case

Knowledge acquisition sessions with the expert, revolved around individual cases. To represent the knowledge as examples and background knowledge, certain specific types of information were acquired for each case:

1. The names of the important domain concepts associated with the case. These are the common elements in both the example and background knowledge.
2. The example structures associated with the case and the expert's identification procedure, which consist of:

- Which of the concepts are features used to describe the case at a level of abstraction natural to the expert.
- Which of the concepts are inferred from the features.
- Which of the inferred concepts can be considered final reportable diagnoses for the case.

3. General relationships which exist between the concepts:

- Relationships which relate the features to the inferred concepts.
- Relationships which relate these features to other correlated features in this case or to similar features in other cases.
- Any other relationships which may occur.

The names of concepts and the example structures are somewhat subjective and probably unique to a particular expert and his identification procedure. Thus, in acquiring this knowledge, interference by the interviewer can be a problem and should be avoided.

The background relationships, on the other hand, are derived from the expert's introspection and explanation of his reasoning processes. Their recall is not necessarily triggered during classification, and forcing the expert to reason about relationships at this time can interfere with obtaining a reliable trace of this procedure. This general knowledge is not as subjective, and can be reliably augmented from other sources such as textbooks.
5.1.3 Two interviewing modes

To achieve the somewhat conflicting goals of gaining reliable information about the expert’s identification procedure, and adequately covering the background knowledge, the interviews were divided into two parts. These parts were thought to correspond to two separate modes of reasoning performed by the expert. The first part of the interview attempted to shed some light on the expert’s natural structure for his problem solving knowledge. The second part, on the other hand, attempted to actively prod the expert’s analytical reasoning ability to gain a more complete description of the general background knowledge. This technique is very similar to that described in [KK84] for developing causal models from interviews.

5.1.3.1 “Thinking aloud” interview phase In this phase of knowledge acquisition, the expert was given a report containing all available information on a case. This description consisted of a standard case history and test results, such as in Figure 5.2 (from [JJ81]).

The expert was then asked to evaluate the case, while attempting to think aloud, and asked to try to articulate all concepts which came to mind, without necessarily explaining their relevance. He was not interrupted or prompted for additional information during this phase, unless he seemed to stop saying what he was thinking. This technique was used to minimize the amount of mode switching by the expert, thus permitting him to rely most heavily on his compiled, performance-oriented knowledge. The relevant portions of an actual transcript of this interviewing mode, for the case from Figure 5.2 are shown in Figure 5.3, in which “I” is the interviewer and “E” is the expert.
CASE: 13-1, a 31-year-old female.

The patient complains of a gradually progressive hearing loss in the left ear for approximately 10 months. For the past 2 months, she has noticed a feeling of fullness in the left ear. Approximately 6 years ago, she experienced a transient hearing loss in her left ear that resolved without medical attention. She does not have any hearing complaints about her right ear. She recalls frequent ear infections and temporary hearing loss during childhood.

On otolaryngologic examination, the left tympanic membrane is noted to be unusually thin in the middle portion and unusually thick around the annulus. General tympanic membrane landmarks cannot be identified. The remainder of the examination is within normal limits.

Figure 5.2: Data presented to the expert for Case 13-1
I: This is case 13-1, a 31-year-old female, the problem is in the left ear.

E: The patient is complaining about a gradually progressive hearing loss in the left ear for about ten months, and for the last two months, she's noticed a feeling of fullness in the left ear.

A few ears ago she had a transient loss in the left ear, that resolved without medical attention. She had frequent ear infections, with consequent temporary hearing loss, during childhood.

Let's look at the audiometric test results, here. In the left ear we see a substantial air-bone gap, a 60-dB air-bone gap. That's a big one, about as big as you get.

When we look at speech audiometry, it's what we expect in the presence of a conductive hearing loss. That is, she has a threshold shift, but as soon as she gets the signal 20 dB or so above that shifted threshold, her performance is 100% correct. So it looks like we're probably looking at some kind of conductive hearing loss, here.

Pure tone average, 65 dB in that left ear, kind of what I said when I eye-scanned the audiogram, 60 dB or so loss.

Let's look at tympanometric results. In the left ear, we have a significantly flattened tympanogram, consistent with some sort of fixation of the elements in the middle ear. Things aren't moving with the compliance that they do in a normal ear.

When we look at the acoustic reflex information, we see a pattern in which, for cross reflexes, if we put the signal in the left ear, we have no reflex in the right, and if we put the signal in the right ear, we have no reflex in the left. This is consistent with some fixation, again, in the ossicular chain, so that, even though the reflex may be coming from the opposite side of the head, it can't overcome whatever the fixation is, that's holding the ossicular chain motionless. If we look at the ipsilateral response, again, the information is consistent with that diagnosis. Ipsilateral to the right ear, we have a normal response, ipsilateral to the left ear, we have no acoustic reflex.

Based on the history, and absence of any visible fluid or infection, behind the tympanic membrane in the affected ear, my diagnosis would be that she has some sort of ossicular chain fixation.

Figure 5.3: Transcript of the “thinking aloud” phase for Case 13-1
5.1.3.2 Active questioning interview phase In the second part of the interview, the expert was questioned about the relevance of the concepts which he had mentioned in the first part. He was asked about underlying mechanisms, causal relationships and the significance of data which he did not mention. In particular, he was asked for explanations linking all relevant features of the case to all concepts which were inferred from them during the first part of the interview.

During this portion of the interview, it was supposed that the expert was almost exclusively in his analytical reasoning mode, and that although the knowledge was useful domain knowledge, it was not necessarily closely related to the performance-oriented knowledge. Background knowledge obtained in this manner is preferable to that acquired from textbooks or other experts, since it uses the same concept language. However, it can probably be augmented fairly reliably with knowledge obtained from these other sources. In our case, this augmentation was found to be necessary, to fill in certain inevitable gaps and details, and took the form of textbook data.

Some typical types of questions which were asked during this portion of the interview were:

- "Is Concept C a reportable final diagnosis for this case?"
- "How did you know that Concept C was present in this case?"
- "What is the relevance of Feature F to Concept C?"
- "How is Concept C useful in evaluating this case?"
- "Is this case an example of Concept C?"
• "How would you describe the information in Data D, and how is it relevant to this case?" (where Data D might be the history, pure-tone, tympanometric, speech, or otoscopic data).

• "Did you reach any partial diagnosis before Diagnosis D occurred to you?"

• "Did you form any tentative hypothesis which you later discarded?"

• "Was there any surprising feature present in this case, and if so, how did it affect your diagnosis?"

• "Was any of the data from the other ear relevant to this diagnosis?"

The relevant portions of an actual transcript of this interviewing mode for Case 13-1 (Figure 5.2) are shown in Figure 5.4.

It was recognized that most general commonsense knowledge was not communicated explicitly, and that this knowledge must be used to interpret and augment the domain specific knowledge. This knowledge was necessarily and inescapably generated by the knowledge engineer himself.
I: At what point did you arrive at this diagnosis of ossicular chain fixation?

E: When you look at the tympanometric data, and the stiffened response of the middle ear system to the pressure, here, in the tympanogram. You have a flattened peak, or the absence of a peak, a flat tympanogram.

I: There wasn't anything that tipped you off earlier?

E: What I should say is that you also see this pattern of response in a middle-ear effusion, for example, where the tympanic membrane is stiffened by the fact that it has fluid behind it, instead of air at ambient air pressure behind it. But that's where information like what's available from the otoscopic examination and the patient's history aids you. Because if in fact there's some fluid behind the tympanic membrane, you can typically see that fact upon otoscopic examination. That sort of a pattern, in the absence of any indication of fluid in the middle ear, leads you to an interpretation, that the stiffness is due to some other source. And the only other stiffening mechanism, is to actually fix one of the joints in the ossicular chain in some fashion.

I: In what ways would you characterize the speech audiometry graph, if you were trying to describe it?

E: Basically, the threshold is shifted, so that you basically get 0% correct, until you get to some amplitude of speech, at which point you get a few correct, and that corresponds, basically to the pure-tone average, the hearing loss that you observed. And then, you go up roughly 20 dB from that point, and there already, they get 100% correct. In the presence of a sensorineural hearing loss, that's not what happens to you. You see the threshold shift, and you see the performance increase, with the increase in stimulus level, but not reaching 100%.

I: So, you look at the percent correct. Is it relevant, what amplification it takes to get that 100% correct?

E: Yes, but that's sort of an internal consistency check. You know what their pure-tone average was, when you found their thresholds. So you look and see where they start to detect a few of the speech sounds, and they are able to hear it. It turns that your threshold for 50% correct identification, should be within 5 dB of your pure tone average threshold.

Figure 5.4: Transcript of the active questioning phase for Case 13-1
I: Back to the audiometric data, how would you characterize it? You characterized it in terms of the air-bone gap. Are there any other features there, that jump out at you?

E: Well, when you see a consistent large air-bone gap, across the frequency range, the way we do here, I think there's probably a tendency to look ahead to some kind of fixation, rather than fluid, for two reasons. First, because of the sheer magnitude of the air-bone gap. Second, because it is frequently the case, when you have a effusion in the middle ear, in addition to not having such a large air bone gap, frequently it's on the order of 20-40 dB, the pattern of loss is not exactly what you see here. You often see a rising threshold pattern for the air conductive response. So even if you see a 40-dB loss at 250 and 1000, it may be 35 at 2000 and it may be 30, or so, at 4000.

So I guess, immediate examination of the pattern, and seeing that consistent loss across frequencies, is also a step on the ladder of the diagnosis for fixation.

I: In the tympanometric data, what features would you look at here? You talked about flattened shape, I believe, is there anything else, you would characterize it as?

E: No, that's the strongest evidence.

I: Was there any surprising feature in this case, which made you question your initial hypothesis?

E: No.

I: Did any of the information from the other ear, the fact that it's normal, affect your diagnosis?

E: No, the fact that it's normal meant that basically, I didn't have to worry about it.

Figure 5.4: Transcript of the active questioning phase (continued)
5.2 Knowledge representation

All interviews were tape-recorded, then transcribed and analyzed to extract relevant knowledge and translate it into the system’s representation. This consisted of:

1. Identifying and representing audiologic concepts mentioned by the expert.
2. Assembling some of the concepts into example structures.
3. Identifying and representing relationships between concepts mentioned by the expert.

The concept representation task was performed first, since both the example structures and background knowledge were built up from simple concepts. The other two tasks were somewhat independent, and were carried out by performing a single additional pass through the transcript data.

5.2.1 Concept representation

In this implementation, all audiologic concepts were viewed as assertions regarding a single ear of the patient. All concepts referred to the ear being diagnosed, unless the concept name began with “other”, meaning that the patient’s other ear was being referred to.

The names for these concepts were represented as predicates or propositions. This implementation did not make use of the information contained in the predicate structures, and they were treated as if they were atoms. However, this approach was seen as a way to capture the information in a potentially more useful form than if it was simply recorded as long atoms.
In the left ear, we have a significantly flattened tympanogram, consistent with some sort of a fixation of the elements in the middle ear. Things aren't moving with the compliance that they do in a normal ear.

- tympanogram-shape(flattened)
- fixation(ossicular-chain)
- tympanogram-compliance(low)

When we look at the acoustic reflex information, we see a pattern in which, for cross reflexes, if we put the signal in the left ear, we have no reflex in the right, and if we put the signal in the right ear, we have no reflex in the left. This is consistent with some fixation, again, in the ossicular chain, so that, even though the reflex may be coming from the opposite side of the head, it can't overcome whatever the fixation is, that's holding the ossicular chain motionless. If we look at the ipsilateral response, again, the information is consistent with that diagnosis. Ipsilateral to the right ear, we have a normal response, ipsilateral to the left ear, we have no acoustic reflex.

- contralateral-acoustic-reflex(absent)
- other-contralateral-acoustic-reflex(absent)
- fixation(ossicular-chain)
- other-ipsilateral-acoustic-reflex(normal)
- ipsilateral-acoustic-reflex(absent)

Figure 5.5: Extraction of concept names from text

An extension to this work might involve procedures which decompose these predicates.

During the first pass through the transcript, the concepts were identified, and names for them were written below the containing sections of text, so as to be available during the second pass. An example of this is given in Figure 5.5, in which the left ear was being described.

In the first few interviews, before many concepts had been identified,
She had frequent ear infections, with consequent temporary hearing loss, during childhood.

```
past-history(intermittent-ear-infections)
past-history(fluctuating-hearing-loss)
```

Figure 5.6: Extraction of concept names from text

the concept representation procedure was necessarily very unconstrained, and the only translation performed was the lumping of several words into a single atomic name. However, after certain patterns and commonly used predicate functors were recognized, these were used whenever applicable. For example, in the first paragraph of Figure 5.5, the “fixation of the elements of the middle ear” described by the expert was represented as fixation(ossicular-chain) since this is how it had been described previously. Thus, the process became increasingly efficient, as the system’s vocabulary was established.

As in this example, functors were typically for one-place predicates, in which the functor acted as a slot of a frame, which imparted context to the argument which was its value. Usually, the information which distinguished the concept was contained primarily in the variable name, while the value simply refined it. For example, in tympanogram-compliance(low), the variable tympanogram-compliance was sufficient to distinguish this concept from most other concepts, while the value low simply differentiated it from other degrees of compliance.

However, in some cases, the situation was reversed, with the functor being very general, and the argument which instantiated it being very specific. An example of this is contained in the excerpt in Figure 5.6. Here, the
history(gradually-progressive-hearing-loss)
feeling-of-fullness
past-history(fluctuating-hearing-loss)
past-history(intermittent-ear-infections)
air-bone-gap(severe)
speech-reception-threshold(severe)
speech-intelligibility-score(normal)
hearing-loss-type(conductive)
pure-tone-average(severe)
tympanogram-shape(flattened)
fixation(ossicular-chain)
tympanogram-compliance(very-low)
ipsilateral-acoustic-reflex(absent)
contralateral-acoustic-reflex(absent)
other-contralateral-acoustic-reflex(absent)
other-ipsilateral-acoustic-reflex(normal)
acoustic-reflex-blocked
air-bone-gap
hearing-loss-type(conductive)
middle-ear-effusion
otoscopic-evidence(middle-ear-effusion)
incompressible-fluid-in-middle-ear
hearing-loss(severe)
hearing-loss-type(sensorineural)
speech-intelligibility-score(abnormal)
air-conduction-shape(flat)
bone-conduction-pta(normal)
air-bone-gap-shape(flat)
air-bone-gap-shape(upwardly-sloping)

Figure 5.7: Concept names identified for Case 13-1
Let’s look at the audiometric test results, here. In the left ear we see a substantial air bone gap, a 60-dB air bone gap, that’s a big one, about as big as you get.

\[
\text{air-bone-gap(\text{severe})}
\]

Figure 5.8: Level of abstraction for features

argument was a valid domain concept, while the functor past-history-of was very general and could be instantiated by any domain concept.

Figure 5.7 shows the concepts identified for Case 13-1, from the transcripts in Figures 5.3 and 5.4. After the concepts in a transcript were identified, a second pass was made to determine how the concepts should be structured into examples, and to identify the described relationships between concepts.

5.2.2 Constructing example structures

Each example was formed by grouping a list of feature concepts with a single category concept in an example structure. Although several examples were sometimes derived from a case, they were viewed as having the same set of features, which were the features of the case. Concepts were determined to be features if they appeared to be “directly observable” by the expert.

The level of abstraction at which concepts could no longer be considered features was determined somewhat intuitively. However, the expert seemed to begin his reasoning at a level somewhere above that of numeric data. The following criteria were used to identify this level:

1. Features must be concepts which were quickly identified by the expert
CATEGOR\text{Y}: fixation(ossicular-chain)

FEATURES:

feeling-of-fullness
air-bone-gap(severe)
air-bone-gap-shape(flat)
air-conduction-shape(flat)
bone-conduction-pta(normal)
contralateral-acoustic-reflex(absent)
history(gradually-progressive-hearing-loss)
ipsilateral-acoustic-reflex(absent)
other-contralateral-acoustic-reflex(absent)
other-ipsilateral-acoustic-reflex(normal)
past-history(fluctuating-hearing-loss)
past-history(intermittent-ear-infections)
pure-tone-average(severe)
speech-intelligibility-score(normal)
speech-reception-threshold(severe)
tympanogram-compliance(very-low)
tympanogram-shape(flattened)

Figure 5.9: Example structure derived for Case 13-1

in describing a case.

2. Any abstraction of actual data must involve only (a) visual processing of figures and graphs, (b) logical combination of observable features, or (c) procedural manipulation and tokenization of numeric data.

For example, in the transcript excerpt in Figure 5.8, the expert quickly performed some abstraction from the graphically represented pure tone data. He first estimated that the average difference between the air conduction and bone conduction curves was approximately 60 dB and called
this the “air bone gap”. He then indicated that he thought of this as being a “substantial” gap “about as big as you get”. After this almost instantaneous preprocessing, which was regarded as part of observation, it seemed that all reasoning was done with the concept of this large gap. In a previous interview, the expert had described this degree of gap as being “severe”.

In this implementation, the only examples created for each case were those of the final diagnostic categories. In the transcripts in Figures 5.3 and 5.4, the only diagnosis assigned by the expert was fixation(ossicular-chain). Thus, the only example derived for Case 13-1 was that shown in Figure 5.9.

5.2.3 Structuring concepts into a network of relationships

Capturing the background knowledge associated with a case, required representing as many as possible of the relationships described in the interview. At this point, if some obvious deficiencies in the knowledge were identified, they were reduced by questioning the expert or by examining textbooks. Deficiencies became apparent if an inferred concept could not be adequately related to any other concept, or if a relevant feature could not be related to any of the inferred concepts. Background knowledge was represented as a semantic network in which concept nodes were connected by relationships.

Agonizing over explanations was not thought to be necessary during the early stages of knowledge acquisition. It was only important that relationships not be made unsoundly strong. If the relationship between two concepts was not well understood, and there was no immediate way to resolve the uncertainty, the relationship was simply not represented. This might have prevented the system from making as strong an inference as possible,
but any serious problem would eventually be remedied by the natural process of knowledge addition.

### 5.2.3.1 The relation language

Relationships were described using the following set of general link types:

1. Current to successor state, temporal relationships: *causes, is caused by*  
   \[ \text{fixation(malleus)} \xrightarrow{\text{causes}} \text{tymanogram-compliance(very-low)} \]

2. Structural to functional mappings: *has function, is function of*  
   \[ \text{ossicular-chain} \xrightarrow{\text{has function}} \text{conduct-sound(tympanic-membrane,cochlea)} \]

3. Set inclusion relationships: *has generalization, has typical specialization*  
   \[ \text{fixation(malleus)} \xrightarrow{\text{has generalization}} \text{fixation(ossicular-chain)} \]

4. Part to whole relationships: *has part, is part of*  
   \[ \text{ossicular-chain} \xrightarrow{\text{has part}} \text{malleus} \]

5. Logical inference relationships: *implies, if and only if*  
   \[ \text{perforation} \xrightarrow{\text{implies}} \text{tymanogram-data(unobtainable)} \]

6. Feature to category relationships: *is sufficient for, requires*  
   \[ \text{otoscopic-evidence(middle-ear-effusion)} \xrightarrow{\text{sufficient for}} \text{middle-ear-effusion} \]

7. Feature to feature or category to category relationships: *co-occurs with,*  
   \[ \text{pta(severe)} \xrightarrow{\text{equivalent to}} \text{pure-tone-average(severe)} \]

In order to provide a finer-grained link description language, and to permit representing quantitative estimates, the following *qualifiers* were used for relationships. Qualifiers captured additional knowledge by clarifying the source and degree of uncertainty in inferences using the relation.
1. Qualifiers describing estimates of frequency: *always, usually, sometimes, occasionally, rarely*

   immobility(outer-middle-ear) \(\xrightarrow{\text{sometimes caused by}}\) fixation(malleus)

2. Qualifiers describing strength of relationship: *strongly, moderately, weakly, very weakly*

   air-bone-gap-shape(flat) \(\xrightarrow{\text{weakly sufficient for}}\) fixation(ossicular-chain)

3. Qualifiers describing certainty that the relationship exists: *certainly, probably, possibly, conceivably*

   cholesteatoma \(\xrightarrow{\text{possibly caused by}}\) longterm-history( intermittent-ear-infections)

4. Qualifiers describing temporal characteristics: *instantly, quickly, gradually*

   longterm-history( intermittent-ear-infections) \(\xrightarrow{\text{possibly gradually causes}}\) cholesteatoma

Figure 5.10 shows some relationships identified from the transcripts in Figures 5.3 and 5.4.

The relation language preserved more of the background knowledge provided by the expert than was possible in rule-based systems. Although their semantics were not well defined, explicitly representing the relationships made them available to be used as effectively as the current understanding permitted.

In this implementation, the primary interpretation of a relationship was a numeric estimate of the strength of any inference which it supported. For example, given

\[
A \xrightarrow{\text{relation}} B
\]

the number derived from this relation was an estimate of the certainty with which concept B could be inferred, given that concept A was known to be
past-history(ear-infections)\text{ causes }\text{past-history(fluctuating-hearing-loss)}

air-bone-gap(\text{severe})\text{ sometimes sufficient for }\text{fixation(ossicular-chain)}

\text{co-occurs with }\text{hearing-loss(\text{severe})}\text{ co-occurs with }\text{speech-reception-threshold(\text{severe})}

\text{co-occurs with }\text{pure-tone-average(\text{severe})}

\text{has generalization to }\text{air-bone-gap}

\text{co-occurs with }\text{hearing-loss-type(\text{conductive})}\text{ co-occurs with }\text{air-bone-gap}

\text{incompressible-fluid-in-middle-ear}\text{ causes }\text{tympanogram-shape(\text{flattened})}

\text{fixation(ossicular-chain) causes tympanogram-shape(\text{flattened})}

\text{fixation(ossicular-chain) causes tympanogram-compliance(\text{very-low})}

\text{incompressible-fluid-in-middle-ear causes tympanogram-compliance(\text{very-low})}

\text{middle-ear-effusion requires otoscopic-evidence(middle-ear-effusion)}

\text{fixation(ossicular-chain) causes acoustic-reflex-blocked}

\text{incompressible-fluid-in-middle-ear causes acoustic-reflex-blocked}

\text{acoustic-reflex-blocked causes ipsilateral-acoustic-reflex(absent)}

\text{acoustic-reflex-blocked causes other-contralateral-acoustic-reflex(absent)}

\text{weakly sufficient for air-bone-gap-shape(\text{flat})\text{ causes fixation(ossicular-chain)}}

\text{middle-ear-effusion causes incompressible-fluid-in-middle-ear}

\text{middle-ear-effusion usually requires air-bone-gap-shape(upwardly-sloping)}

\text{speech-intelligibility(abnormal) co-occurs with hearing-loss-type(sensorineural)}

\text{Figure 5.10: Relationships derived for Case 13-1}
true. The functions for evaluating the inference strengths of relationships were developed by examining instances of their use in various domains. These functions are described in Section 5.5.

5.2.3.2 Conjunction and negation In this semantic network representation, all relations entering a node were implicitly disjunctive. Any single concept could be used to infer another concept, if they were connected by a strong enough relationship.

This simplification was not viewed as a serious problem, since conjunctive relationships could have been added later, and would simply have represented stronger but more specific knowledge. Furthermore, any realizable conjunction of concepts can be viewed as a generalization of an example, so that this knowledge should eventually be captured by the example knowledge base.

In this implementation, all relationships were positive in nature, and no relationships were represented which disconfirmed concepts. The closed-world assumption was in effect, with concepts being assumed to be absent or false unless there was positive evidence for their existence.

This simplification did not affect an ultimate diagnosis, since all hypotheses were tested, but hypotheses could not be discarded during hypothesis selection. As with conjunction, this negative knowledge could always have been added, without the need to revise the existing knowledge. Furthermore, the expert seemed to use such a positive reasoning strategy, since he typically mentioned only positive evidence for a hypothesis.
5.2.3.3 Representational assumptions and simplifications To summarize, some of the representational assumptions and simplifications made in this implementation were:

1. All concepts can be assertions about a patient from the perspective of a single ear.

2. A relatively small set of general relations is sufficient to compute inference strengths between related concepts.

3. Explanations can be represented using a loose network of context-free relations.

4. Only positive relationships need be expressed to achieve adequate performance. Concepts are assumed to be absent unless there is positive evidence for their existence.

5. Conjunctive relationships do not need to be expressed in the explanation language to achieve adequate performance. Conjunction of concepts is captured by the conjunction of features in examples.

These simplifications did not prevent good classification performance from being achieved, because as more examples were added to the system, less generalization was needed to permit the examples to cover the space of categories, and less reliance was placed on the explanation generation mechanism.

5.3 Knowledge entry

After a few sessions involving several cases, the derived concept language was assumed to be stable enough that knowledge-entry could begin. It should be noted that the stabilization process was desirable only to permit parsimony of concept names, and was not necessary in principle.
A knowledge management tool was developed to facilitate the entry, inspection and editing of knowledge. This tool had two modes, one for unfocused exploration and editing of concepts and relationships, and one for case-oriented entry of examples and associated background knowledge. After a transcript was analyzed, the relationships were first entered using the unfocused mode, then the example was entered, and any missing background knowledge was filled in using the example-oriented mode.

5.3.1 Entering background knowledge

The knowledge entry tool permitted the background knowledge to be viewed in two different ways: (1) from the perspective of a single concept, examining and editing its relationships to neighboring concepts, or (2) in the context of the relationships between two concepts, with the ability to find and edit explanations between them. The first viewing technique was found to be more useful for “browsing” through the knowledge base, moving from concept to concept. On the other hand, the second method was more useful in entering or improving useful relationships.

Whenever a relationship was entered, the user was prompted for any qualifiers which should be added to it. This permitted the user to enter relationships which he was not certain of, or which might not always apply. To make the relationship useful in both directions, the tool also requested qualifiers for the inverse relationship. If no qualifier was supplied for a relationship, the tool attached a “typical” strength to it. Explicitly encouraging the user to qualify relationships permitted sounder inference by the system, since the qualifiers represented a form of domain-specific knowledge. The process also forced the user to more deeply examine the relationship, triggering
Figure 5.11: Relationships from Case 13-1 after addition of qualifiers
realizations about additional, deeper knowledge, which better explained the relationship.

Figure 5.11 shows how the relationships from Figure 5.10 were augmented to include qualifiers and inverse relationships.

5.3.2 Adding examples

The knowledge-entry tool permitted adding a new example to the knowledge base by either entering all of its features, one at a time, or by copying a closely related example and editing it. After the features were added, it was necessary to ensure that there was sufficient background knowledge to explain their relevance. This knowledge was needed to permit the example to be retrieved and generalized appropriately.

5.3.2.1 Determining the relevance of features As described in Chapter 4, the relevance of each feature can be described in terms of its reminding strength and importance. If a satisfactory explanation of both these forms of relevance did not already exist, then the background knowledge was augmented so as to include one. What constituted a satisfactory explanation, depended on the feature involved, and was determined by the user. Spurious features did not require any explanations, and had no reminding power or importance.

In this implementation, it was assumed that reminding and importance strengths could be estimated by evaluating the relevance of features to the example’s category. Although in some cases, a feature appeared to be more relevant to the example than to its category, there was no way to compute this greater relevance, since explanations could be formed only between
concepts. Such a problem usually was an indication that the current category concept was too general, and should be divided into subcategories. This technique was in fact found to be useful, and is described in the Section 5.4 on knowledge base refinement.

Reminding strength was estimated by computing the numeric certainty with which the diagnostic category could be inferred from a feature. Importance, on the other hand, was approximated by the numeric certainty with which a feature could be inferred from the diagnostic category. This was actually a computation of expectation value, but if a feature was very important, then it could be expected to be present with a high degree of certainty. Inference strengths were computed by forming explanations and evaluating them, using a method described in Section 5.5.

To avoid needlessly recomputing reminding and importance values, they were stored as reminding and importance links. This action represented a form of computation caching, analogous to some of the indexing which occurs in human memory. If values had been previously computed, they were checked for agreement with the current best explanations.

Explanations generated by the system during computation of remindings and importances were very useful in identifying deficiencies or problems in the knowledge base. In fact, much of the background knowledge was entered as a consequence of this process. After the first example of each category, the amount of additional background knowledge needed for subsequent examples was typically not very great, and limited to information unique to them.

Figure 5.12 shows a portion of the background knowledge containing explanations of the relevance of features of the example in Figure 5.9 to the
contralateral-acoustic-reflex(absent) 
  occasionally caused by 
    usually causes 
      
conductive-hearing-loss(severe) 
  always co-occurs with 
    always co-occurs with 
      
air-bone-gap(severe) 
  sometimes sufficient for 
    usually requires 
      weakly sufficient for 
        usually requires 
          has moderately typical specialization 
            has generalization 
              
fixation(malleus) 
  always causes 
    sometimes caused by 
      
fixation(ossicular-chain) 
  always causes 
    sometimes caused by 
      
acoustic-reflex(blocked) 
  sometimes caused by 
    always causes 
      
other-contralateral-acoustic-reflex(absent) 
  sometimes caused by 
    always causes 
      
tympanogram-shape(flattened) 
  sometimes caused by 
    

tympanogram-compliance(very-low) 
  sometimes caused by 
    

Figure 5.12: Some explanations relevant to the example from Case 13-1
<table>
<thead>
<tr>
<th>FEATURE</th>
<th>REMINDING</th>
<th>IMPORTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>feeling-of-fullness</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>air-bone-gap(severe)</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>air-bone-gap-shape(flat)</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>air-conduction-shape(flat)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>bone-conduction-pta(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>contralateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>0.64</td>
</tr>
<tr>
<td>history(gradually-progressive-hearing-loss)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ipsilateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>1.0</td>
</tr>
<tr>
<td>other-contralateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>1.0</td>
</tr>
<tr>
<td>other-ipsilateral-acoustic-reflex(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>past-history(fluctuating-hearing-loss)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>past-history(intermittent-ear-infections)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pure-tone-average(severe)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>speech-intelligibility-score(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>speech-reception-threshold(severe)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>tympanogram-compliance(very-low)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>tympanogram-shape(flattened)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Figure 5.13: Remindings and importances for the example from Case 13-1

category fixation(ossicular-chain). These explanations were the strongest ones present in the knowledge base after the above process of explanation retrieval and editing for each feature. The numeric reminding and importance values computed from these explanations are shown in Figure 5.13.

Since the importance of features was closely related to their expectation value, it was used to estimate the prototypicality of the example. A rough prototypicality value was computed by calculating the average importance over the features of the example. Thus, using the values in Figure 5.13, the prototypicality of the current example was estimated to be 0.32. The use of expectation value to estimate prototypicality is analogous to the representativeness heuristic demonstrated to exist in humans [TK74]. This heuristic is effective because in the absence of actual statistical information, it is a fairly
good estimate of probability. This value was used by the system to select the example of a category most likely to match a new object, or to generate the most typical example of a category.

5.4 Knowledge refinement

In expert systems, knowledge refinement typically consists of (1) correcting inconsistencies, errors, or deficiencies in the knowledge base, and (2) modifying the knowledge base to improve performance.

5.4.1 Correcting problems

In this implementation, correcting problems in the knowledge base was not a significant task. Performance errors were typically due to deficiencies in the knowledge and were correctable by the normal process of knowledge addition. This was especially true for examples, whose addition always made the coverage of the knowledge base more complete.

Errors in the background knowledge sometimes occurred when explanations were added with a particular context in mind. In these situations, the explanations were too strong and needed to be weakened or made more specific. This type of error was usually discovered quickly, since the system constantly generated explanations during classification, computation of remindings and importances, or unfocused knowledge base inspection. A significant problem resulted in unsound explanations, which were easily identified and corrected.
5.4.2 Improving performance

The knowledge-based-modification techniques employed to improve performance at the classification task were:

1. Dividing a category into subcategories
2. Merging examples of a category

These techniques were aimed primarily at reducing the number of examples in a category. This was useful because reminders were computed using only background knowledge, and therefore led only to categories. From this point it was still necessary to select a single model from the examples of the category. If these were great in number, examining them all could be prohibitively expensive.

Furthermore, the computed estimates of importance used to match an example were most reliable if the space of objects covered by an example most nearly matched the space of its category concept. Thus, a goal in knowledge-base refinement was to gradually and soundly move toward a knowledge base in which categories were represented by a small number of prototypical examples.

5.4.2.1 Dividing a category into subcategories Dividing a category into subcategories reduced the number of examples per category by increasing the number of categories. The examples of the original category were partitioned, and moved to the new subcategories. The new categories had concepts associated with them, and were therefore related to other concepts in the knowledge base, so that more specific reminding and importance estimates could be computed. Since this process was driven by the acquisition
Figure 5.14: Specializations of fixation(ossicular-chain)

<table>
<thead>
<tr>
<th>FEATURE</th>
<th>fixation(oss)</th>
<th>fixation(mal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>feeling-of-fullness</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>air-bone-gap(severe)</td>
<td>0.6</td>
<td>0.36</td>
</tr>
<tr>
<td>air-bone-gap-shape(flat)</td>
<td>0.4</td>
<td>0.24</td>
</tr>
<tr>
<td>air-conduction-shape(flat)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>bone-conduction-pta(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>contralateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>0.144</td>
</tr>
<tr>
<td>history(gradually-progressive-hearing-loss)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ipsilateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>0.144</td>
</tr>
<tr>
<td>other-contralateral-acoustic-reflex(absent)</td>
<td>0.24</td>
<td>0.144</td>
</tr>
<tr>
<td>other-ipsilateral-acoustic-reflex(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>past-history(fluctuating-hearing-loss)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>past-history(intermittent-ear-infections)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pure-tone-average(severe)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>speech-intelligibility-score(normal)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>speech-reception-threshold(severe)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>tympanogram-compliance(very-low)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>tympanogram-shape(flattened)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Figure 5.15: Effects of moving Example 13-1 to fixation(malleus)
of knowledge, it was sound, and was viewed as extending the background knowledge “into” the space of a category.

This type of knowledge-base refinement was generally performed whenever the knowledge to do it became available. It was in fact applied to the case of the example in Figure 5.9. This became possible when knowledge was entered describing the ways in which fixations could involve different elements of the ossicular chain. The example was moved to the subcategory fixation(malleus) and the reminding and importances were re-evaluated in this context. The effects of this process are shown in Figures 5.14 and 5.15. The reminding and importance values for fixation(ossicular-chain) were not discarded, because reminding to supercategories were used by the classification procedure, and importance values could usually be inherited by subcategories.

5.4.2.2 Merging examples Merging examples reduced the number of examples per category by reducing the number of examples directly. Merging examples also usually involved some slight abstraction of features, so that fewer examples were required to cover the space of a category. This abstraction was also beneficial in increasing the average importance of the features. This resulted in a higher prototypicality value for examples which were more frequently matchable.

However, this process involved permanent generalization of examples, and could therefore lead to unreliable knowledge losses. For this reason, it was very rarely performed in this system, and only when the amount of generalization involved was very small and justified by the background knowledge.
5.5 Knowledge use

5.5.1 Explanation and inference

In this system, explanations and inference were very closely related. In fact, inferences were made by constructing explanations and evaluating their strength. As in MYCIN, explanations were simply traces of the problemsolving behavior of the system. However, since the relationships which underlie inferences were explicitly represented, the explanations could rely on these relationships and were much sounder. Thus, explanations were the justifications for inferences performed by the system, and inference strength was estimated, on the basis of underlying relationships.

This technique for representing and using relationships split every relationship between two concepts into two unidirectional components which were represented separately. These served as justifications for the inferences which were performed in the corresponding directions. For example, a causal relationship was represented using two directed components as follows:

\[
A \xrightarrow{\text{causes}} B \xleftarrow{\text{caused by}}
\]

The causes relationship could be used to justify and quantify the inference of B from A, while the caused by relationship lent support to inferences in the other direction.

In this system, the simplifying assumption was made that the degree to which a concept could be inferred from another, could be estimated solely from relationships which existed between the concepts. It was also assumed that the inference strength supported by a composite relationship, involving intermediate concepts, could be estimated by combining the independently
evaluated inference strengths associated with the component primitive relationships.

Thus, for the composite relationship

\[ C_1 \xrightarrow{R_1} C_2 \xrightarrow{R_2} C_3 \]

the certainty with which \( C_3 \) could be inferred from \( C_1 \) depended on the independently computed inference strengths for \( R_1 \) and \( R_2 \). Furthermore, the final belief in \( C_3 \) depended on the original belief in \( C_1 \) from which it was inferred. Since inference strengths and certainties are conceptually similar to conditional and simple probabilities, they were represented using numbers between 0 and 1. These numbers were combined in a way which paralleled Bayes' rule, so that the belief in \( C_2 \) derived from the belief in \( C_1 \) was estimated by

\[
\text{belief in } C_1 \times \text{inference strength for } R_1
\]

and the belief in \( C_3 \) derived from \( C_1 \) through the above path was approximately:

\[
\text{belief in } C_2 \text{ derived from } C_1 \times \text{inference strength for } R_2
\]

or

\[
\text{belief in } C_1 \times \text{inference strength for } R_1 \times \text{inference strength for } R_2
\]

More generally, given an explanation,

\[ C_1 \xrightarrow{R_1} C_2 \xrightarrow{R_2} \ldots \xrightarrow{R_{n-1}} C_n, \]

the belief in \( C_n \) derived from belief in \( C_1 \) was computed as:

\[
\text{belief in } C_1 \times \prod_{i=1}^{n-1} \text{inference strength for } R_i
\]
ipsilateral-acoustic-reflex(absent) \( \xrightarrow{\text{sometimes caused by}} \) acoustic-reflex-blocked
\[ \text{link strength} = 0.6 \]

acoustic-reflex-blocked \( \xrightarrow{\text{occasionally caused by}} \) fixation(ossicular-chain)
\[ \text{link strength} = 0.4 \]

ipsilateral-acoustic-reflex(absent) \( \xrightarrow{\text{composite explanation}} \) fixation(ossicular-chain)
\[ \text{explanation strength} = 0.6 \times 0.4 = 0.24 \]

Figure 5.16: Computation of inference strength from explanations

In this system, inference strengths for primitive relationships were computed by the following procedure:

IF there are no modifiers

THEN return default strength of base relation

ELSE return product of modifier strengths

The underlying idea was that if the relation had been qualified by modifiers, these should be used to evaluate inference strength, otherwise, the "typical" inference strength associated with the relation should be used. The modifier and default strengths were tabulated, general estimates for these values. For example:

- The inference strength of \textit{causes} was 0.9, because the default strength of \textit{causes} was 0.9.
- The inference strength of \textit{usually causes} was 0.8, because the qualifier strength of \textit{usually} was 0.8.
- The inference strength of \textit{usually gradually causes} was 0.48, because the modifier strength of \textit{gradually} was 0.6.
In finding explanations, a uniform-cost search with cycle detection was employed, with cost being measured using the above function. An example of how an inference strength was computed is shown in Figure 5.16 for one of the explanations from Figure 5.12.

The only deviation from the independent-combination assumption involved certain pruning heuristics, used to remove paths containing unsound explanation patterns. The only patterns identified for use in this implementation were the following:

\[
\begin{align*}
A & \xrightarrow{\text{has generalization}} B \xrightarrow{\text{has specialization}} C \\
A & \xrightarrow{\text{causes}} B \xrightarrow{\text{caused by}} C
\end{align*}
\]

These heuristics were based on the fact that the \textit{has specialization} and \textit{caused by} links were interpreted under the assumption that no previous specialization or cause, respectively, was known for the current concept. However in the above patterns, this assumption is invalidated.

5.5.2 Classification

Classification was a cycle involving two alternating phases of hypothesis generation and hypothesis testing. The generation phase was a data directed, unfocused process, in which remindings were used to select an example to serve as a model. The testing phase was a more constrained, model-directed process, in which an attempt was made to match the example to the data.

The high-level classification algorithm was:
compute reminding to categories
WHILE user is not satisfied and there are untried reminding DO
    select best untriedcategory reminding
    WHILE user is not satisfied
        and there are untried examples of category DO
        select best untried example of category
        compute match between new object and example
    ENDWHILE
ENDWHILE

In this procedure, computing reminding, selecting the best category
reminding and selecting the best category example were all part of hypothesis
selection, while computing the match between the new object and the example
was hypothesis testing. In this implementation, all hypotheses selected and all
matches computed, were subject to user approval. This was due partly to the
experimental nature of this system and partly to the fact that a high degree of
interaction with the user was desirable during knowledge-base construction.
However, it is easy to visualize how these decisions could be automated using
various user-specified thresholds.

5.5.2.1 Hypothesis generation

Computing category reminding The procedure for computing reminding
from a list of features of the new object was as follows:

FOR each feature in new object DO
    FOR each reminding link from feature to a category DO
        IF this is first reminding to category
            THEN add category and reminding strength to list of reminding
        ELSE combine new reminding strength with previous one
    ENDFOR
ENDFOR
Figure 5.17: Reminding strengths and specialization links for Example 13-1

Multiple remindings to the same category were combined so that the overall reminding strength always increased, but could never exceed 1.0. Remindings were combined incrementally as follows:

\[ R_{new} = R_{previous} + R_{current} - R_{previous} \times R_{current} \]

Figure 5.17 shows the remindings computed from the features of Example 13-1 (Figure 5.9), together with the associated specialization/generalization hierarchy.

Selecting the best category reminding  Selecting the best category reminding consisted of initially selecting the strongest reminding from the list of untried remindings. However, this process was slightly more complicated, since sometimes the strongest reminding was to a category which did not have any explicitly stored examples. This could happen because, as described in
Section 5.4, a category was sometimes split into subcategories, with the examples being divided among the subcategories, but the reminding links to the parent category were preserved. Thus, the complete procedure was as follows:

select strongest category reminding
IF category has examples
    THEN return category
ELSE return best subcategory reminding

Choosing the best subcategory reminding was simply a recursive call, in which the new set of remindings to be considered was the set of remindings which were subcategories of the initial reminding.

In Figure 5.17, the strongest combined reminding from the features of the new example was to fixation(ossicular-chain). However, if this category had no examples, the system would have looked for an example in fixation(malleus), which was the subcategory of fixation(ossicular-chain) with the strongest reminding.

Selecting the best example of a category  The best example of a category was assumed to be the untried example which was most prototypical. As described in Section 5.3, prototypicality was a precomputed and cached value, which ranked examples of a category by expectation value.

5.5.2.2 Hypothesis testing After an example reminding had been selected, the system computed a match between the example and the new object using the following procedure:
FOR each feature of example DO

find best explanation for inferring feature given new object features
combine explanation strengths into overall match strength

Explanations were computed using a backward-chaining uniform-cost search, using the inference computing function described in Section 5.5.1. Thus, for each feature of the example, the strength with which it could be inferred from the set of features of the new object was computed. The importance of example features was used to reduce search effort for less important features. Search was only performed until the strength of the inference dropped below the unimportance of the feature (where, unimportance = 1 – importance). Thus, an explanation strength could be no lower than the unimportance of the feature, so that less important, unmatched features had less effect on the match than more important, unmatched features.

Explanation strengths were combined into an overall match strength, using a multiplicative function. Thus, for an example having the features $F_1 \ldots F_n$, the strength of the match was given by:

$$\prod_{i=1}^{n} \text{strength of best explanation for } F_i$$

Figure 5.18 shows how a malleus fixation example from another case was matched to Example 13-1 (Figure 5.9). In this case, all the features were either matched perfectly or unimportant, except for air-bone-gap(severe). This feature could only be matched through a weakly equivalent relation to air-bone-gap(moderate), with an inference strength of 0.4. Therefore, the overall match strength was also 0.4.

An important point is that just as independence of the individual inferences was assumed in computing the strength of an explanation, indepen-
dence of the features of an example was assumed in computing the strength of a match. Since this assumption is generally not valid in the probabilistic sense, the computed values could not be assumed to be actual probabilities. However, these values were effective as relative, heuristic measures of goodness of fit.
FEATURES OF NEW EXAMPLE

- air-bone-gap(moderate)
- air-bone-gap-shape(flat)
- air-conduction-shape(flat)
- bone-conduction-pta(normal)
- contralateral-acoustic-reflex(absent)
- ipsilateral-acoustic-reflex(absent)
- other-contralateral-acoustic-reflex(absent)
- other-ipsilateral-acoustic-reflex(absent)
- pure-tone-average(moderate)
- speech-intelligibility-score(normal)
- speech-reception-threshold(moderate)
- tympanogram-compliance(very-low)
- tympanogram-shape(flattened)

MATCH TO STORED EXAMPLE

<table>
<thead>
<tr>
<th>EXEMPLAR FEATURE</th>
<th>MATCH DESCRIPTION</th>
<th>MATCH STRENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>feeling-of-fullness</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>air-bone-gap(severe)</td>
<td>weakly equivalent</td>
<td></td>
</tr>
<tr>
<td>air-bone-gap-shape(flat)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>air-conduction-shape(flat)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>bone-conduction-pta(normal)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>contralateral-acoustic-reflex(absent)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>history(gradually-progressive-hearing-loss)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>ipsilateral-acoustic-reflex(absent)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>other-contralateral-acoustic-reflex(absent)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>other-ipsilateral-acoustic-reflex(normal)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>past-history(fluctuating-hearing-loss)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>past-history(intermittent-ear-infections)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>pure-tone-average(severe)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>speech-intelligibility-score(normal)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>speech-reception-threshold(severe)</td>
<td>unimportant</td>
<td>1.0</td>
</tr>
<tr>
<td>tympanogram-compliance(very-low)</td>
<td>matched</td>
<td>1.0</td>
</tr>
<tr>
<td>tympanogram-shape(flattened)</td>
<td>matched</td>
<td>1.0</td>
</tr>
</tbody>
</table>

OVERALL MATCH STRENGTH

0.4

Figure 5.18: Matching a new example to Example 13-1
Chapter 6

Conclusion

This thesis has identified some of the problems with current expert-system techniques which make construction needlessly difficult and result in knowledge bases having limited utility. The common element in these problems is the failure to retain certain types of relevant knowledge during construction of the knowledge base. This has resulted in problems, because knowledge usefulness has been evaluated in too narrow-minded a way, and even in this limited context it could not be evaluated reliably. The information loss has made the resulting knowledge base incapable of supporting anything it was not specifically designed for, including its own construction. Furthermore, the unreliability of this process has resulted in the need for continuous revision, just to achieve good performance at the classification task.

The proposed solution to these problems can be viewed on three levels of abstraction: (1) The general theme which underlies this research, (2) the representation scheme and general techniques for using it, and (3) the specific expert-system implementation and techniques, as described in the previous chapter. Thus, in discussing and evaluating this solution, it is useful to do so on each of these levels.
6.1 The general theme

The general idea behind this research is that the most cost effective approach to knowledge-base construction is one of acquiring and representing all available knowledge. This philosophy maximizes the benefits realized from the effort involved, since the task does not need to be repeated whenever the knowledge base is put to a different use. Furthermore, as has been demonstrated in this thesis, this approach facilitates the process of initial knowledge-base construction.

Construction of any knowledge base requires that a representation language be selected. Since this selection must be done early, it potentially poses the biggest obstacle to the implementation of these ideas. Some general properties which the representation should have are:

- It should be flexible and complete enough so as to not restrict the knowledge which can be retained.
- The knowledge should be in a form which is acquirable easily and incrementally, with a minimum amount of required transformation.
- It should be capable of supporting all identifiable types of expert tasks.

Although in practice there will always be limitations to the way any representation can be used, the above goals should be foremost in the selection process. Once the representation has been selected, then it is the responsibility of the knowledge engineer to make maximal use of its capabilities during knowledge-base construction, so that the greatest benefits can be achieved. As discussed in Chapter 4, the representation described in this thesis has these characteristics in the context of diagnostic expert systems.
6.2 The representation and techniques

The general approach used in this research consists of acquiring classification examples and explanations, and retaining them in an integrated network of examples and relations between concepts. This scheme is useful for acquiring, representing and using most of the knowledge reliably available from experts in most domains. There are, however, some general limitations to its applicability.

It is restricted to domains and tasks in which the classification space can adequately be covered by examples

The proposed approach relies on matching to past examples, and so requires knowledge of this set of possible solutions. Therefore, it is only applicable to domains in which the categories are enumerable, and each can be covered by a reasonably small number of examples.

Expert systems which operate in domains not meeting these conditions, must be capable of constructing a solution [Cla85]. Since the proposed technique does not include operators for synthesizing a new example\(^1\), it does not have this capability. The background knowledge and its inference mechanisms could, in principle, be sophisticated enough to support this type of solution construction. However, it is not apparent that this knowledge could be acquired incrementally from explanations.

\(^1\)Modifying another example does not qualify as construction in the sense described by Clancey.
It is restricted to domains in which explanation knowledge is available.

The domain must also be one in which explanations between concepts can be generated. The expert can perform classification by reasoning \textit{categorically}, relying primarily on compiled knowledge and without needing to generate explanations. However, the evolving expert system must have the ability to also reason \textit{probabilistically} [SP78]. This is necessary because (1) the knowledge can never be complete, (2) the knowledge must be used in unforeseen ways, and (3) the compiled categorical associations cannot be learned directly.

Probabilistically quantified associations can be produced by the expert only through the process of generating an explanation, and evaluating its strength. In general, an expert cannot reliably estimate statistical probabilities directly [TK74]. These estimates are accurate only to the extent to which the expert can generate reliable explanations. Therefore, the domain must be one in which the theories are strong enough that fairly sound explanations are possible.

\textbf{6.3 The implementation}

The actual expert system developed using these ideas was described in detail in the previous chapter. As is always the case with prototype systems, some assumptions and compromises were made limiting its effectiveness. The remainder of this section describes some directions for future research aimed at correcting some of the weaknesses of this implementation.
Richer representation for relationships

The most significant limitations of this implementation were a consequence of the oversimplified language for the background knowledge. This language forced a certain amount of transformation to be performed on the background knowledge to represent it. Many explanations generated by the expert could not really be captured at all. Although the resulting knowledge loss was less than it would have been for a rule-based representation, it could be reduced significantly if the representation were augmented in the following ways:

- **The set of primitive relationships could be richer.** There were many relationships expressed by the expert, such as “suggests”, “results in” and “permits”, which were not present in the relation language, and are probably not exactly expressible using it. \(^2\)

- **The language could be extended to permit conjunction.** Currently, all relationships are implicitly disjunctive; however, there appeared to be explanations involving relationships between conjunctions of concepts rather than single concepts.

- **The language could be extended to permit negation.** In this implementation, all relationships must be captured in a positive way. Capability for negating concepts could be valuable, and the set of relations should be augmented to include negative relationships such as “prevents”, or “is mutually exclusive with.”

\(^2\)This extension still would assume that a domain-independent set of relations could be used. Removing this assumption, although perhaps useful, is viewed as an extension to the paradigm rather than the implementation.
• The language could be extended to permit statements about the context in which a relationship is valid. This context might deal with the current goals, the current data or other facts relevant to the current state of the system.

These extensions would be useful in permitting much more of the available knowledge to be captured during knowledge representation. However, to make use of this additional knowledge, the knowledge-use algorithms would need to be improved accordingly.

**Sounder inference mechanism for generating explanations**

In addition to extensions necessary to make use of a richer background knowledge-representation language, the explanation-generation mechanism could be improved to make better use of the information already available:

• **The mechanism could make use of additional heuristics involving the whole explanation.** With the exception of the two pruning heuristics described previously, explanations were evaluated for goodness by independently evaluating the component relationships. Although effective, this was obviously an oversimplification, and there are probably many other general explanation patterns which could cause the explanation to be strengthened, weakened or pruned.

• **The mechanism could make use of heuristics involving the current context.** In addition to attaching a context to relationships, the system could know of generally applicable heuristics dealing with the current state of the system.
• The mechanism could make use of the information about the nodes connected by the relation links. There was a considerable amount of knowledge about the concepts represented, but not used by the explanation-generating procedures. This knowledge includes the structure of the concept-name predicates, the relationships with neighboring concepts, and any examples connected to the concept through has example or feature of links. One can easily conceive of explanation-evaluating heuristics which make use of this type of lateral information while passing through a concept.

More powerful representation for classification procedures

In this implementation, knowledge of the expert’s classification was retained only in examples, and these examples were represented simply as a flat list of features. A more powerful representation may sometimes be needed to fully describe examples or capture procedural information:

• Examples could be represented in a more structured way. The features of an example could be structures, rather than simple concepts. This might be necessary to describe the physical subassemblies of an object, such as the ears of a patient. It might also be useful to describe complex, identification subgoals.

• The identification mechanism for some concepts could be defined procedurally. An expert may sometimes consciously use a procedure to process data into higher level concepts. This is the case with numeric audiologic data, for which some standard computations, such as averaging are often performed. There could be a way to represent and invoke this procedural information.
Why these extensions were not necessary

The view in building this system, was that the approach was sound enough that many of the weaknesses would be overcome by the simple process of continued incremental addition of examples and background knowledge. This hypothesis proved to be correct, since even after only a small number of examples had been added, the system was able to perform well for most other cases having the same diagnoses. More significantly, overall performance seemed to continuously improve as more examples were added, so that the technique should be capable of incrementally achieving truly expert levels of performance.

Furthermore, the knowledge base which resulted from these efforts seems to be a valuable collection of understandable facts about the domain of audiology. The general nature of this knowledge makes it a powerful resource in its own right. Even if the processing components were found to need extensive revision, all the effort of knowledge-base construction would be retained in this flexible knowledge base. The same situation would apply if the knowledge were applied to new uses.

The portion of this approach which requires the greatest human effort is the construction of the background knowledge base, but this work can be retained across changes in the various algorithms which use and compile the knowledge. Therefore, future research should probably first address development of the knowledge-representation language, so that a very complete knowledge base can be constructed (along the lines of [PMS*87,LMS86]). Experimentation with the various algorithms and heuristics can then be performed more meaningfully, and without the need to repeat the laborious task of knowledge-base construction.
BIBLIOGRAPHY


89


Programming Project, Computer Science Department, Stanford University, 1978.


in Medical Artificial Intelligence, pages 131–159, Addison-Wesley, 1984.


