1 Introduction

Auscultation is a technique used in cardiac physical examination to detect irregularities by analyzing heart sounds. Interpretation and analysis of the auscultatory findings require expertise because of the wide range of applicable information. A heart murmur, for example, can be interpreted in terms of its intensity, location, and duration in the cardiac cycle; supplementary data on its quality, pattern, directions of radiation, and changes with physiologic maneuvers provide additional evidence of the possible causes. Presence of multiple murmurs, extra heart sounds, and variations in the first and second heart sounds further complicate the analysis.

This paper introduces **Murmur Clinic** [4], an expert system for auscultation. Development of the system is motivated by the desire, in accordance with the general research goals of Artificial Intelligence in Medicine [6], to:

1. understand and formalize the reasoning process of expert physicians;
2. study and apply Artificial Intelligence techniques; and
3. provide a clinically useful decision-making aid in auscultation.

This paper will first present the design of the system, and then describe the diagnostic algorithm. A brief summary of ongoing effort will also be included.

2 The Design

Empirical studies on the physician's problem-behavior showed that diagnostic reasoning is a hypothesize-and-test process [5]. **Murmur Clinic** was designed to simulate an expert physician. Its diagnostic reasoning process is modeled after the description of Ledley and Lusted [3]:

1. Acquire input information on case;
2. Evaluate the relative importance of the different signs and symptoms; and
3. Generate and evaluate hypotheses, then produce decisions by eliminating unlikely candidates.

Hypothesis generation involves retrieval of hypotheses (abnormalities) from the physician's long-term memory as he detects one or more manifestations in the input information [5]. Evidence for and against competing hypotheses is gathered until a hypothesis seems clearly better. When inconsistent findings are detected, the physician tries to explain the inconsistency or rejects the hypothesis in question. The **Murmur Clinic** diagnostic algorithm captures these phenomena through the maintenance of an *active* and a *dormant* set of hypotheses, indicating whether the hypotheses are under consideration. Hypotheses are moved from one set to the other when conflicting evidence is detected. Cases like the following can be well handled by the system: Suppose only a holosystolic murmur is observed, together with an opening snap. This indicates the co-existence of mitral regurgitation (MR) and mitral stenosis (MS), but the MS murmur was too faint to be noticed. Two sets of evidence are applicable here: MS would not be considered initially because the MS murmur is diastolic, and only one systolic murmur is heard. Since the opening snap is pathognomonic for MS, however, MS should actually be considered as one of the topmost decisions.

In hypothesis evaluation, each active hypothesis is evaluated by matching the pattern of findings to the physician's concepts of manifestations of the abnormality [2]. The matching process draws on the associations between the findings and the causal manifestations. It is, therefore, incorporated into the algorithm as pair-wise comparisons among possible values of the manifestations.

Moreover, the algorithm captures an important practice in auscultatory diagnosis: If multiple murmurs are present, which indicate presence of multiple abnormalities, diagnosis of the diastolic murmur, if any, is carried out first. This is because conclusions on the diastolic murmur can be evidencing to the other, presumably systolic, ones. This feature reflects the physician's tendency to search for a basis for diagnostic expectation, and his ability to first identify the relevant clinical features that should prove fruitful for further investigation [2].

The design of the system can be illustrated by answering the following questions [4]:

- How much knowledge should the system have?

- How should the knowledge be organized and applied?

- How should the system interact with the user?

The answers are discussed in turn in the following subsections.

2.1 Scope Of The Knowledge Base

- *How much knowledge should the system have?*

The system must have enough knowledge to make a diagnosis from auscultatory findings. Three types of knowledge are involved in the process, emphasizing different epistemological structures: the *associational knowledge* of symptomatic relations, the *discriminatory knowledge* about the degree of importance of symptomatic entities, and the *judgmental knowledge* of medical facts.

Interpretation and analysis of auscultatory findings involve mostly symptomatic comparisons between the findings and the manifestations of the causal abnormalities. An illustration of such associational knowledge is the degree of *similarity* between a finding and a manifestation of the same category, *e.g.*, the intensity of a heart murmur.

Different causal manifestations, and hence different findings, have different degrees of discriminatory importance. For example, in deciding about mitral stenosis (MS), presence of the opening snap is pathognomonic evidence; a grade 1 intensity is less important. The degree of importance of a particular manifestation can be supplemented by the frequency with which it appears in patients with the corresponding abnormality. Discriminatory knowledge is needed to determine this aggregate significance.

Retaining the most likely hypotheses as diagnostic results involves the exercise of medical facts and experience. These are captured in the judgmental knowledge, which is applied in situations like the following: When only a diastolic murmur is observed, all abnormalities that would manifest a systolic murmur should be excluded from consideration.

The associational knowledge incorporated into the **Murmur Clinic** system is mainly extracted from various medical textbooks and publications; the discriminatory and the judgmental knowledge involve work done by interrogating human experts.

DS	Value	Interpretation
1	10	Almost pathognomonic.
2	7	Important-- Strongly weighted evidence for the corresponding abnormality.
3	3	Negligible--Characteristic but not a discriminating finding.

Table 1: Interpretation of Discriminatory Significance.

2.2 Knowledge Representation

- *How should the knowledge be organized and applied?*

Modularity and explicitness are emphasized in the knowledge representation because the three categories of knowledge in the system are to be maintained and updated independently.

The associational and discriminatory knowledge are represented by discrete conceptual structures. These include representations for individual causal abnormalities, and tables of similarity between the auscultatory findings and the causal manifestations.

Both the characteristics of the possible causal abnormalities and the auscultatory findings reported by the user are encoded in structures called *murmurs*. A murmur is an entity with the following attributes: *intensity, quality, pattern, location,* accompanying *heart sounds, maneuver* and *radiation*. Each of these attributes has one or more manifestation values. Except in murmurs that represent input findings, each manifestation is assigned a *discriminatory significance* (DS) and an *occurrence frequency* (OF). The DS indicates the importance of the manifestation in distinguishing the corresponding abnormality from the others; the OF indicates how often a patient with the abnormality exhibits the manifestation. The representation of a causal abnormality, called a *reference murmur*, is shown in Figure 1. Table 1 and Table 2 show the interpretations of the DS and OF values. The numerical values indicated are used to compute the *decision-factor* to be described later in this paper.

The pair-wise similarity scores of all possible manifestation values of a particular murmur attribute are held in a *score-table*. Mechanisms for manipulating the score entries in the table are provided.

Figure 1: A murmur

OF	**Value**	**Interpretation**
A	10	Always present.
U	7	Usually present.
S	3	Sometimes present (Rare).

Table 2: Interpretation of Occurrence Frequency.

The judgmental knowledge, consisting of medical facts, is most naturally represented as production rules of the "IF...THEN..." form. These rules can be classified into several subcategories, each performing a well defined subtask such as checking input consistency, recognizing further differentiability of the hypotheses, ruling out certain possibilities, etc. In some of these subtasks, a single evidence can override the rest; in others, a few or all of the facts should be considered before making a conclusion. Consequently, the rules are characterized as either "terminating" or "non-terminating". Execution of a terminating rule inhibits activation of all other rules in the same group; execution of a non-terminating rule simply performs its actions as side effects. A rule responsible for detecting further differentiability among the top few decision choices is an example of a terminating rule, while a rule for checking input consistency a non-terminating one.

Because of the choice of representations, the **Murmur Clinic** knowledge base lacks the compactness and conceptual simplicity of conventional rule-based systems [1]. The gain in explicitness, however, enables the diagnostic reasoning process of a physician to be modeled closely.

Figure 2: Input Panel Of The Murmur Clinic

2.3 User-Interface and Explanations

- *How should the system interact with the user?*

Since descriptions of auscultatory findings involve entities such as heart sounds and murmur patterns, locations and areas of radiations, pictorial representations are much more accurate and convenient than textual ones. Therefore, the user is provided with an input-panel where he can, by using the mouse, choose from a list of findings in pictorial representations. In order to provide as “user-friendly” a user-interface as possible, keyboard input from the user is avoided during interaction in general; the user is usually asked to choose from a list of possible answers from a menu for a particular question. The user-interface makes extensive use of the Symbolics 3600 Lisp Machine graphics facilities, and the implementation issues can be found in [4]. The screen display during a session is shown in Figure 2.

With the goal of a clinically useful decision-making aid in mind, the design of **Murmur Clinic** incorporates an explanatory component that enables it to justify and explain its reasoning to the user. This component is built on top of two features that keep track of the decision history of the diagnostic process: *evidence-record* and *score-template*. The evidence-record keeps track of whether a particular abnormality is being considered for diagnosis, and the evidence that dictate this status. It is useful for explaining to the user why and how a decision is made in the face of conflicting evidence. The score-template records the similarity scores between the findings and the manifestations of a hypothesis. It is useful for explaining to the user why a particular abnormality is ranked over others.

In addition to describing the individual reference murmurs and the judgmental rules encoded, four types of questions can be answered by the system about the reasoning in the diagnostic procedure. The questions, pertaining to each list of possible causes produced (*i.e.*, corresponding to each input murmur), are as follows:

- WHY is [abnormality] ranked first?
- WHY is [abnormality] NOT ranked first?
- WHY is [abnormality] ranked as it is?
- WHY is [abnormality-1] ranked OVER [abnormality-2]?

The answers are generated by tracing the decision history. Figure 3 shows an example of the answers produced in a session.

3 The Algorithm

The diagnostic algorithm of the **Murmur Clinic** decides on the possible causal abnormalities from the input auscultatory findings, verifying the existence of single or multiple abnormalities. The decisions are presented as a ranked list of possible causes for each abnormality determined to be present in the case. Mechanisms for explaining and justifying the reasoning involved are also provided.

The algorithm, which can be summarized in Figure 4, is as follows:

Figure 3: Example Of A Query

Step 1

Initial auscultatory findings are entered by the user. These findings are described in terms of heart murmurs, each with characteristics of pattern, location, intensity, quality and optional areas of radiation, changes with physiologic maneuvers, and variations of and/or presence of extra heart sounds. If more than one murmur is present in the input information, the following steps are carried out successively once for each input murmur, starting with the diastolic component, if any.

Step 2

The input findings, corresponding to a single murmur from here on, undergo a consistency test. If inconsistent or incomplete input is detected, *e.g.*, only the intensity of a heart murmur is given, the user will be warned and asked if he wishes to change or add more information. If the input is consistent and reasonable, or the user does not wish to change or add information after being warned of inconsistency, the procedure enters the next step. Otherwise new input is obtained and goes through the same consistency test again.

Step 3

The input findings are checked for evidence to rule out each reference murmur in the database as a possible cause. Based on the evidence from the input, each abnormality is classified into either the active or the dormant state, indicating whether it is suitable for consideration in the diagnosis. An evidence-record is generated for each of these abnormalities, keeping track of the evidence which induced the classification. Positive evidence is recorded together with the evidence-score in the evidence-record. The higher the evidence-score, the stronger the indication of the abnormality.

Step 4

For each abnormality in the list of active candidates for diagnostic considerations, a score-template is generated. Filling in a score-template involves the following steps:

1. The similarity score of each manifestation with respect to the corresponding finding is determined from the score-table. There are seven score-tables for the seven attributes of a heart murmur. The similarity scores are encoded in a scale from 0 to 10, with 10 indicating an exact match, and 0 for inconsistency.

   Each of the similarity scores are scaled by a *decision-factor* on a scale from 0 to 1. The decision-factor is determined by the discriminatory significance (DS) and the occurrence frequency (OF), of the particular manifestation (see Table 1 and 2).

2. The weighted similarity scores of a particular attribute is summed and entered into the corresponding slot in the template. Only those attributes in which input information is provided are evaluated; otherwise a default value is used.

3. The *overall similarity score* is generated by summing up the similarity scores of all its attributes and the evidence-score from the evidence-record for the abnormality.

Step 5

The abnormalities are ranked according to their overall similarity scores. The abnormality with the highest score is the most likely cause of the murmur.

Step 6

The ranked list is checked for further differentiability as follows: If the first two or three abnormalities are further differentiable by some physiologic maneuvers, and input information in this regard is not provided, the user is prompted for additional information. If additional maneuver information is provided, the process returns to **Step 2**. If the user does not provide more information, or the differentiability test failed, the process returns to **Step 2** and continues with the next set of murmur findings. Notice that each tentative diagnosis can provide further evidence for the diagnosis of the murmurs that follow. This is most true of diastolic murmur evidence followed by systolic murmurs. The order in which co-existing systolic murmurs are diagnosed is unimportant. The cycle terminates when the list of input murmurs is exhausted.

Step 7

At this point a ranked list of the possible causal abnormalities with respect to each input heart murmur is available. The evidence-records for these abnormalities are then checked to see if those with pathognomonic evidence are among the highest ranked choices. If not, each of the non-highest ranked abnormalities with pathognomonic evidence is separated into a new ranked list by itself.

Step 8

All the ranked lists go through a final test for discrepancies, and are displayed as the diagnostic results, together with a brief description on how the decisions were made.

Figure 4: The Algorithm

4 Conclusion

Although evaluation and refinement are still under way, **Murmur Clinic** is an operational system showing competence in its performance. It is capable of handling sophisticated cases of multiple murmurs and conflicting evidence, and shows promise of improvement as the knowledge base is enhanced and refined. Its explicit knowledge representation and modular design have greatly eased the development process and the study of the diagnostic reasoning mechanisms of a physician.

Immediate goals for the ongoing work include upgrading the explanatory component and enhancing the knowledge base. When the knowledge base is more complete, definitive diagnosis, *i.e.*, declaring particular causes as being the most consistent with the evidence, will replace the current tentative diagnosis that produces a ranked list of possible causes for each set of findings.

5 Acknowledgments

I wish to thank Bill Long and Ramesh Patil for their guidance and encouragement, and Bob Jayes for his valuable comments on the system.

References

[1] Davis, R., Buchanan, B., and Shortliffe, E. "Production Rules as a Representation for a Knowledge-Based Consultation Program", in *Artificial Intelligence, 8*, pp15-45, North-Holland, 1977.

[2] Kassirer, J. P. and Gorry, G. A. "Clinical Problem Solving: A Behavioral Analysis", in Reggia, J. A. and Tuhrim, S. (Eds), *Computer-Assisted Medical Decision Making, Vol. 2*, pp 85-107, Springer-Verlag, 1985.

[3] Ledley, R. S. and Lusted, L. B. "Reasoning Foundations of Medical Diagnosis", in *Science, 130*, July 3, 1959, pp 9-21, American Association for the Advancement of Science, 1959.

[4] Leong, T. *Murmur Clinic: An Auscultation Expert System*, TM-319, MIT Laboratory for Computer Science, 1987.

[5] Reggia, J. A. and Tuhrim, S. "An Overview of Methods for Computer-Assisted Medical Decision Making", in Reggia, J. A. and Tuhrim, S. (Eds), *Computer-Assisted Medical Decision Making, Vol. 1*, pp 3-45, Springer-Verlag, 1985.

[6] Szolovitz, P. "Artificial Intelligence and Medicine", in Szolovitz, P. (Ed), *Artificial Intelligence in Medicine*, pp 1-19, Westview, 1982.