Submitted version CV10

REPRESENTING A MEDICAL KNOWLEDGE BASE FOR MULTIPLE USES

Paul D. Clayton, Peter J. Haug, T Allan Pryor, and Ove B.Wigertz[1]

Dept of Medical Informatics, LDS Hospital/ University of Utah, Salt Lake City Utah 84143

1) Dept of Medical Informatics, Linkoping University, S-581 83 Linkoping Sweden

It is possible to use medical expertise which was originally developed to diagnose disease to accomplish additional tasks as well. We used a knowledge base which had been developed to diagnose pulmonary diseases to measure the amount of diagnostic information contributed by various sources of clinical and historical information and to measure the effect of variability in radiologist's performance. Using this source of medical expertise as a standard, we found that substantial variation in diagnostic accuracy resulted from the differences in reported findings among several radiologists. We are also using the knowledge base to implement a model for generating a patient-specific consult. Based upon these experiences and the accomplishments of others, we now espouse a model in which the medical expertise in the knowledge base is represented in a high-level, procedural format which 1) can be accessed by multiple compilers or interpreters to accomplish a variety of different tasks and 2) can be shared among users in multiple institutions.

Introduction

If the knowledge base (medical expertise) of a system truely contains sufficient power/ knowledge to accurately diagnose disease, then this same knowledge can be used to accomplish alternative ojectives as well. Based upon this philosophy, specific modes of the inference engine in the MYCIN and INTERNIST-1 programs used the diagnostic expertise stored in the knowledge base to explain the reasoning which lead to a conclusion or to decide which questions to ask as part of the diagnostic process. This notion of multiple use was expanded as Clancey conceived GUIDON (1) and Warner developed the LERN program (2); both of these applications used the medical expertise in the knowledge base to teach students.

In the present article, we will briefly describe our experience with three examples of additional alternative uses of a knowledge base which was originally constructed to diagnose pulmonary diseases: 1) the ability to measure the degree to which different pieces of information affect diagnostic accuracy, 2) the ability to measure how variability in physician performance can affect the ultimate diagnostic conclusion, and 3) the ability to selectively convey to a physician the important facts which are already known about a patient. We first describe the features of our method of representing the knowledge base and then conclude that the same approach which facilitates the use of the knowledge base to accomplish multiple objectives may also provide a practical way of sharing knowledge among groups which have used different software tools to implement decision-making applications.

Methods

Several features in the design of the knowledge based system known as HELP(3) have contributed to its routine use in patient care: 1) the decision-making system is imbedded within a comprehensive clinical information system, 2) the knowledge base, though separate from the decision-making program, is constructed in a very modular, procedural representation, and 3) the appropriate decision logic is activated whenever patient-specific data (facts) referenced in the decision criteria are stored in the clinical information system.

The HELP oriented knowledge base consists of knowledge modules which were originally called "sectors" and now are increasingly referred to as frames. Each frame contains the logic necessary to make one specific decision and exists in at least two different formats: 1) the high level representation which is an ASCII format with a slot oriented structure, and 2) compiled "HELP frames" which are executable. By using compilers or interpreters designed with alternative objectives in mind, the expertise in the high level frames can be compiled into formats which allow multiple uses of the expertise contained in the knowledge base.

While the HELP system has always had two levels of representation, the current evolution of the system (4) tends to make the high level of representation higher and the low level lower. By this statement we mean that the original HELP "sectors" were created and stored in a specific format which could only be accessed by a special purpose editor (HCOM). The current representation is an ASCII format and could be created by almost any editor as long as the knowledge base author observed the proper syntax. Instead of intepreting the low level representation of the "sector," we are now compiling the logic for more efficient execution. Depending upon the design of the compiler, the object code could run on a variety of different hardware or operating systems.

This more general type of approach is possible because the frames themselves contain the logic which determines how they are to be evaluated. Because of the procedural nature of the frames, the medical knowledge base is able to support a variety of the many mature decision-making models (IF...THEN...rules, probabalistic or other scoring algorithms which rank differential diagnoses, query for "important" missing data, etc.) rather than relying upon one specific type of inference engine which contains the procedural information about the decision model.

The following frame is an example of the high level syntax which is used in our most recent representation of decision logic. This diagnostically oriented frame was originally developed to calculate a Bayesian likelihood score that a specific patient has pneumonia.

Title: Pneumonia diagnosis (7.141.1).

Type: Diagnosis

Author: Peter Haug.

Date: 12/12/86

Message: "<disease_prob (val; #.##)> Pneumonia (history)".

Variables: chest_pain as (DO YOU HAVE CHEST PAIN?),
cough as (HAVE YOU HAD A COUGH WITH THIS ILLNESS?),

fever_or_chills as MAX(fever, chills)
where fever is (HAVE YOU RECENTLY HAD A FEVER?)
and chills is (HAVE YOU HAD CHILLS RECENTLY?)
if Exist (fever) or Exist (chills),

•
•
•

Statistics: for fever_or_chills with Sensitivity(YES, 0.85; NO, 0.15),
and Specificity(YES, 0.3; NO, 0.7),
for cough with Sensitivity(YES, 0.9; NO, 0.1),
and Specificity(YES, 0.2; NO, 0.8),

•
•
•

Logic: disease_prob = 0.014.
If Exist(fever_or_chills) then disease_prob = Bayes(disease_prob, fever_or_chills).
If Exist(cough) then disease_prob = Bayes(disease_prob, cough).

•
•
•

If disease_prob LT 0.014 then finish.

Ask: Patients(fever, chills, cough) Heirarchical.

Evoke: If chest_pain EQ YES or fever EQ YES or chills EQ YES or cough EQ YES.

Urgency: 5/9

Gold Standard: If ICD_pneumonia and pneumonic_infiltrate

The various slots in the frame are labeled with bold face names and the logical criteria in this case are a series of Bayesian probability revisons of the initial likelihood of 1.4% that an inpatient in our hospital will have pneumonia. In order to properly qualify the terminology which is used in the logical rules, the terms (variables/symbols) used in the rules are declared. These declarations refer to terms (denoted by capitalized words) which are defined in the system data dictionary and this mapping would have to be done individually for non-HELP oriented users with different types of software and database systems. These declarations may use chronological constraints as well as qualification regarding the value for a variable. The procedural logic is written in a slightly structured, but understandable, language. The relative "urgency "of recognizing this disease is given along with gold standard criteria which allow evaluation of the accuracy of the logic.

We will now describe examples which illustrate how we used this representation of knowledge to investigate the information content of diagnostic procedures, to evaluate the performance of radiologists and to assist the radiologist to interpret chest films.

<u>Information content</u>. Using this frame for diagnosing pneumonia and similar frames for 29 additional pulmonary diseases, we measured the information which a specific item or class of items contributed to the diagnostic process by turning on and off the Bayesian calculation in the logic section which references that particular item or set of items. We compiled the altered frames, ran the logic on a representative population with and without using the information contributed by the specified item and analyzed the accuracy of the resultant lists of differential diagnoses. We compared the magnitude of the likelihood score and the rank of the actual disease in a population of 517 patients (220 of whom had a variety of pulmonary diseases and 297 of whom showed no evidence of pulmonary disease). We examined the accuracy of the lists by simulating situations in which limited subsets of data would be available: 1) the thirteen "most important" history questions, 2) a comprehensive history without the benefit of radiographic findings, 3) radiographic findings without history information, and 4) a combination of history and radiographic information about the patient.

<u>The effect of radiologist's performance</u>. We were also able to use the basic diagnostic expertise to evaluate the differences in diagnostic contributions of several radiologists who all read the same set of chest films. It has long been recognized that there are substantial differences in the findings which are reported by radiologists. These differences appear to be the result of different confidence thresholds, judgements about the importance of a particular finding, and visual perception disparities. We investigated the overall effect of these variations by modifying the logic so that it would retrieve only the set of radiographic findings reported by a specified radiologist. We then ran the logic on a population of 100 patients (52 normals) using the clinical history and the radiographic findings reported by each of four different radiologists.

<u>Patient-specific consultation</u>. Our final example of the alternate uses of the knowledge base is a model in which the computer decides which facts about a patient should be presented to a radiologist who is interpreting chest films. It is recognized that human interpretation of tests (radiographs, ECGs, etc) is strongly influenced by the currently known clinical facts about the patient. We are currently implementing a compiler that will parse through the knowledge base and construct an inverted file for each item of information. For each indicant in each disease, a weight based upon the sensitivity and specificity obtained from the statistics section of the frame is calculated. The overall importance (I_k) of each indicant (k) is calculated as the sum of a product of the probability of the existence of a disease (Di) multiplied by the weight (w_{ik}) of that indicant for that disease summed over all diseases: $I_k = \sum D_i * w_{ik}$.

At the time the radiologist reviews a film, the likelihoods of each disease will be calculated for a specific patient and, using the inverted file constructed by the compiler, we will calculate the "most important" indicants which are known about the patient and present this "intelligent" selection of information to the radiologist. We hypothesize that this consultation will improve the diagnostic performance of the radiologist.

of measuring radiologists' performance have been compounded because there is an associated degree of uncertainty. The radiologist may over or under-read films. By integrating the findings into the total decision-making process, one can ascertain the end effect of the differences. It should be noted that we have asked the radiologist to enter discreet objective findings into the computer; there may be cognitive interpretations which, based upon the experience of the radiologist, represent a higher level of knowledge than the disassociated collection of objective pieces of the puzzle. We are in the process of gathering data to investigate this possibility.

Based upon our experience in these three circumscribed applications, we feel that there is substantial benefit to having a high level representation of the knowledge base which can be accessed by multiple special purpose compilers which can translate the expertise in the knowledge base into a variety of applications. In addition to the multiple uses of the knowledge base for explanation, education, consultation, studies of efficacy (information content), evaluation of physician performance, critiquing, etc., such a high level representation has the additional benefit of being transferable among different institutions. Each systems group which presently has their own data structure and operating system environment could write a compiler to transfer the high level representation into their operational environment. The variables declaration section of each decision frame would have to be mapped to the symbol table or data dictionary which is peculiar to each installation; other than this requirement, the logic could be translated into the appropriate language for each system. Of course the fact that our knowledge base representation is procedurally oriented would make difficult the assimilation by systems in which the procedural knowledge for one specific type of decision model is contained in the inference engine rather than the knowledge modules (frames).

With these reservations we feel that future accomplishments in the field would be hastened if different institutions could share knowledge in a high level format and then compile it into the specific types of applications or local representations which are desired. It is generally agreed that the content of the knowledge base is the source of power, not the particular mode of representation. Our experience, as well as the work of others, illustrates the potential for generating multiple applications from a rich knowledge base.

References:

1. Clancey WJ. The use of MYCIN's rules for tutoring. Int J Man-machine Studies 11:25-49 (1979)

2. Warner HR. Computer assisted medical decision-making New York:Academic Press 1979 p.182-184.

3. Pryor TA, Gardner RM, Clayton PD, and Warner HR. The HELP system J Med Syst 7:87-102. 1983

4. Ostler MR, Stansfield JD, Pryor TA. A new, efficient version of HELP. Proceedings 9th SCAMC, pp 296-297 November, 1985.

5. Clayton PD, Shelton P, Haug PJ, Harada SK, Tocino I, Morrison WJ, Elliott CG, Collins DV, Frederick PR. Differences in diagnostic outcome when chest films are read by different physicians. Radiology 161(P):327, 1986.