Interfacing the Heart Disease Program to Coded Data

William J. Long[1], Hamish Fraser[1, 2], Shapur Naimi[2]
[1]MIT Lab for Computer Science, Cambridge, MA,
[2]Tufts-New England Medical Center, Boston, MA

This poster examines the problems of interfacing the Heart Disease Program to other online resources by translating the vocabulary into one of the commonly used medical dictionaries.

The Heart Disease Program (HDP)[1] is an expert system to help physicians diagnose heart disease. It takes a patient description from the physician, including history, symptoms, physical examination, and pertinent laboratory results. The program provides a differential diagnosis consisting of hypotheses accounting for the findings provided which are summarized for the user. The program was developed as an independent tool, but it is clear that to incorporate it into the world of the physician and make it useful, it has to take data from existing sources and provide results in forms that can be used by the other tools in a medical information system. The first step in this process is to provide a translation from existing medical vocabularies to that of the HDP input and from the HDP output into that vocabulary.

THE CODING PROCESS

This poster examines the relationship between the vocabulary of the HDP and that provided by the UMLS[2] and READ[3]. The output vocabulary of HDP includes over 200 terms and the input vocabulary includes about 500 terms, although if the available modifier combinations were countered the numbers would be far higher. We used the Web accessible version of the UMLS and concentrated on the commonly used dictionaries it contains: ICD9, SNOMED, COSTART, and WHO. With these and READ we have coded the terms using the closest available dictionary terms and present the results of this process.

PROBLEMS OF CODING

Issues confronting the coding process include:

1. Codes sometimes do not make distinctions made by the HDP. Eg., HDP considers renal failure to be more severe than renal insufficiency while UMLS considers them equivalent.
2. Some of the concepts used for non-cardiac diagnoses represent groupings not present in the dictionaries. Eg., the concept of upper GI disease to cover the range of diseases that can mimic anginal chest pain and the concept of pulmonary parenchymal disease that causes symptoms that mimic pulmonary congestion. Both of these represent groupings that require multiple terms in any of the dictionaries.
3. The biggest challenge is the number of properties that are important for representing clinical distinctions in input concepts but are not available in the vocabularies. Eg., murmurs are very important to cardiac diagnosis and the HDP allows extensive specification of the murmurs while only a few of the dictionaries cover physical examination findings and of those only READ allows as many as 8 different kinds of murmurs.

The different vocabularies have different strengths and weaknesses. Vocabularies such as READ and SNOMED that allow combinations of terms provide more flexibility in representing terms whose specificity is determined by the needs of a domain-specific decision support system. Some of the dictionaries focus almost entirely on diseases, while others cover the findings, drugs, and procedures as well. Because of these differences we also examine the coverage of the dictionaries with respect to the different kinds of information used by the HDP.

CONCLUSIONS

These analyses allow us to generate reasonable expectations and strategy for developing the kinds of interfaces to domain specific expert systems like the HDP and the role of a controlled vocabulary as provided by general medical dictionaries. This analysis also suggests some ways in which the dictionaries should expand to anticipate the needs of decision support systems.

References

1. Long WJ, Naimi S, Criscitiello MG. Evaluation of a New Method for Cardiovascular Reasoning. J Am Med Inform Assoc. 1994;1:127-141.

2. Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The Unified Medical Language System: an informatics research collaboration. J Am Med Inform Assoc. 1998; 5: 1-11.

3. Robinson D, Schulz E, Brown P, Price C. Updating the Read Codes: user-interactive maintenance of a dynamic clinical vocabulary. J Am Med Inform Assoc. 1997;4: 465-472.