OntoQuest: A Physician Decision Support System based on Ontological Queries of the Hospital Database

Mihail Popescu and Gerald Arthur
Health Management and Informatics Dept., University of Missouri, Columbia, MO

ABSTRACT

OntoQuest is a physician decision support system that mines the hospital data base for previous decisions made in cases similar to the current one. For example, OntoQuest displays a list of the medications prescribed to similar historical patients from which the physician may compare his choice of medication for the current patient. This information retrieval is accomplished using ontological queries. Unlike a regular database query, an ontological query is able to account for semantic similarity between patients. To implement the ontological query, we propose a method for computing the ontological similarity between patients represented by sets of ICD-9 diagnoses. We have tested the OntoQuest prototype on a pilot data set of 2077 patients. Finally, we compare the OntoQuest performance to conventional database queries. We believe that OntoQuest can be extended to compare services, quality and outcomes among patient and provider groups.

INTRODUCTION

In an Institute of Medicine report released in 1999 it was stated that there are 44000 to 98000 deaths caused by medical errors every year. With the introduction of computerized systems in each health care institution a huge data repository was created that represents, in itself, an invaluable asset. In this work we propose a physician decision support system (PDSS) that uses the hospital data repository to reduce medical errors. The PDSS is data-driven and is based on just-in-time data mining (JITDM) of the hospital medical data repository. The main advantages of a data driven PDSS over a knowledge driven PDSS, such as expert systems, are that they reflect the current state of medical knowledge and the current practice in a particular hospital.

The proposed OntoQuest system answers queries like: show me diagnoses for patients with SIMILAR symptoms, show me medications for patients with SIMILAR diagnoses, show me information from the medical record related to this condition. In a sense, OntoQuest is a query by example (QBE) system. In this paper we present a pilot version of the OntoQuest system that deals with patient medications. By identifying the medication recommended to similar patients (by him/her or by physicians from his/her own organization) the physician has a method of verifying his/her own therapeutic recommendations. The medication previously prescribed for similar patients is displayed in a drop-down box that is used by the physician to choose the appropriate medication for the current (query) patient.

Ontological similarity techniques have been previously used in fuzzy databases, web searching and bioinformatics. The ontological query concept was introduced by Andreasen in the context of fuzzy databases [1]. Ontological methods related to web searching were mentioned in [2]. Lord [3] and Popescu [4] used Gene Ontology to compute ontological similarities between gene products. Various rule-based expert systems have been proposed in the literature [5]. In contrast to these, our approach is data-driven and is based on just-in-time data mining (JITDM) of the hospital medical data repository.

Methods

The key concept of the above system is patient similarity which may be computed using an ontology and ontology related algorithms. Ontology is understood, in this work, to indicate a controlled vocabulary together with a term hierarchy. A controlled vocabulary is necessary for matching similar or related medical terms from the patient record for better retrieval sensitivity. On the other hand, a term hierarchy can improve the retrieval sensitivity of a database search by finding patients with diagnoses that are medically "close" in the given hierarchy. In this work we use the ICD-9 codes for describing patient diagnoses.

Example 1.

The similarity between two patients P_1 and P_2 with the following ICD-9 diagnoses, P_1={250.01, 401.1} and P_2={250.02 and 401.9}, is 0 if an exact (database search) match is used. That is, patient P_2 will never be retrieved when a search for patients similar to P_1 is conducted in the patient database. However, knowing that

250.01 (Diabetes mellitus type I) is related to 250.02 (Diabetes mellitus type II) and 401.1 (Benign hypertension) is related to 401.9 (Hypertension NOS) it becomes obvious that the similarity between the two patients is quite high. The problem consists of identifying a method to compute this similarity.

Consider two patients P_1 and P_2 described by ICD-9 diagnosis terms as: $P_1=\{T_{11},\ldots,T_{1N}\}$ and $P_2=\{T_{21}, \ldots,T_{2M}\}$. In [4] we investigated various ontology enhanced similarity measures for objects described by Gene Ontology terms. In this work we use the following measure [4] for computing the similarity s_a between P_1 and P_2:

$$s_a(P_1,P_2) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} s_p(T_{1i},T_{2j})}{NM}, \qquad (1)$$

where $s_p(T_{1i}, T_{2j})$ is the pair-wise similarity between diagnosis T_{1i} and diagnosis T_{2j} computed in the ICD-9 ontology. In other words, s_a is an average of the pair-wise similarity between the diagnoses that describe the two patients P_1 and P_2. To make $s_a(P_1,P_1)$ (self-similarity) equal 1 we need the normalize the above similarity as:

$$s(P_1,P_2) = \frac{s_a(P_1,P_2)}{\max\{s_a(P_1,P_1),s_a(P_2,P_2)\}} \qquad (2)$$

Calculating the above similarity requires first the computation of pair-wise diagnosis similarities $s_p(T_{1i}, T_{2j})$. There are numerous ways of computing similarity (distance) between ontology terms [6]. The main concept behind computing the similarity between two patients described by terms from an ontology, is the information content of the diagnoses. As a consequence, two patients are more similar if they share the same rare (in the database) diagnosis than if they both carry the identical, but common diagnosis, for example hypertension. The information content of a diagnosis can be computed based on the diagnosis frequency in the hospital database or based on the depth in the ontology (deeper terms are more specific). Since we do not have access to the whole hospital database to compute the frequency of diagnoses, we have adopted a depth (level) based approach in this paper.

The information content (IC), of a term in the ICD-9 taxonomy (see figure 1) is computed as IC=1-1/n where n={1,2,3,4,5} is the hierarchical level of the term. For example (see figure 1) the IC of diagnosis code 241 (Goiter) is IC=1-1/4=0.75.

Now, returning to the problem from the beginning of this paragraph, the similarity of two

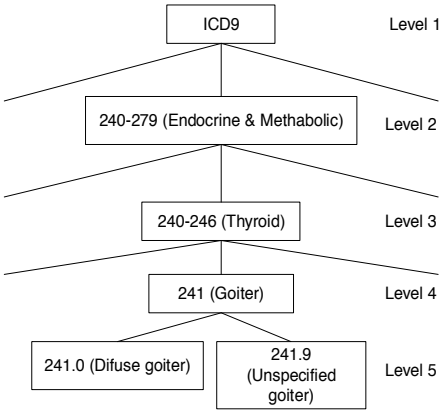

Figure 1. Partial view of the ICD9 taxonomy.

diagnosis terms, $s_p(T_{1i}, T_{2j})$, is defined as [7]:

$$s_p(T_{1i}, T_{2j})=IC\{NCA(T_{1i}, T_{2j})\}. \qquad (3)$$

where NCA is the nearest common ancestor of the two terms. In figure 1, the NCA of 241.0 and 241.9 is 241, hence the similarity between the two diagnoses is $s_p(241.0, 214.9)=0.75$.

Revisiting example 1, we can now compute $s(P_1,P_2)$ as:

$$s(P_1,P_2) = \frac{\frac{0.75+0+0.75+0}{4}}{\max\{\frac{1+0+1+0}{4},\frac{1+0+1+0}{4}\}} = 0.75$$

since, for example, $s_p(250.01, 250.02)=0.75$ (NCA at level 4), $s_p(250.01, 250.01)=1$ (NCA at level 5) and $s_p(250.01, 401.9)=0$ (NCA level 1).

Assume now that we are interested in the distribution of prescribed medications for a set of N patients, $\{P_i\}_{i=1,N}$ with a disease similar to our query P_q. If we consider only one drug (D) per patient and the patients P_i as having identical diagnoses to P_q, then the percentage of patients that take the given drug is:

$$p_D = \frac{\#\{P_i \mid P_i \text{ takesD}, i=1,N\}}{N}100. \qquad (4)$$

The problem with this approach is that N is either too small (when all the diagnoses of P_i are identical to P_q) or too large (when only one diagnosis is identical between P_i and P_q). When we extend the above relation to account for the ontological similarity, relation (4) becomes:

$$p_D = \frac{\sum_{i=1}^{N} s(P_q,P_i)\delta(D,P_i)}{N}100, \qquad (5)$$

where $\delta(D,P_i)=1$ if the patient P_i gets the drug D and 0 otherwise. In other words, the distribution

gets weighted by the similarity between patients. In addition, the number of similar patients N can be controlled using a threshold on the similarity between patients, as we will describe next.

In this work we consider (unrealistically) that each patient takes only one drug. This assumption comes in part from the fact that the medication was artificially assigned to the patients in this pilot database (see next section for details). However, (5) can be easily extended for multiple drugs per patient.

Dataset

After obtaining IRB approval, we extracted a set of 2077 patients from the hospital patient record database. The only information extracted for each patient was two ICD-9 codes representing diagnoses assigned to the patient in our hospital database. Drugs were associated with patients by searching the web site www.medicinenet.com using the primary ICD-9 diagnosis. In this way 26 drugs were assigned to our 2077 patients (one drug per patient). The dataset contains 520 distinct cases (that is, distinct associations of the 2 diagnoses). About 40% of the 520 cases are unique (that is the combination of their diagnoses is unique) and only about 5% have a frequency of 10 or more. Obviously, the distribution will be altered in a real hospital database that is typically 1000-2000 times bigger than our dataset. However, since the number of diagnoses per patient is usually higher than in our dataset the above distribution might not be completely unrealistic. The point here is that for each case that has a unique combination of diagnoses our decision support system will not retrieve any similar case, hence there will be no support for a decision. A possible alternative would be to find matches for each of the diagnoses in the query patient. In this case, the number of retrieved cases will be huge and not necessarily related to the query case. The above cases are discussed in more details in the next section for the case of our data set.

Results

1. Ontological similarity calculation

To analyze the effect of the similarity on the patient retrieval we computed the patient similarity matrix using different similarity measures. In figure 2.a, we show the case where we consider two patients similar if they have an identical set of diagnoses. This case is equivalent to the database (intersection) query:

"Select P from patients where P.diagnosis1=Query.diagnosis1 and P.diagnosis2=Querry.diagnosis2".

In figure 2.b, we show the case where we consider two patients similar if they share one diagnosis. This case is equivalent to the database (union) query:

"Select P from patients where P.diagnosis1=Query.diagnosis1 or P.diagnosis2=Querry.diagnosis2 or P.diagnosis2=Query.diagnosis1 or P.diagnosis1=Querry.diagnosis2".

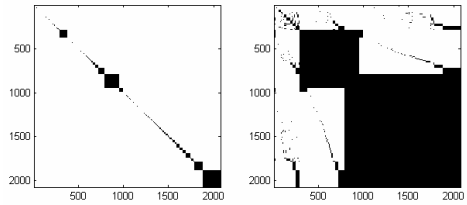

Figure 2. a.) Intersection similarity matrix b.)Union similarity matrix (black=similar)

While in the intersection case (figure 2.a) we retrieve a median number of 15 similar patients for each query, in the union case (figure 2.b) we retrieve a median number of 1296 patients per query. In the context of our support system one could use the medication of the similar patients to populate a drop-down list from which the physician could choose the one indicated for the query patient. In such an application, it is clear that 15 options are not enough and 1296 are too many. Moreover, there is no parameter that one can modify in order to obtain a reasonable number of items in the drop-down list. On the other hand, if the patient similarity were a continuous variable in [0,1], then one can choose a threshold that will retrieve a "reasonable" number of patients. Next we show how we can achieve this goal using the ontological query.

If we compute the patient similarity using formula (2), then the similarity matrix is as shown in figure 3.

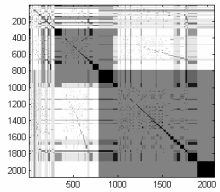

Figure 3. The patient similarity matrix for the ontological similarity (black=similar)

As we can see, the similarity is much less binary than the similarities based on the pure database searches (figure 2.a and 2.b).

Moreover, the median number of patients retrieved can be controlled using a similarity threshold (see figure 4). If one considers all the patients that are more than 80% similar with the query patient, then the median number of the similar patients retrieved is 25. This is about 60% more than in the case of the intersection query (identical patients) but only about 5% of the patients retrieved by the union query.

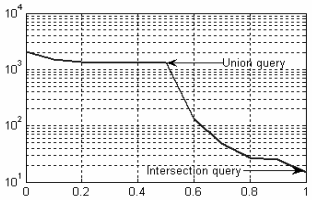

Figure 4. Variation of the median number of patients retrieved by an ontological query versus the similarity threshold.

It is interesting to note in figure 4 that the ontological query can retrieve more patients than the union query (the maximum that can be retrieved using a regular database query strategy). For example, at a similarity threshold of 0.1, the query based on the ontological similarity retrieves about 18% more patients. This is due to the ability of the ontology to relate different diagnoses that retain some degree of similarity in the taxonomy. This ontological effect will have an impact on the cases that are unique in the database (no exact match can be found). Using an ontological query one will be able to find cases that can be "close enough" in order to have some support for the required decision. As a consequence, the drop-down list with related medication will not be empty.

2. Ontological similarity validation

To validate the ontological similarity introduced in the previous section, three physicians (a pathologist, a pediatrician and an ENT physician) were asked to assess the similarity for 100 patient pairs chosen at random from the 2077 patient data set. Each physician was asked to score the similarity between the two patients based on their diagnoses as displayed on the screen.

The plot of the median of the three physician similarities versus the ontological similarity for the 100 patient pairs (fewer points are visible due to identity) is shown in figure 5. The overall coefficient of correlation between the median physician similarity and the ontological similarity is 0.88, which shows a very good correlation between the two similarities.

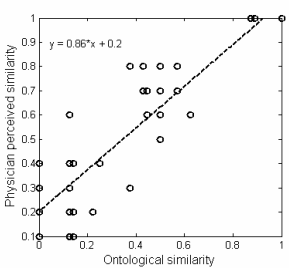

Figure 5. The plot of the median of the physician perceived patient similarity and ontological similarity.

In addition, in figure 5, the linear regression for the points representing the similarity pairs has a slope of 0.86 (close to the desired coefficient of 1) and an intercept of 0.2. The intercept shows that the physician perception of similarity is higher than that given by the ontological approach. It follows that the ontological similarity is more conservative than the physician perceived similarity. One reason for this result is that in our approach we ignored the significance of the diagnoses. For example, for two patients P_1={diabetes, cough} and P_2={diabetes, leg pain}, the ontological similarity is 0.5 while the physician assigned similarity was 0.8. In this case, the physician considered that the more serious diagnosis (diabetes) makes the patients more similar than 0.5. In future work we plan to assign an importance coefficient to each disease and incorporate it in our similarity measure. Another reason for the differences in similarity could be that the ICD-9 hierarchy is not granular enough to capture the many subtleties of medical knowledge. In future work we plan to use larger ontologies such as Snomed or UMLS to compute the ontological similarity.

3. OntoQuest prototype

A prototype interface for a medication decision support system called OntoQuest is shown in figure 6. The patient "John Doe1996" is found to be similar to 11 other patients in the 2077 dataset at a similarity threshold of about 0.8. The medications found for all 11 patients are displayed in the "Medication" list box. The percentage of the patients that take each drug listed is shown in the pie chart shown below the medication list-box (equation (5)). At this point, the physician can choose the appropriate medication for "John Doe1996" from the "Medication" list box. The approach may increase the confidence of junior physicians in prescribing medication. Moreover, by having a

list of medications commonly prescribed in similar cases this approach has the potential for limiting the inadvertent prescription of drugs inappropriate for the disease.

In the example from figure 6, the similarity is based on the fact (inferred from the ICD-9 hierarchy) that the similarity between diagnoses with ICD9 codes {786.5-Chest pain, 786.2-Cough, 786.05-Shortness of breath} is 0.75 (level 4 similarity).

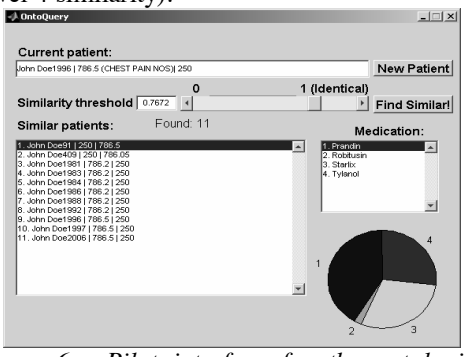

Figure 6. Pilot interface for the ontological query system for a 0.77 patient neighborhood.

On the other hand, only 4 patients are retrieved for the same patient if a threshold of 1 (identity) is specified. By changing the similarity threshold, the number of the items in the medication list box can be adapted to the rarity of the case. This adaptive ("intelligent") behavior of the medication list box allows the physician to avoid looking for the right combination of drugs by performing extensive searches.

Conclusions

In this paper we introduced an ontological query method for computing the similarity between two patients based on their ICD-9 diagnosis sets and found that this correlates very well with the physician perceived patient similarity.

We used the ontological similarity for building a prototype of a medication support system, OntoQuest. The design goals of OntoQuest are the reduction of medication prescription errors and the facilitation of physician interaction with computerized physician order entry (CPOE) systems.

As future work, before testing OntoQuest in a clinical setting, we plan to address several issues: first, the development of a method of assigning the relative importance of diagnoses; second, the implementation of the Snomed and UMLS controlled terminology databases for computing patient similarity; third, the need for the inclusion of demographic variables in the computation of patient similarity which is necessary in order to differentiate, for example, between a pediatric and a geriatric patient. fourth: the necessity of using a more realistic data set for preliminary testing where each patient has a larger and more realistic number of drugs and diagnoses assigned.

Another potential research direction is to extend the ontological query methodology to a diagnosis support system. This system would retrieve the diagnoses existent in the hospital database for patients "similar" to the query patient based on their specific set of signs, symptoms, laboratory data and other findings.

Acknowledgement

The work was supported by NLM grant 2-T15-LM07089-14. The authors thank to Patricia Alafaireet for providing the patient data.

References

[1] Andreasen T., Bulskov H., Knappe R., "On ontology-based querying", pp. 53-59 in Heiner Stuckenschmidt (Eds.): 18th International Joint Conference on Artificial Intelligence, Acapulco, Mexico, August 9, 2003.

[2] Hartmann J., Stojanovic N., Studer R., Schmidt-Thieme L., "Ontology-Based Query Refinement for Semantic Portals", From Integrated Publication and Information Systems to Virtual Information and Knowledge Environments 2005, 41-50.

[3] Lord P.W., Stevens R.D., Brass A., Goble C.A., "Semantic similarity measure as a tool for exploring the gene ontology",Pacific Symposium on Biocomputing, pp. 601-612, 2003.

[4] Popescu M., Keller J.M., Mitchell J.A., "Fuzzy Measures on the Gene Ontology for Gene Product Similarity", IEEE Trans. Computational Biology and Bioinformatics, accepted for publication 2005.

[5] Buchanan B.G., Shortliffe E.H. (eds.) Rule-Based Expert Systems: the MYCIN experiments of the Stanford Heuristic Programming Project, Addison-Wesley, Reading MA, (1984).

[6] Jiang J.J., Conrath D.W., "Semantic Similarity Based on Corpus Statistics and Lexical Ontology", Proc. of Int. Conf. Research on Comp. Linguistics X, 1997, Taiwan.

[7] Resnick P., "Semantic similarity in a taxonomy: an information-base measure and its application to problems of ambiguity in natural language", Journal of Artificial Intelligence Research (JAIR), 11, pp. 95-130, 1999.