Using Prototypes and Adaptation Rules for Diagnosis of Dysmorphic Syndromes

Rainer Schmidt and Tina Waligora

Institute for Medical Informatics and Biometry, University of Rostock, Germany
rainer.schmidt@medizin.uni-rostock.de

Abstract. Since diagnosis of dysmorphic syndromes is a domain with incomplete knowledge and where even experts have seen only few syndromes themselves during their lifetime, documentation of cases and the use of case-oriented techniques are popular. In dysmorphic systems, diagnosis usually is performed as a classification task, where a prototypicality measure is applied to determine the most probable syndrome. These measures differ from the usual Case-Based Reasoning similarity measures, because here cases and syndromes are not represented as attribute value pairs but as long lists of symptoms, and because query cases are not compared with cases but with prototypes. In contrast to these dysmorphic systems our approach additionally applies adaptation rules. These rules do not only consider single symptoms but combinations of them, which indicate high or low probabilities of specific syndromes.

1 Introduction

When a child is born with dysmorphic features or with multiple congenital malformations or if mental retardation is observed at a later stage, finding the correct diagnosis is extremely important. Knowledge of the nature and the etiology of the disease enables the pediatrician to predict the patient's future course. So, an initial goal for medical specialists is to diagnose a patient to a recognised syndrome. Genetic counselling and a course of treatments may then be established.

A dysmorphic syndrome describes a morphological disorder and it is characterised by a combination of various symptoms, which form a pattern of morphologic defects. An example is Down Syndrome which can be described in terms of characteristic clinical and radiographic manifestations such as mental retardation, sloping forehead, a flat nose, short broad hands and generally dwarfed physique [1].

The main problems of diagnosing dysmorphic syndromes are as follows [2]:

- more than 200 syndromes are known,
- many cases remain undiagnosed with respect to known syndromes,
- usually many symptoms are used to describe a case (between 40 and 130),
- every dysmorphic syndrome is characterised by nearly as many symptoms.

Furthermore, knowledge about dysmorphic disorders is continuously modified, new cases are observed that cannot be diagnosed (it exists even a journal that only publishes reports of observed interesting cases [3]), and sometimes even new

P. Perner (Ed.): ICDM 2006, LNAI 4065, pp. 1 – 9, 2006.

syndromes are discovered. Usually, even experts of paediatric genetics only see a small count of dysmorphic syndromes during their lifetime.

So, we have developed a diagnostic system that uses a large case base. Starting point to build the case base was a large case collection of the paediatric genetics of the University of Munich, which consists of nearly 2000 cases and 229 prototypes. A prototype (prototypical case) represents a dysmorphic syndrome by its typical symptoms. Most of the dysmorphic syndromes are already known and have been defined in the literature. And nearly one third of our entire case base has been determined by semiautomatic knowledge acquisition, where an expert selected cases that should belong to same syndrome and subsequently a prototype, characterised by the most frequent symptoms of his cases, was generated. To this database we have added cases from "clinical dysmorphology" [3] and syndromes from the London dysmorphic database [4], which contains only rare dysmorphic syndromes.

1.1 Diagnostic Systems for Dysmorphic Syndromes

Systems to support diagnosis of dysmorphic syndromes have already been developed in the early 80's. The simple ones perform just information retrieval for rare syndromes, namely the London dysmorphic database [3], where syndromes are described by symptoms, and the Australian POSSUM, where syndromes are visualised [5]. Diagnosis by classification is done in a system developed by Wiener and Anneren [6]. They use more than 200 syndromes as database and apply Bayesian probability to determine the most probable syndromes. Another diagnostic system, which uses data from the London dysmorphic database was developed by Evans [7]. Though he claims to apply Case-Based Reasoning, in fact it is again just a classification, this time performed by Tversky's measure of dissimilarity [8]. The most interesting aspect of his approach is the use of weights for the symptoms. That means the symptoms are categorised in three groups – independently from the specific syndromes, instead only according to their intensity of expressing retardation or malformation. However, Evans admits that even features, that are usually unimportant or occur in very many syndromes sometimes play a vital role for discrimination between specific syndromes.

In our system the user can chose between two measures of dissimilarity between concepts, namely of Tversky [8] and the other one of Rosch and Mervis [9]. However, the novelty of our approach is that we do not only perform classification but subsequently apply adaptation rules. These rules do not only consider single symptoms but specific combinations of them, which indicate high or low probabilities of specific syndromes.

1.2 Case-Based Reasoning and Prototypicality Measures

Since the idea of Case-Based Reasoning (CBR) is to use former, already solved solutions (represented in form of cases) for current problems [10], CBR seems to be appropriate for diagnosis of dysmorphic syndromes. CBR consists of two main tasks [11], namely retrieval, which means searching for similar cases, and adaptation, which means adapting solutions of similar cases to the query case. For retrieval usually explicit similarity measure or, especially for large case bases, faster retrieval

algorithms like Nearest Neighbour Matching [12] are applied. For adaptation only few general techniques exist [13], usually domain specific adaptation rules have to be acquired.

In CBR usually cases are represented as attribute-value pairs. In medicine, especially in diagnostic applications, this is not always the case, instead often a list of symptoms describes a patient's disease. Sometimes these lists can be very long, and often their lengths are not fixed but vary with the patient. For dysmorphic syndromes usually between 40 and 130 symptoms are used to characterise a patient.

Furthermore, for dysmorphic syndromes it is unreasonable to search for single similar patients (and of course none of the systems mentioned above does so) but for more general prototypes that contain the typical features of a syndrome. Prototypes are a generalisation from single cases. They fill the knowledge gap between the specificity of single cases and abstract knowledge in the form of cases. Though the use of prototypes had been early introduced in the CBR community [14, 15], their use is still rather seldom. However, since doctors reason with typical cases anyway, in medical CBR systems prototypes are a rather common knowledge form (e.g. for antibiotics therapy advice in ICONS [16], for diabetes [17], and for eating disorders [18]).

So, to determine the most similar prototype for a given query patient instead of a similarity measure a prototypicality measure is required. One speciality is that for prototypes the list of symptoms is usually much shorter than for single cases.

The result should not be just the one and only most similar prototype, but a list of them – sorted according to their similarity. So, the usual CBR methods like indexing or nearest neighbour search are inappropriate. Instead, rather old measures for dissimilarities between concepts [8, 9] are applied and explained in the next section.

2 Diagnosis of Dysmorphic Syndromes

Our system consists of four steps (fig.1). At first the user has to select the symptoms that characterise a new patient. This selection is a long and very time consuming process, because we consider more than 800 symptoms. However, diagnosis of dysmorphic syndromes is not a task where the result is very urgent, but it usually requires thorough reasoning and afterwards a long-term therapy has to be started. Since our system is still in the evaluation phase, secondly the user can select a prototypicality measure. In routine use, this step shall be dropped and instead the measure with best evaluation results shall be used automatically. At present there are three choices. As humans look upon cases as more typical for a query case as more features they have in common [9], distances between prototypes and cases usually mainly consider the shared features.

The first, rather simple measure (1) just counts the number of matching symptoms of the query patient (X) and a prototype (Y) and normalises the result by dividing it by the number of symptoms characterising the syndrome.

This normalisation is done, because the lengths of the lists of symptoms of the various prototypes vary very much. It is performed by the two other measures too.

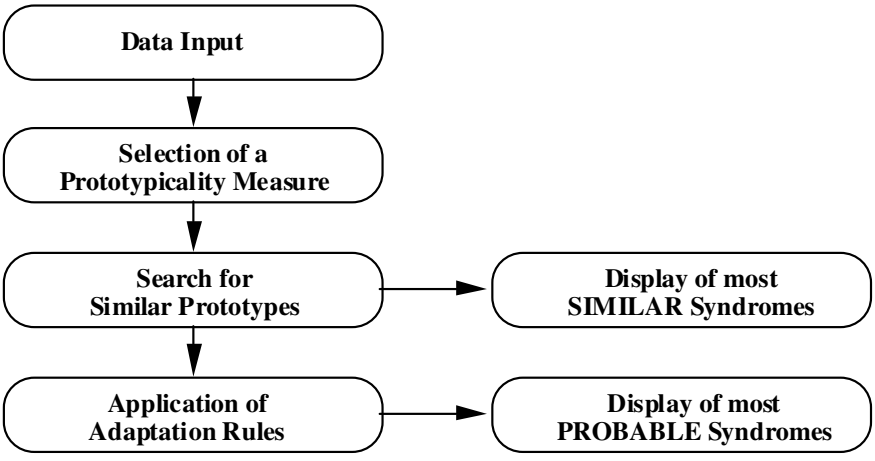

Fig. 1. Steps to diagnose dysmorphic syndromes

The following equations are general (as they were originally proposed) at the point that a general function "f" is used, which usually means a sum that can be weighted. In general these functions "f" can be weighted differently. However, since we do not use any weights at all, in our application "f" means simply a sum.

$$D(X,Y) = \frac{f(X+Y)}{f(Y)} \tag{1}$$

The second measure (2) was developed by Tversky [8]. It is a measure of dissimilarity for concepts. In contrast to the first measure, additionally two numbers are subtracted from the number of matching symptoms. Firstly, the number of symptoms that are observed for the patient but are not used to characterise the prototype (X-Y), and secondly the number of symptoms used for the prototype but are not observed for the patient (Y-X) is subtracted.

$$D(X,Y) = \frac{f(X+Y) - f(X-Y) - f(Y-X)}{f(Y)} \tag{2}$$

The third prototypicality measure (3) was proposed by Rosch and Mervis [9]. It differs from Tversky's measure only in one point: the factor X-Y is not considered:

$$D(X,Y) = \frac{f(X+Y) - f(Y-X)}{f(Y)} \tag{3}$$

In the third step to diagnose dysmorphoic syndromes, the chosen measure is sequentially applied on all prototypes (syndromes). Since the syndrome with maximal

Table 1. Most similar prototypes after applying a prototypicality measure

Most Similar Syndromes	**Similarity**
Shprintzen-Syndrome	0.49
Lenz-Syndrome	0.36
Boerjeson-Forssman-Lehman-Syndrome	0.34
Stuerge-Weber-Syndrome	0.32

similarity is not always the right diagnosis, the 20 syndromes with best similarities are listed in a menu (table 1).

2.1 Application of Adaptation Rules

In the fourth and final step, the user can optionally choose to apply adaptation rules on the syndromes. These rules state that specific combinations of symptoms favour or disfavour specific dysmorphic syndromes. Unfortunately, the acquisition of these adaptation rules is very difficult, because they cannot be found in textbooks but have to be defined by experts of paediatric genetics. So far, we have got only 10 of them and so far, it is not possible that a syndrome can be favoured by one adaptation rule and disfavoured by another one at the same time. When we, hopefully, acquire more rules, such a situation should in principle be possible but would indicate some sort of inconsistency of the rule set.

How shall the adaptation rules alter the results? Our first idea was that the adaptation rules should increase or decrease the similarity scores for favoured and disfavoured syndromes. But the question is how. Of course no medical expert can determine values to manipulate the similarities by adaptation rules and any general value for favoured or disfavoured syndromes would be arbitrary.

So, instead the result after applying adaptation rules is a menu that contains up to three lists (table 2).

On top the favoured syndromes are depicted, then those neither favoured nor disfavoured, and at the bottom the disfavoured ones. Additionally, the user can get information about the specific rules that have been applied on a particular syndrome (e.g. fig. 2).

Table 2. Most similar prototypes after additionally applying adaptation rules

Probable prototypes after application of adaptation rules	**Similarity**	**Applied Rules**
Lenz-Syndrome	0.36	Rule-No.6
Dubowitz-Syndrom	0.24	Rule-No.9
Prototypes, no adaptation rules could be applied		
Shprintzen-Syndrome	0.49	
Boerjeson-Forssman-Lehman-Syndrome	0.34	
Stuerge-Weber-Syndrome	0.32	
Leopard-Syndrome	0.31	

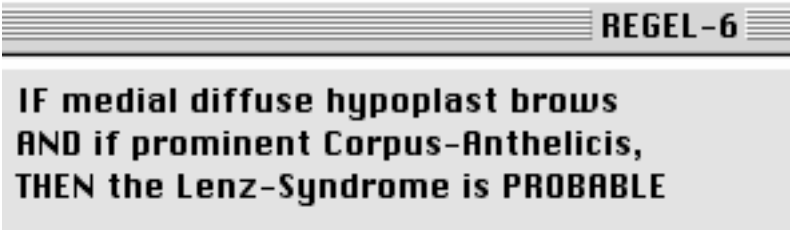

Fig. 2. Presented information about the applied adaptation rule

In the example presented by tables 1 and 2, and figure 2 the correct diagnosis is Lenz-syndrome. The computation of the prototypicality measure of Rosch and Mervis determines Lenz-syndrome as the most similar but one syndrome (here Tversky's measure provides a similar result, only the differences between the similarities are smaller). After application of adaptation rules, the ranking is not obvious. Two syndromes have been favoured, the more similar one is the right one. However, Dubowitz-syndrome is favoured too (by a completely different rule), because a specific combination of symptoms makes it probable, while other observed symptoms indicate a rather low similarity.

3 Results

Cases are difficult to diagnose when patients suffer from a very rare dysmorphic syndrome for which neither detailed information can be found in literature nor many cases are stored in our case base. This makes evaluation difficult. If test cases are randomly chosen, frequently observed cases resp. syndromes are frequently selected and the results will probably be fine, because these syndromes are well-known. However, the main idea of the system is to support diagnosis of rare syndromes. So, we have chosen our test cases randomly but under the condition that every syndrome can be chosen only once.

For 100 cases we have compared the results obtained by both prototypicality measures (table 3).

Table 3. Comparison of prototypicality measures

Right Syndrome	Rosch and Mervis	Tversky
on Top	29	40
among top 3	57	57
among top 10	76	69

The results may seem to be rather poor. However, diagnosis of dysmorphic syndromes is very difficult and usually needs further investigation, because often a couple of syndromes are very similar. The first step is to provide the doctor with

information about probable syndromes, so that he gets an idea about which further investigations are appropriate. That means, the right diagnose among the three most probable syndromes is already a good result.

Obviously, the measure of Tversky provides better results, especially when the right syndrome should be on top of the list of probable syndromes. When it should be only among the first three of this list, both measures provide equal results.

Adaptation rules. Since the acquisition of adaptation rules is a very difficult and time consuming process, the number of acquired rules is rather limited, namely at first just 10 rules. Furthermore, again holds: the better a syndrome is known, the easier adaptation rules can be generated. So, the improvement mainly depends on the question how many syndromes involved by adaptation rules are among the test set. In our experiment this was the case only for 5 syndromes. Since some had been already diagnosed correctly without adaptation, there was just a small improvement (table 4).

Table 4. Results after applying adaptation rules

Right Syndrome	Rosch and Mervis	Tversky
on Top	32	42
among top 3	59	59
among top 10	77	71

Some more adaptation rules. Later on we acquired eight further adaptation rules and repeated the tests with the same test cases. The new adaptation rules again improved the results (table 5).

Table 5. Results after applying some more adaptation rules

Right Syndrome	Rosch and Mervis	Tversky
on Top	36	44
among top 3	65	64
among top 10	77	73

It is obvious that with the number of acquired adaptation rules the quality of the program increases too. Unfortunately, the acquisition of these rules is very difficult and especially for very rare syndromes probably nearly impossible.

4 Conclusion

Diagnosis of dysmorphic syndromes is a very difficult task, because many syndromes exist, the syndromes can be described by various symptoms, many rare syndromes are still not well investigated, and from time to time new syndromes are discovered.

We have compared two prototypicality measures, where the one by Tversky provides slightly better results. Since the results were rather pure, we additionally have applied adaptation rules (as we have done before, namely for the prognosis of influenza [19]). We have shown that these rules can improve the results. Unfortunately, the acquisition of them is very difficult and time consuming. Furthermore, the main problem is to diagnose rare and not well investigated syndromes and for such syndromes it is nearly impossible to acquire adaptation rules.

However, since adaptation rules do not only favour specific syndromes but can be used to disfavour specific syndromes, the chance to diagnose even rare syndromes also increases by the count of disfavouring rules for well-known syndromes. So, the best way to improve the results seems to be to acquire more adaptation rules, however difficult this task may be.

References

1. Taybi, H., Lachman, R.S.: Radiology of Syndromes, Metabolic Disorders, and Skeletal Dysplasia. Year Book Medical Publishers, Chicago (1990)
2. Gierl, L., Stengel-Rutkowski, S.: Integrating Consultation and Semi-automatic Knowledge Acquisition in a Prototype-based Architecture: Experiences with Dysmorphic Syndromes. Artificial Intelligence in Medicine 6 (1994) 29-49
3. Clinical Dysmorphology. http://www.clyndysmorphol.com (last accessed: April 2006)
4. Winter R.M., Baraitser M., Douglas J.M.: A computerised data base for the diagnosis of rare dysmorphic syndromes. Journal of medical genetics 21 (2) (1984) 121-123
5. Stromme P.: The diagnosis of syndromes by use of a dysmorphology database. Acta Paeditr Scand 80 (1) (1991) 106-109
6. Weiner F., Anneren G.: PC-based system for classifying dysmorphic syndromes in children. Computer Methods and Programs in Biomedicine 28 (1989) 111-117
7. Evans C.D.: A case-based assistant for diagnosis and analysis of dysmorphic syndromes. International Journal of Medical Informatics 20 (1995) 121-131
8. Tversky, A.: Features of Similarity. Psychological Review 84 (4) (1977) 327-352
9. Rosch E., Mervis C.B.: Family Resemblance: Studies in the Internal Structures of Categories. Cognitive Psychology 7 (1975) 573-605
10. Kolodner, J.: Case-Based Reasoning. Morgan Kaufmann Publishers, San Mateo (1993)
11. Aamodt, A., Plaza, E.: Case-Based Reasoning: Foundation issues, methodological variation, and system approaches. AICOM 7 (1994) 39-59
12. Broder, A.: Strategies for efficient incremental nearest neighbor search. Pattern Recognition 23 (1990) 171-178
13. Wilke, W., Smyth, B., Cunningham, P.: Using configuration techniques for adaptation. In: Lenz, M. et al. (eds.): Case-Based Reasoning technology, from foundations to applications. Lecture Notes in Artificial Intelligence, Vol. 1400, Springer-Verlag, Berlin Heidelberg New York (1998) 139-168
14. Schank, R.C.: Dynamic Memory: a theory of learning in computer and people. Cambridge University Press, New York (1982)
15. Bareiss, R.: Exemplar-based knowledge acquisition. Academic Press, San Diego (1989)
16. Schmidt, R., Gierl, L.: Case-based Reasoning for antibiotics therapy advice: an investigation of retrieval algorithms and prototypes. Artificial Intelligence in Medicine 23 (2001) 171-186

17. Bellazzi, R., Montani, S., Portinale, L.: Retrieval in a prototype-based case library: a case study in diabetes therapy revision. In: Smyth, B., Cunningham, P. (eds.): Proc European Workshop on Case-Based Reasoning. Lecture Notes in Artificial Intelligence, Vol. 1488, Springer-Verlag, Berlin Heidelberg New York (1998) 64-75
18. Bichindaritz, I.: From cases to classes: focusing on abstraction in case-based reasoning. In: Burkhard, H.-D., Lenz, M.: (eds.): Proc German Workshop on Case-Based Reasoning, University Press, Berlin (1996) 62-69
19. Schmidt, R., Gierl, L.: Temporal Abstractions and Case-based Reasoning for Medical Course Data: Two Prognostic Applications. In: Perner P (eds.): Machine Learning and Data Mining in Pattern Recognition, MLDM 2001. Lecture Notes in Computer Science, Vol. 2123, Springer-Verlag, Berlin Heidelberg New York (2001) 23-34