Diagnosis of Dysmorphic Syndromes Using Prototypes and Adaptation Rules

Tina Waligora, Rainer Schmidt
Institute for Medical Informatics and Biometry
University of Rostock
Rembrandtstr. 16/17, D-18055 Rostock, Germany
[tina.waligora / 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. Our system 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 paediatrician 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. It is characterised by a combination of various symptoms, which form a pattern of morphologic defects. The main problems of diagnosing dysmorphic syndromes are as follows [Gierl *et al*., 1994]:

- existence of more than 200 syndromes,
- 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, and sometimes even new syndromes are discovered.

We have developed a diagnostic system that uses a large case base. Starting point to build-up the case base was a large case collection of the paediatric genetics of the University of Munich, which consists of nearly 2,000 cases and 229 prototypes. A prototype (prototypical case) represents a dysmorphic syndrome by its typical symptoms. Many dysmorphic syndromes have been defined in literature. Additionally, nearly one third of our case base was determined by semiautomatic knowledge acquisition, where an expert selects cases that should belong to the same syndrome and subsequently a prototype, characterised by the most frequent symptoms of it's cases, is generated.

In our system the user can choose between two measures of dissimilarity between concepts, namely one measure proposed by Tversky [Tversky, 1977], the other one by Rosch and Mervis [Rosch *et al*., 1975]. 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.

2 Diagnosis of Dysmorphic Syndromes

Our system performs four steps. At first the user has to select symptoms that describe a new patient. This selection is strenuous and time consuming, because more than 800 symptoms are considered. However, diagnosis of dysmorphic syndromes is not a task requiring great speed, but it usually requires thorough reasoning and is followed by a long-term therapy. Since our system is still in the evaluation phase, the user can select a prototypicality measure. In routine use, this step shall be dropped and instead the measure with better evaluation results shall be used automatically. There are two choices.

As humans look upon cases as more typical for a query case with increasing numbers of common features [Rosch *et al*., 1975], distances between prototypes and cases usually mainly consider the shared features. The first measure was developed by Tversky [Tversky, 1977]. It is a measure of dissimilarity of concepts. From the number of features shared by the query case and the prototype two numbers are subtracted. 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.

$$D_{Tversky}(X,Y) = \frac{f(X + Y) - f(X - Y) - f(Y - X)}{f(Y)}$$

The second prototypicality measure was proposed by Rosch and Mervis [Rosch *et al*., 1975]. It differs from Tversky's measure only in one point: the factor X-Y is not considered:

$$D_{Rosch,\ Mervis}(X,Y) = \frac{f(X + Y) - f(Y - X)}{f(Y)}$$

In the third step to diagnosis dysmorphic syndromes, the chosen measure is sequentially applied on all prototypes (syndromes). Since the syndrome with maximal similarity is not always the right diagnosis, the 20 syndromes with highest similarity are presented ranked according similarity.

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. For example, this is an adaptation rule favouring Lenz-Syndrome:

```
IF medial diffuse hypoplast brows
AND IF prominent Corpus-Anthelicis
THEN the Lenz-Syndrome is probable
```

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.

The question is how shall adaptation rules alter the results. Our first idea was that the similarity values should be changed. A syndrome that is favoured by an adaptation rule might get a higher similarity. But we had no idea how much an adaptation rule shall increase a similarity. Of course no medical expert can help here and a general value for favoured or disfavoured syndromes by adaptation rules would be arbitrary. So, instead the result after applying adaptation rules is a menu that contains up to three lists. 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.

3 Results

Cases are difficult to diagnose when patients suffer from a very rare dymorphic 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 syndromes will be frequently selected and the results will probably be fine, because these syndromes are well-known. However, the main idea of our 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, before and after applying adaptation rules (table 1).

Table 1. Comparison of prototypicality measures

			With Adaptation	With Adaptation
Right Syndrome	Rosch, Mervis	Tversky	Rosch, Mervis	Tversky
on Top	29	40	32	42
among top 3	57	57	59	59
among top 10	76	69	77	71

Obviously, the measure of Tversky provides just very slightly better results, especially when the right syndrome should be on top of the list of probable syndromes. Since the acquisition of adaptation rules is very difficult and time consuming, the number of acquired rules is rather limited, namely 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 with five syndromes. Since some of them had already been diagnosed correctly without adaptation, the improvement by adaptation rules is very small.

References

[Gierl *et al*., 1994] Lothar Gierl and Sabine Stengel-Rutkowski,. Integrating Consultation and Semi-automatic Knowledge Acquisition in a Prototype-based Architecture: Experiences with Dysmorphic Syndromes. *Artificial Intelligence in Medicine*, 6:29-49, 1994

[Rosch *et al*., 1975] Elenor Rosch and C.B. Mervis. Family Resemblance: Studies in the Internal Structures of Categories. *Cognitive Psychology* 7:573-605, 1975

[Tversky, 1977] A Tversky. Features of Similarity. *Psychological Review* 84:327-352, 1977