

# AN INTERPRETATION OF A MEDICAL SCORE TEST AS A BASIC PROBABILITY ASSIGNMENT

Ewa Straszecka<sup>\*</sup>, Joanna Straszecka<sup>\*\*</sup>

<sup>\*</sup>Div. of Biomedical Engineering, Inst. of Electronics, Silesian Technical University  
16 Akademicka St., 44-100 Gliwice, Poland  
phone (48)(32)2371725, fax (48)(32)2372225, e-mail ewa@biomed.iele.polsl.gliwice.pl

<sup>\*\*</sup>Department of Internal Diseases and Rheumatology  
Medical University of Silesia  
102 Edukacji St., 43-100 Tychy, Poland  
phone (48)(32)2276071-215, fax (48)(32)2277723

**ABSTRACT:** In the paper a method of replacing the algebraic sum in medical score tests by the basic probability assignment in Dempster-Shafer theory of evidence is proposed. The problem of diagnosis is divided into two sub-tasks: to evaluate chances for the disease and for health. A method of calculating basic probability assignment on original score as well as using frequency of a disease symptom occurrence is suggested. Conditions in which combining the assignments is desirable and a suitable formula are mentioned. The study is illustrated by an example of Crooks index used in hyperthyroidism diagnosis transformation.

## 1. INTRODUCTION

Many algorithms in medicine take a form of score tests in which a patient's parameter is evaluated by a number. An algebraic sum of such numbers indicates a diagnostic conclusion. A straightforward computerisation of the algorithms is disadvantageous for two major reasons:

- a rigid description of symptoms requires an answer "yes" or "no", neglecting the situation when symptoms are unknown;
- an attempt of supplementing the test with new parameters change the global sum in an unpredictable way.

It would be profitable to introduce a method of automatic evaluation of importance of a finding and of differentiation between absent and unknown symptoms. What is more a score test is designed basing on experience of physicians in a determined time and place. Though its essential sense remain unalterable, it often happens that in medical practice in various regions and other conditions some symptoms change their meaning. For the reasons the global algebraic sum which is the main diagnostic factor could be misleading. In the paper a method of replacing the algebraic sum by basic probability assignment in Dempster-Shafer theory of evidence (DST) is proposed. Combining score test knowledge with database information is described. Results of calculations for Crooks' index in hyperthyroidism diagnosis are presented.

## 2. SCORE TESTS IN MEDICINE

Researchers building expert systems appreciate algebraic sum as a simple tool which can solve many problems [4]. However, though a straightforward computerization of score tests is easy it is also disadvantageous for above-mentioned reasons. The simplification and speeding up of score tests and flexible computerization of medical diagnostic algorithms are the reasons for interest in DST as a tool for medical knowledge description.

There are two main types of medical score tests:

### Type I

*Lack of intensive physical exercise* + 3 points  
*Frequent stress situations at work* + 1 point to conclusion score  
*Systolic blood pressure within* + 4 points

[170, 179] mmHg

Generally:

**IF**  $X_i$  is  $A_i$  **THEN** +  $s_i$  to conclusion score

## Type II

Heart beating felt	+ 2 points	
pulse below 80/min	- 3 points	to conclusion score
pulse within [80, 90] /min	+ 0 point	
pulse above 90/min	+ 3 points	

Generally:

<b>IF</b> $X_i$ is $A_i$ <b>THEN</b>	+ $s_i$	<b>to conclusion score</b>
<b>IF</b> $X_j$ is $A_j$ <b>THEN</b>	- $s_j$	<b>from conclusion score</b>

The first example is taken from a heart infarct risk test, while the second is a part of Crooks index for hyperthyroidism. They differ in sign of points that are added. In the first type only positive points are added while in the second both positive and negative are comprised. In both tests a diagnosis depends on a global algebraic sum for all parameters. The tests can be transformed into the same form of rules, provided:

- 1) for negative points an opposing conclusion (i.e.  $\neg$  conclusion) is defined,
- 2) in reasoning, conclusions are considered as an alternative.

Thus a unification of symptoms formulation is satisfied and the following method can be used for each of the score tests types.

There are 22 parameters in Crooks' index. The score for an individual parameter is in [-5, 5] interval. Negative weights contradict while positive confirm hyperthyroidism. It is assumed that negative weights turn into positive if the conclusion is health (more precisely: euthyroidism). Therefore the parameters are divided for the disease ( $d$ ) and health ( $h$ ) symptoms denoted by  $\{s_{d1}, \dots, s_{dk}\}$  and  $\{s_{h1}, \dots, s_{hn}\}$ , respectively. In terms of DST symptoms are focal elements and conclusions are hypotheses which are linked with focal elements by the basic probability assignment. Two more focal elements are defined that represent unknown symptoms (if there are any in a diagnostic process). They are  $e_d$  and  $e_h$ . Thus links symptoms  $\Rightarrow$  diagnosis are as follows:

$$S_d = \{s_{d1}, \dots, s_{dk}, e_d\} \Rightarrow d$$

$$S_h = \{s_{h1}, \dots, s_{hn}, e_h\} \Rightarrow h$$

In this way the problem of the diagnosis is divided into two sub-tasks: to evaluate chances for the disease and for health. As often some symptoms are unknown, two factors of imprecision will be necessary to estimate belief of a hypothesis and its certainty. The Bel factor is defined in traditional way [1] while Cert =  $1 - m(\emptyset)$  [2], [3].

### 3. BASIC PROBABILITY ASSIGNMENT DETERMINATION

The most important element of DST is the basic probability assignment (bpa). The assumptions for bpa are relatively weak, so it is quite easy to normalize Crooks' index score to fulfill them. Let's denote bpa for hyperthyroidism  $m_d$  and for euthyroidism  $m_h$ . It is enough to divide original score by the maximum possible score to satisfy  $\sum_{s \in S_d} m_d(s) = 1$ ;  $\sum_{s \in S_h} m_h(s) = 1$ . Such an innovation brings only one profit to the original algorithm of

diagnosis – it is possible to define Cert of a hypothesis. The Bel factor directly resembles the algebraic sum. However, if we define another bpa for the same symptoms it will be possible to combine knowledge from the two sources of information. The other source could be a database.

It sometimes happens that occurrence of a disease is different in various regions of the country. Such a situation is observed for thyroid gland diseases. It would be justified to introduce a bpa which resembles the differences. It could be obtained for a database. Unfortunately the database will have to be small (as large-scale investigations in individual regions are not possible) and often changing (as the occurrence of the disease is connected with changing living conditions). However there are not obstacles against constructing a bpa on it. Therefore we propose to define a bpa relevant to frequency of symptoms occurrence as an interpretation of symptom's importance:

$$m_{d2}(s_{di}) = \begin{cases} \frac{P(s_{di}/d) - P(\neg s_{di}/d)}{\sum_{(s_{di} \Rightarrow d)=t} P(s_{di}/d) - P(\neg s_{di}/d)} & \text{if } P(s_{di}/d) - P(\neg s_{di}/d) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$m_{h2}(s_{hi}) = \begin{cases} \frac{P(s_{hi}/h) - P(\neg s_{hi}/h)}{\sum_{(s_{hi} \Rightarrow h)=t} P(s_{hi}/h) - P(\neg s_{hi}/h)} & \text{if } P(s_{hi}/h) - P(\neg s_{hi}/h) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Bpa defined on original Crooks' score ( $m_{d1}$ ) and calculated for a small database of 83 patients ( $m_{d2}$ ) are presented in table 1.

Table 1.

Symptom	$m_{d1}$	$m_{d2}$	$m_{d1}-m_{d2}$	$m_{h1}$	$m_{h2}$	$m_{h1}-m_{h2}$
1.	.024	.072	-.047			
2.	.049	.073	-.024			
3.	.049	.074	-.025			
4.				.208	.060	.149
5.	.122	.071	.051			
6.	.073	.069	.004			
7.	.073	.036	.037			
8.				.125	.060	.065
9.				.125	.060	.065
10.	.073	.068	.005			
11.	.049	.074	-.025			
12.	.073	.069	.004	.125	.071	.054
13.	.049	.023	.026	.083	.167	-.083
14.	.049	.015	.033	0	0	0
15.	.024	.033	-.008	0	0	0
16.	.098	.058	.039	.083	.155	-.071
17.	.024	.072	-.047	0	0	0
18.	.049	.060	-.011	.083	.167	-.083
19.	.024	.067	-.042	.042	.167	-.125
20.	.098	.004	.094	0	0	0
21a				.125	.095	.030
21b	0	0	0	0	0	0
21c	.073	.062	.011	0	0	0

As  $m_{d1}-m_{d2}$  and  $m_{h1}-m_{h2}$  show there are not great differences between bpa values in both assignments. Still, some symptoms are more important in  $m_{d2}$  and  $m_{h2}$  than in original score. Hence, combining knowledge from the two sources of information is desirable.

#### 4. COMBINING BASIC PROBABILITY ASSIGNMENTS

As differences between  $m_{d2}$  and  $m_{h2}$  exist, it is justified to combine the both assignments. It is done according to the formula [1], [3]:

$$\forall_{s_{di} \neq f} m_d(s_{di}) = \frac{m_{d1}(s_{di}) \cdot m_{d2}(s_{di})}{\sum_{s_{dj} \neq f, j=1, \dots, k} m_{d1}(s_{dj}) \cdot m_{d2}(s_{dj})}$$

For euthyroidism,  $m_h$  is calculated in a similar way. It is worth noticing that the bpas :

- excludes symptoms which are either not found or not investigated in a population
- can be combined with even more bpas.

Contradictive bpas, i.e. situation when  $m_{d1}(s_{di}) = \frac{1}{m_{d2}(s_{di})}$  cannot happen as both  $m_{d1}$  and  $m_{d2}$  are less than 1.

However, strong differences among bpas result in flattening of  $m_d$  which could be a disadvantage of the assignment.

It is obvious that a combination of almost the same bpas causes blowing up importance of some symptoms. Hence it must be carefully investigated if combining is fruitful.

For the considered example the combination seems to be justified. The result of reasoning for descending numbers of present disease symptoms and the combined bpas in comparison to the original Crooks index sum are given in table2.

Table 2.

Number of present symptoms	Crooks' score	Bel ( $d$ ) calculated for $m_d$
18	45	1
17	43	.970
15	39	.847
13	31	.614
11	25	.483
9	20	.334
7	16	.402
5	11	.191
3	8	.111
1	3	.077
0	0	0

Though the order of symptoms was arbitrary chosen (and is the same as in original index), it is clear that the value of Bel is changing together with the original score. Still some differences are noticeable. The diagnosis of hyperthyroidism is valid if the original score is greater than 20. The equivalent threshold could be  $Bel(d)=0.4$ . However, further investigations show that much more reliable analysis consist in investigation of mutual relations between  $Bel(d)$  and  $Bel(h)$ .

### 5. RESULTS

Sample results for 10 patients healthy (in euthyroidism) and suffering form hyperthyroidism are in fig.1. and 2. Original score (in a changed scale for the clarity of drawing) as well as  $Bel(d)$  and  $Bel(h)$  for euthyroidism (in fig.1) and for hyperthyroidism (in fig.2) are presented. The threshold at 0.4 level seems not to be the right

categorising criterion as it causes that one of the cases for each group is misjudged. However observing  $Bel(d)$  and  $Bel(h)$  mutual relations we are able to find the right diagnosis: if  $Bel(d) > Bel(h)$  then a patient is hyperthyroid, on the contrary – euthyroid. The distances between  $Bel(d)$  and  $Bel(h)$  values for each case are great enough to exclude misinterpretation. A little flattening of Bel in comparison to original score is observable as the result of the combined bpa.

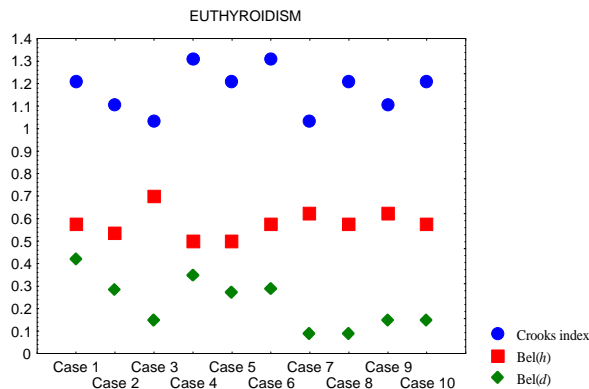


Fig.1.

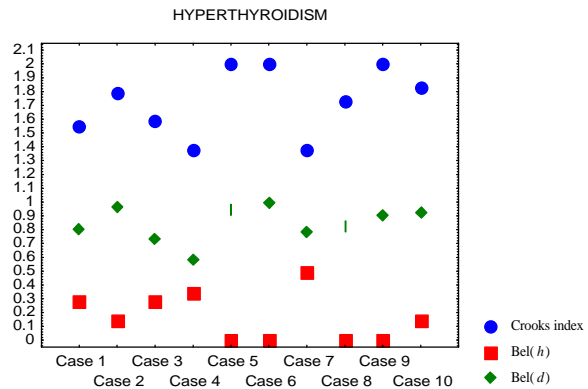


Fig.2.

### 6. CONCLUSIONS

Particular features of Dempster-Shafer theory (DST): the possibility of diagnosis certainty interpretation, replacing probability distribution by the basic probability assignment (bpa) and the ability to combine bpas make it a suitable tool for probability-based measures calculations. A comparison of belief measure values for a disease and health hypotheses render a diagnosis support. Since thyroid gland diseases occur with different frequencies in various regions what involves the necessity of specific bpa calculations, hyperthyroidism was chosen as an example of calculation. An investigation of replacing a score test which is Crooks index was carried out.

First calculation results show that:

- it is possible to interpret a medical score test by means of DST;
- it is desirable to combine a score-test-based bpa with an adequate bpa for a population;
- relations between  $Bel(d)$  and  $Bel(h)$  values should replace the original medical test interpretation consisting in the threshold categorization.

The method allows excluding symptoms that don't occur or are not investigated in a population from the original score. In the same way new symptoms can be added. Still precautions must be taken against blowing up importance of some symptoms.

There are many score tests in medical handbooks. It's most probable they can be computerised in a flexible way using DST.

#### REFERENCES

- [1] Gordon J., Shortliffe E.H. - The Dempster-Shafer Theory of Evidence, Rule -Based Expert Systems, Buchanan B.G., Shortliffe E.H. eds, Addison Wesley 1984, pp.272-292
- [2] Straszecka E., Straszecka J. - Dempster-Shafer theory of evidence in application to medical diagnosis support, Proc. Int. Workshop BEMI'97 Sept. 97, Gliwice, Poland pp 40-44
- [3] Straszecka E., Straszecka J. - Dempster-Shafer theory and basic probability assignment in diagnosis support, Proc. Joint EUROFUSE-SIC'99, in print
- [4] Xiao S., Peng C., Wang Z., Wang F. Zhiwei N. (1996). Using algebraic sum method in medical expert systems. IEEE Engineering in Medicine and biology Vol.15, No.3, pp 80-82