

# Diagnostic Rule Extraction Using the Dempster-Shafer Theory Extended for Fuzzy Focal Elements

Sebastian Porebski<sup>(✉)</sup> and Ewa Straszecka

Division of Biomedical Electronics, Faculty of Automatic Control,  
Electronics, and Computer Science, Institute of Electronics,  
Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland  
{sebastian.porebski, ewa.straszecka}@polsl.pl

**Abstract.** The Dempster-Shafer theory along with the fuzzy set theory are suitable tools for the medical diagnosis support. They can deal with medical knowledge uncertainty and data imprecision. This paper presents a study of medical knowledge representation by means of the Dempster-Shafer theory extended with the fuzzy set theory and introduces the new rule selection algorithm. The presented method gives an opportunity of interpretable and reliable rule extraction. The method is elaborated and its performance is tested on a popular medical data set. Results show that the presented method can be useful for the knowledge engineer and diagnostician cooperation due to the simple rule base and clear inference method.

**Keywords:** Rule extraction · Dempster-Shafer theory · Fuzzy sets · Medical diagnosis support · Thyroid disease

## 1 Introduction

Medical knowledge extraction is the second most popular area in data mining applications [2]. A diagnosis support tool should deal with the diagnosis uncertainty and medical data imprecision. Simultaneously, the support tool should offer not only diagnosis efficiency but also an inference that is easy to understand for physicians and a reasonable size of extracted knowledge [3]. These are inevitable requirements of a knowledge engineer and a diagnostician cooperation. In medicine even reliable tools can not substitute a diagnostician, they can only support him. The most popular medical knowledge extraction approaches are decision trees [5], support vector machines [4] and neural networks [1]. Their inference parameters (weights or structure) can be translated to the readable IF-THEN rules with additional methods. In this paper the connection of the Dempster-Shafer theory (DST) and the fuzzy set theory (FST) is used to perform diagnostic inference. With the basic probability assignment (critical component of DST) it is possible to represent the knowledge uncertainty. On the other

hand, FST is used as well-known method to represent data imprecision in many fields. As a result of a study on these methods, new rule extraction algorithm is described. The main requirement in the proposed algorithm is to provide a consistent and easy to interpret diagnostic rule base, also paying attention to its reliability. Experimental results evaluate the approach from the point of view of a diagnostician and knowledge engineer cooperation.

## 2 Medical Diagnosis Support Based on the DST and FST

Medical diagnosis support is usually based on with the IF-THEN rules [2], which shape the inference from symptoms (conditions in a premise) and diagnosis (conclusion). Let us denote  $\mathbf{x}$  as the medical data case represented with  $p$  symptoms:

$$\mathbf{x} = [x_1, x_2, \dots, x_j, \dots, x_p]^\top. \quad (1)$$

A simultaneous use of the DST and FST [8] makes it possible to define the fuzzy focal element, which is a simple or complex logical condition related to the specific medical data symptoms, which can be test results or natural language statements like “*high temperature*” [8]. The simplest focal element considers only one symptom and the most complex focal element considers  $p$  symptoms. It is clear that the possible number of fuzzy focal elements for one diagnosis equals  $2^p - 1$ . General form of the fuzzy focal element can be represented in the following way:

$$s^{(r_l)} : X_j \text{ is } A_j^{(r_l)} \wedge \dots \wedge X_k \text{ is } A_k^{(r_l)}, \quad (2)$$

where  $A_j^{(r_l)}$  is the fuzzy set described by the fuzzy membership function  $\mu_j^{(r_l)}$ , ( $j = 1, \dots, p$ ,  $r_l = 1, \dots, n^{(l)}$ ,  $l = 1, \dots, C$ ). This membership function describes  $j$ -th symptom values characteristic for the  $l$ -th diagnosis. The  $m_l(s^{(r_l)})$  value is assigned to this focal element. The  $m_l(s^{(r_l)})$ ,  $r_l = 1, \dots, n^{(l)}$ , make the basic probability assignment. In the decision support area the latter can be treated as the uncertainty of diagnostic rules.

### 2.1 Basic Probability Assignment

The BPA can be given by a human expert, i.e. diagnostician, but it can be also calculated with the training data. It must fulfil the following conditions:

$$m_l(f) = 0, \quad \sum_{s^{(r_l)} \in \mathcal{S}_l} m_l(s^{(r_l)}) = 1, \quad (3)$$

where  $f$  is a false premise and  $\mathcal{S}_l$  is the fuzzy focal element set for the  $l$ -th diagnosis. In the [6] the basic probability assignment is calculated as a measure of correspondence between  $s^{(r_l)}$  focal element and suitable  $l$ -th diagnosis cases. Its calculation can be briefly described by the following formula:

$$m_l(s^{(r_l)}) = \frac{\sum_{i=1}^{n_l} \gamma_i^{(r_l)}}{\sum_{r_l=1}^{n^{(l)}} \left( \sum_{i=1}^{n_l} \gamma_i^{(r_l)} \right)}, \quad (4)$$

where  $n_l$  is  $l$ -th diagnosis data set size,  $n^{(l)}$  is the  $\mathcal{S}_l$  size and

$$\gamma_i^{(r_l)} = \begin{cases} 1 & \text{for } \eta^{(r_l)} > \eta_{BPA}, \\ 0 & \text{for } \eta^{(r_l)} \leq \eta_{BPA}, \end{cases} \quad (5)$$

where  $\eta^{(r_l)}$  is the matching level of the  $s^{(r_l)}$  with the data case. According to (4), the basic probability value is the number of cases conformed with the  $l$ -th diagnosis cases. This number is normalized with the sum of all  $s^{(r_l)}$  conformation numbers. This operation provides truth of (3). The conformation condition is based on the matching level, which can be calculated in the following way:

$$\eta^{(r_l)} = \min_{j \in \mathbb{J}^{(r_l)}} \mu_j^{(r_l)}(x_j), \quad (6)$$

where  $\mathbb{J}^{(r_l)}$  is the set of considered symptoms in the  $s^{(r_l)}$  and  $x_j$  is the  $j$ -th symptom value of data case  $\mathbf{x}$ . According to (5), the  $\eta_{BPA}$  influences the basic probability assignment. Lower values of the threshold allow more uncertainty of the focal elements, which are treated as conformed. On the other hand, higher threshold values cause that only the best adjusted focal elements can obtain the highest BPA value.

## 2.2 Diagnosis Belief Measure

The DST-based diagnosis support is related to the diagnosis belief calculation, which adds the basic probability values of the  $s^{(r_l)}$  elements, which are sufficiently conformed with the medical data case. The belief measure is calculated as:

$$Bel_l(\mathbf{x}) = \sum_{r_l=1}^{n^{(l)}} \gamma_{Bel}^{(r_l)} m_l(s^{(r_l)}), \quad (7)$$

where

$$\gamma_{Bel}^{(r_l)} = \begin{cases} 1 & \text{for } \eta^{(r_l)} > \eta_T, \\ 0 & \text{for } \eta^{(r_l)} \leq \eta_T. \end{cases} \quad (8)$$

According to (7) and (8),  $m_l(s^{(r_l)})$  is added to  $Bel_l$ , if the  $s^{(r_l)}$  matching level with data case is higher than determined threshold  $\eta_T$ . Belief measure should be calculated for all considered diagnoses ( $D_l$ ). Final diagnosis (denoted as  $D$ ) is chosen according to the following formula:

$$D(\mathbf{x}) = D_l \text{ for } l = \operatorname{argmax}_{1 \leq l \leq C} (Bel_l(\mathbf{x})). \quad (9)$$

It is possible that the highest  $Bel$  value can be obtained for more than one diagnosis. In this situation medical diagnosis should not be stated arbitrarily. Hence the data case is treated as undiagnosed. Similarly to the  $\eta_{BPA}$ , the  $\eta_T$  threshold is the parameter, which tunes the inference and allow to use in the belief calculations more or less suited focal elements.

### 3 The Design of Membership Functions

Membership functions for fuzzy focal elements can be designed through statistical analysis of training data. In [6] different membership function shapes are considered and discussed. In this paper, trapezoidal membership functions are used. Their shape can be described as:

$$\mu(x, [a, b, c, d]) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x \leq b, \\ 1, & b < x \leq c, \\ \frac{d-x}{d-c}, & c < x \leq d, \\ 0, & x > d. \end{cases} \quad (10)$$

Let us denote  $\mathbf{X}_l$  as the subset of the medical data set  $\mathbf{X}$ , for which every data case is assigned to the  $l$ -th diagnosis. Hence,  $\mathbf{X}_l$  can be denoted as:

$$\mathbf{X}_l = \{\mathbf{x}_{l1}, \dots, \mathbf{x}_{li}, \dots, \mathbf{x}_{ln_l}\}, \quad (11)$$

where  $n_l$  is the  $\mathbf{X}_l$  set size. Moreover, let us denote  $\mathbf{x}_{jl}$  as  $j$ -th symptom values from  $\mathbf{X}_l$  as:

$$\mathbf{x}_{jl} = \{x_{j1} \dots, x_{ji} \dots, x_{jn_l}\}. \quad (12)$$

In the research one membership function is designed for each  $j$ -th symptom and  $l$ -th diagnosis data subset. Hence  $C \cdot p$  membership functions are designed. To this end, statistical measures of the  $\mathbf{x}_{jl}$  should be found:  $\bar{x}_{jl}$  and  $\sigma_{jl}$  as the mean and standard deviation values. This values can be used to treat the  $\mathbf{x}_{jl}$  as normally distributed data. Remaining values, which are used for the trapezoidal function design are: lower and upper quartiles ( $Q_{1_{jl}}, Q_{3_{jl}}$ ) of normal distribution and the intersection point  $x_c^{(l)}$  between two neighbouring ( $l$ -th and  $(l+1)$ -th) distributions. It is clear that the trapezoidal function shapes should be designed individually for data subsets with the lowest, intermediate and highest mean values. The equations used to trapezoidal membership function parameter calculation are presented in the Table 1. The  $z$  and  $Z$  in Table 1 are sufficiently small and big numbers according to minimal and maximal values of  $\mathbf{x}_{jl}$ , respectively. This operation provides, that trapezoidal function are left- and right-opened for the lowest and the highest value of training data, respectively. However, for few data couples it is possible that lower or upper quartile outreach suitable  $x_c^{(l)}$  value. Hence the design must be protected from incorrectly parameter calculation. It can be done with two conditions:

- if  $b \leq x_c^{(l-1)}$ :  
 $a = 0.99 \cdot x_c^{(l-1)}$ ,  
 $b = 1.01 \cdot x_c^{(l-1)}$ ,
- if  $c \geq x_c^{(l)}$ :  
 $c = 0.99 \cdot x_c^{(l)}$ ,  
 $d = 1.01 \cdot x_c^{(l)}$ ,

which recalculate function parameters to obtain sufficiently steep slopes [7].

**Table 1.** Trapezoidal membership function parameters calculated according to the data statistical analysis

$\mathbf{x}_{jl}$ with the lowest $\bar{x}_{jl}$	$\mathbf{x}_{jl}$ with intermediate $\bar{x}_{jl}$	$\mathbf{x}_{jl}$ with the highest $\bar{x}_{jl}$
$a = z \cdot \min_{1 \leq i \leq n_l} (x_{ji})$	$a = b - 2(b - x_c^{(l-1)})$	$a = b - 2(b - x_c^{(l-1)})$
$b = \min_{1 \leq i \leq n_l} (x_{ji})$	$b = Q_{1_{jl}}$	$b = Q_{1_{jl}}$
$c = Q_{3_{jl}}$	$c = Q_{3_{jl}}$	$c = \max_{1 \leq i \leq n_l} (x_{ji})$
$d = c + 2(x_c^{(l)} - c)$	$d = c + 2(x_c^{(l)} - c)$	$d = Z \cdot \max_{1 \leq i \leq n_l} (x_{ji})$

### 4 Rule Selection Algorithm

The main aim of the rule extraction is to find the best subset of  $2^p - 1$  possible focal elements of the  $l$ -th diagnosis. The solution should provide satisfactory results in the view of the diagnosis efficiency and rule number. The rule elimination algorithm was used to obtain simple rules in the [6,8]. According to these results a different way to the rule extraction can be proposed when an acceptable effectiveness must go along with the simplest rules. Here, the rule selection algorithm is described. Its performance is based on iterative focal element adding to the created fuzzy focal element set to obtain the best diagnosis efficiency. The rule selection algorithm can be described in the following steps:

1. Choose training data  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N\}$  as well as the  $\eta_{BPA}$  and  $\eta_T$  thresholds.
2. Design fuzzy membership functions for each symptom and each diagnosis  $\mu_j^{(l)}(x_j), l = 1, \dots, C, j = 1, \dots, p$  through statistical analysis of  $\mathbf{X}$ .
3. Create initial fuzzy focal element sets  $\hat{S}_l$  and calculate the basic probability assignment for each set of diagnostic rules  $s^{(r_l)}, r_l = 1, \dots, n^{(l)}, l = 1, \dots, C$ .
4. For each diagnosis, find the first focal element with the maximal  $m_l(s^{(r_l)})$  value in  $\hat{S}_l$  and move it to the new created focal element set  $S_l$ .
5. Recalculate the basic probability assignment for every fuzzy focal element  $s^{(r_l)}$  in  $S_l, l = 1, \dots, C$ .
6. Calculate the belief measure  $Bel_l(\mathbf{x}_i)$  for every data case,  $i = 1, \dots, N$ . Select the final diagnosis,  $l = 1, \dots, C$ .
7. Calculate the diagnosis error as the number of cases misdiagnosed and cases, for which the diagnosis can not be selected.
8. If it is the first iteration, go to step **11**. Otherwise go to step **9**.
9. If the diagnosis error increased, go to the step **13**. Otherwise go to step **10**.
10. If yet achieved error is the lowest, go to step **11**. Otherwise go to step **12**.
11. Save actual focal element sets  $S_l$  as the best ( $S'_l$ ),  $l = 1, \dots, C$ .
12. For each diagnosis where error occurs, find the focal element with the maximal  $m_l(s^{(r_l)})$  value in  $\hat{S}_l$  and move them to the created focal element set  $S_l$  and go to the step **5**.
13. Terminate the algorithm. Final focal element sets are  $S'_l$ .

In this algorithm the “most valuable” focal elements are gradually moved into the final rule set. This allows for prevalence of focal elements of the best differentiation force among diagnoses. The initial fuzzy focal element set  $\hat{S}_l$  ( $l = 1, \dots, C$ ) can contain all possible combinations of  $p$  symptoms, hence its size grows exponentially with the symptom number. On the other hand, the objective of the algorithm is to extract simple rules. Hence, the  $\hat{S}_l$  size can be limited. It is possible to specify the maximum symptom number in the focal element. Thus, focal element with too many symptom conditions (which have bad influence on their readability) are not created. This detail also affects the reduction of calculations when the high-dimensional data are proceed.

## 5 Experiments

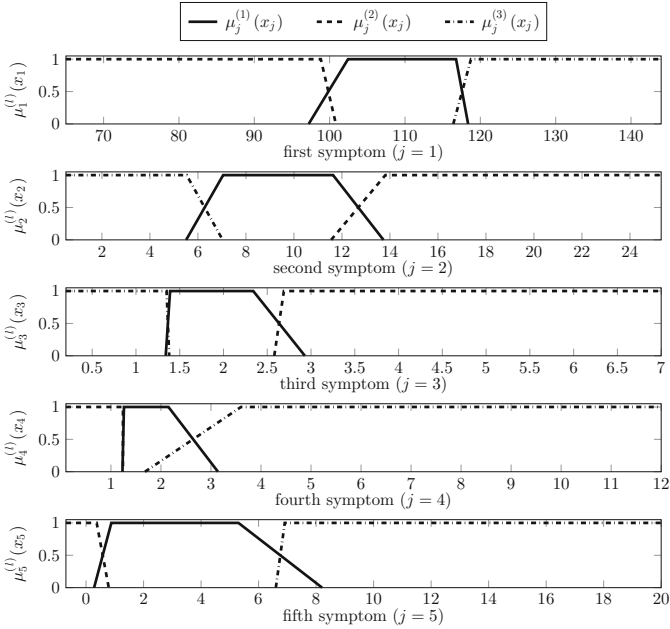
*Thyroid Disease* [9] is commonly known medical data set and often used in the diagnostic rule extraction methods. This source provides 10 different data sets related to the thyroid gland functionality. In this paper the “*new-thyroid*” set is considered. The set contains 215 medical data cases divided into three thyroid diagnoses: normal activity (150 cases), hypothyroidism (30 cases) and hyperthyroidism (35 cases). Each data case is also the set of 5 medical test results. All of them are numerical values related to following hormones:

1. triiodothyronine-resin uptake test (percentage value),
2. total serum thyroxin (T4),
3. total serum thiiodothyronine (T3),
4. total serum thyrotropine (TSH),
5. maximal TSH serum difference after 200mg of TSH-releasing hormone injection.

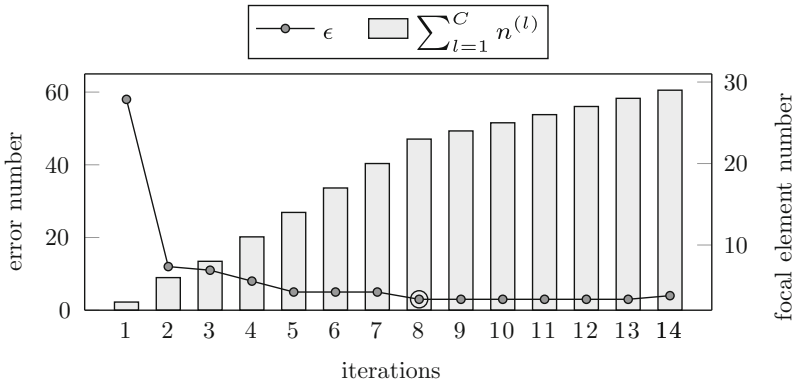
Trapezoidal membership functions designed for the *Thyroid Disease* data are presented in Fig. 1. The  $\eta_{BPA}$  and  $\eta_T$  thresholds influence the rule extraction performance. The couple that provides the highest efficiency is:  $\eta_{BPA} = 0.4$ ,  $\eta_T = 0.25$ . The performance of the algorithm of rule selection is presented in Fig. 2. Diagnosis error and rule number in every iteration are depicted with points and bars, respectively. The former are connected by line for readability. Since data symptom number is low, size of the initial fuzzy focal element sets does not have to be limited. Initial focal element sets ( $\hat{S}_l$ ,  $l = 1, \dots, C$ ) include 31 focal elements for each diagnosis. According to the Fig. 2, fuzzy focal element set obtained in the eighth iteration is chosen as the best set due to only two misdiagnosed cases. There were no further iteration, for which the diagnosis error has decreased. The efficiency of the fuzzy focal element set can be calculated in the following way:

$$E = \frac{N - \epsilon}{N} \cdot 100\% \quad (13)$$

where  $\epsilon$  is the diagnosis error and  $N$  is the data set size. According to (13), the set of 23 fuzzy focal elements created in the eighth iteration provides 99.07% efficiency, which is better results than published in [6]. The rule set for the



**Fig. 1.** Trapezoidal membership functions designed for the *Thyroid Disease* data set. Leftmost, middle and rightmost functions describe “low”, “normal” and “high” symptom value, respectively.

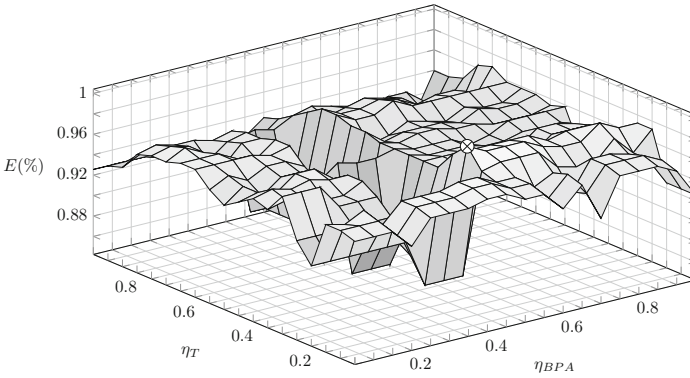


**Fig. 2.** Rule selection algorithm performance with the *Thyroid Disease* data set

thyroid disease diagnosis extracted from the training data are presented in the Table 2. It can be noticed that this efficiency comes at a price of numerous focal elements. This issue is related to the well-known trade-off between high reliability and good interpretability of the extracted knowledge [3]. With the  $\eta_{BPA}$  and  $\eta_T$  threshold there is possibility to tune the extracted knowledge. Diagnosis

**Table 2.** Fuzzy focal element sets obtained as a result of the rule selection algorithm. The  $\eta_{BPA}$  and  $\eta_T$  thresholds are set to 0.5 and 0.25, respectively.

$s^{(r_l)}$	$m_1(s^{(r_l)})$	$m_2(s^{(r_l)})$	$m_3(s^{(r_l)})$
$X_1$ is $A_1^{(1)}$	0.13	0.11	0.14
$X_2$ is $A_2^{(2)}$	0.15	0.14	0.17
$X_3$ is $A_3^{(3)}$	0.13	0.12	0.14
$X_4$ is $A_4^{(4)}$	–	0.11	0.15
$X_5$ is $A_5^{(5)}$	0.14	0.15	0.13
$X_1$ is $A_1^{(6)}$ and $X_2$ is $A_2^{(6)}$	0.11	–	0.13
$X_1$ is $A_1^{(9)}$ and $X_5$ is $A_5^{(9)}$	0.10	–	–
$X_2$ is $A_2^{(10)}$ and $X_3$ is $A_3^{(10)}$	0.11	–	–
$X_2$ is $A_2^{(11)}$ and $X_4$ is $A_4^{(11)}$	–	–	0.14
$X_2$ is $A_2^{(12)}$ and $X_5$ is $A_5^{(12)}$	0.13	0.14	–
$X_3$ is $A_3^{(14)}$ and $X_5$ is $A_5^{(14)}$	–	0.11	–
$X_4$ is $A_4^{(15)}$ and $X_5$ is $A_5^{(15)}$	–	0.11	–



**Fig. 3.** Diagnosis efficiencies obtained for all combination of  $\eta_{BPA}$  and  $\eta_T$  thresholds. The highest efficiency is indicated by  $\otimes$ .

efficiency dependence on the thresholds determination in the range 0.05 to 0.95 with 0.05 step is presented in Fig. 3. The highest obtained efficiency is indicated by mark  $\otimes$ . With the  $\eta_{BPA}$  and  $\eta_T$  thresholds set to 0.05 and 0.1 it is possible to obtain the lowest possible number of focal elements equal six providing slightly worse efficiency equal 95.35%. A lower  $\eta_{BPA}$  threshold allows the simplest rules to obtain higher BPA values and the final efficiency is acquired for very few rules. This rule set is presented in Table 3. It is possible to translate the fuzzy focal elements to the heuristic rules, which represent extracted knowledge from the medical data. According to the Fig. 1 fuzzy membership functions  $\mu_j^{(l)}(x_j)$ ,



**Table 3.** Fuzzy focal element sets obtained as a result of the rule selection algorithm. The  $\eta_{BPA}$  and  $\eta_T$  thresholds are set to 0.2 and 0.3, respectively.

$s^{(r_l)}$	$m_1(s^{(r_l)})$	$m_2(s^{(r_l)})$	$m_3(s^{(r_l)})$
$s^{(2)}$ : $X_2$ is $A_2^{(2)}$	0.5	0.5	0.5
$s^{(3)}$ : $X_3$ is $A_3^{(3)}$	–	0.5	–
$s^{(4)}$ : $X_4$ is $A_4^{(4)}$	–	–	0.5
$s^{(5)}$ : $X_5$ is $A_5^{(5)}$	0.5	–	–

**Table 4.** Heuristic rules for the thyroid gland functionality diagnosis according to the knowledge extracted from the *Thyroid Disease* data (*new-thyroid* set)

1. <i>euthyroid</i> diagnosis ( $n^{(1)} = 2$ ):	
$s^{(2)}$ : “T4 level is normal”	$m_1(s^{(2)}) = 0.50$
$s^{(5)}$ : “TSH change is normal”	$m_1(s^{(5)}) = 0.50$
2. <i>hyperthyroidism</i> diagnosis ( $n^{(2)} = 2$ ):	
$s^{(2)}$ : “T4 level is high”	$m_2(s^{(2)}) = 0.50$
$s^{(3)}$ : “T3 level is high”	$m_2(s^{(3)}) = 0.50$
3. <i>hypothyroidism</i> diagnosis ( $n^{(3)} = 2$ ):	
$s^{(2)}$ : “T4 level is low”	$m_3(s^{(2)}) = 0.50$
$s^{(4)}$ : “TSH level is low”	$m_3(s^{(4)}) = 0.50$

which are used in fuzzy focal elements sets are also describing fuzzy sets related to the natural language statements like “*low*”, “*normal*” and “*high test result*”. Natural language statements occurring in Table 4 are presented by suitable fuzzy sets in Fig. 1.

## 6 Conclusions

Experimental results show that the Dempster-Shafer theory along with the fuzzy set theory can be used to create a diagnostic rule base and to interpret heuristically data-driven medical knowledge. This framework successfully deals with medical data issues like imprecision and uncertainty. As a result of the rule selection algorithm the fuzzy focal element set can be chosen to provide sufficiently effective diagnosis. Extracted knowledge is comprehensive for a human expert thanks to the intuitive understanding of trapezoidal membership functions of fuzzy focal elements. Performance of the rule extraction algorithm can be tuned with thresholds, which determine the BPA calculation and final diagnosis determination. The search of the best thresholds may involve computational complexity but the rule extraction algorithm is always fast. Generally, the solution with best diagnosis efficiency may not be related to the lowest number of focal elements. Hence, the threshold determination is also vague. The aim is to

obtain a good compromise between the rule base reliability and its readability. This paper presents results of the knowledge extraction with the training data. Further research will deal with the mentioned compromise as well as with the rule base generalization quality.

**Acknowledgements.** This research is financed from the statutory activities of the Institute of Electronics of the Silesian University of Technology.

## References

1. Amato, F., López, A., Pena-Meñdez, E.M., Vañhara, P., Hampl, A., Havel, J.: Artificial neural networks in medical diagnosis. *J. Appl. Biomed.* **11**(2), 47–58 (2013)
2. Esfandiari, N., Babavalian, M.R., Moghadam, A.-M.E., Tabar, V.K.: Knowledge discovery in medicine: current issue and future trend. *Expert Syst. Appl.* **41**(9), 4434–4463 (2014)
3. Gacto, M.J., Alcalá, R., Herrera, F.: Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures. *Inf. Sci.* **181**, 4340–4360 (2011)
4. Han, L., Luo, S., Yu, J., Pan, J., Chen, S.: Rule extraction from support vector machines using ensemble learning approach: an application for diagnosis of diabetes. *IEEE J. Biomed. Health Inform.* **19**(2), 728–734 (2015)
5. Liu, X.D., Feng, X., Pedrycz, W.: Extraction of fuzzy rules from fuzzy decision trees: an axiomatic fuzzy sets (afs) approach. *Data Knowl. Eng.* **84**, 1–25 (2013)
6. Porebski, S., Straszecka, E.: Membership functions for fuzzy focal elements. *Arch. Control Sci.* **26**(3), 281–313 (2016)
7. Straszecka, E., Straszecka, J.: Interpretation of Medical Symptoms Using Fuzzy Focal Elements. In: Kurzyński M., Puchała E., Woźniak M., Żolierek A. (eds) *Computer Recognition Systems: Proceedings of the 4th International Conference on Computer Recognition Systems CORES 2005. Advances in Soft Computing*, vol 30, pp. 287–293. Springer, Heidelberg (2005)
8. Straszecka, E.: Combining uncertainty and imprecision in models of medical diagnosis. *Inf. Sci.* **176**, 3026–3059 (2006)
9. UCI Machine Learning Repository: Thyroid Disease Data Set. <https://archive.ics.uci.edu/ml/datasets/Thyroid+Disease>. Accessed 22 Dec 2016