Modelling Medical Uncertainties with Use of Fuzzy Sets and Their Extensions

Patryk Żywica[0000−0003−3542−8982]

Faculty of Mathematics and Computer Science
Adam Mickiewicz University in Poznań, Poland
bikol@amu.edu.pl,
WWW home page: http://min.wmi.amu.edu.pl/en/staff/patryk-zywica/

Abstract. This work presents an approach to deal with uncertainty in patient's medical record. After giving a brief characterisation of possible sources of uncertainty in medical records, the paper introduces fuzzy set based approach that allows modelling of such information. First, heterogeneous data is converted to homogeneous model with the use of Feature Set structure. With such model uncertainty may be represented directly as Fuzzy Membership Function Families (FMFFs). Some theoretical results connecting FMFFs with Hesitant Fuzzy Sets and Type-2 Fuzzy Sets are also given.

Keywords: medical data, hesitant fuzzy sets, imperfect information

Acknowledgement. This work was supported by the Polish National Science Centre grant number 2016/21/N/ST6/00316.

1 Introduction

Over the years, many models have been developed to extend and generalise the fuzzy sets theory. Motivation for a large part of them was the modelling of broadly understood data uncertainty. Most commonly used in practice are: IVFS[1], AIFS[2, 3], HFS[4], and T2FS[1, 5, 6]. All these extensions enrich the set of values that membership function can take, replacing a single number with an interval, set, and even a fuzzy number in $[0, 1]$. Depending on the interpretation, new values of membership can be considered as separate "logical" values (ontic interpretation) or as an approximation of our imperfect knowledge of a certain value (epistemic interpretation) [7, 8].

This work is devoted to uncertainty in an epistemic sense, i.e. one that results from limited, incomplete or imperfect knowledge. The simplest case is the complete lack of knowledge – the case of incomplete data. However, even here one should be careful. Missing values are often marked with a very ambiguous abbreviation *NA*. *NA* – not available – the value is not available, no one conducted an examination or test. This is a perfect example of total epistemic uncertainty. On the other hand, another interpretation of NA – not applicable – has nothing to do with the incompleteness of the data. After all, this is completely certain information that a given parameter is not suitable for describing a particular case. This fact further highlights how important it is to understand the context of modelled data.

The primary objective of this paper is to describe how one can use fuzzy sets for patient record modelling. The two main problems will be addressed. The first involves the process of transforming heterogeneous medical data into a fuzzy set with clearly defined semantics. The second concerns the modelling of epistemic uncertainty.

The rest of the paper is organised as follows. Section 2 gives some background information about fuzzy sets and their extensions. The third section covers the issue of medical data normalisation. Section 4 deals with uncertainty model for medical data. The concept of Fuzzy Membership Function Family (FMFF) is defined and some theoretical results connecting it with Hesitant Fuzzy Sets and Type-2 Fuzzy Sets are given. Finally, in Section 5 some conclusions and areas for further research are given.

2 Definitions

Let $U = \{x_1, x_2, \dots\}$ be a crisp, at most countable universal set. A mapping $A : U \to [0, 1]$ is called a fuzzy set in U. For each i, the value $A(x_i)$ (a_i for short) represents the membership grade of x_i in A. Let $\mathcal{FS}(U)$ be the family of all fuzzy sets in U.

Interval-valued fuzzy set theory, which is a special case of Type–2 Fuzzy Set (T2FS) theory, was introduced by Zadeh [1]. Let $\mathcal{I}$ be the set of all closed subintervals of $[0, 1]$. A mapping $\hat{A} : X \to \mathcal{I}$ is called an Interval-Valued Fuzzy Set. For each $1 \leq i \leq n$, the value $\hat{A}(x_i) = [\underline{A}(x_i), \overline{A}(x_i)] \in \mathcal{I}$ represents the membership of an element x_i in $\hat{A}$. Usually $\underline{A}$ and $\overline{A}$ are called the lower and upper membership functions of $\hat{A}$ respectively. In epistemic approach, interval $\hat{A}(x_i)$ is understood to contain the true membership degree of x_i in some incompletely known fuzzy set A represented by $\hat{A}$. We denote the set of all interval-valued fuzzy sets in U by $\mathcal{IVFS}(U)$.

In 2010, Torra defined Hesitant Fuzzy Sets (HFS, [4]). They perfectly combine the simplicity of IVFS and the ability to model very complex data provided by Type–2 Fuzzy Sets [9]. The important fact about Hesitant Fuzzy Sets is that they were created with the aim of representing the uncertain, epistemic data. A Hesitant Fuzzy Set is a mapping $A^H : X \to 2^{[0,1]}$. It is worth noting that this is an equivalent concept to the notion of Set-Valued Fuzzy Sets [9, 10]. Type–2 Fuzzy Sets were proposed by Zadeh in 1971 [11]. They generalise most of the known extensions of fuzzy sets. A T2FS is defined as a mapping $\tilde{A} : X \to \mathcal{FS}([0, 1])$.

3 Modelling medical data as fuzzy membership function

In applications such as classification or decision support, it is often required that individual instances (patients) should be represented as homogeneous real vectors. Given the wide variety of data types encountered in the patient record, direct conversion / normalisation of data can lead to some anomalies. In this section we will present a procedure for converting heterogeneous medical data into a homogeneous model that allow to preserve full semantics of data.

Table 1. Sample medical data represented with original values.

No	Age	Gender	Tumor in family	WBC	Qualitative assessment	Quantitative assessment
1	38	F	yes	2817.3	1/brown	2/medium
2	18	F	no	2181.8	3/blue	4/very high
3	64	F	no	8611.3	4/yellow	1/low
4	40	F	yes	3017.1	2/grey	2/medium

Table 2. Data from Table 1 after min/max normalisation.

No	Age	Gender	Tumor in family	WBC	Qualitative assessment	Quantitative assessment
1	0.38	1	1	0.282	0.0	0.333
2	0.18	1	0	0.218	0.667	1
3	0.64	1	0	0.861	1.0	0
4	0.40	1	1	0.302	0.333	0.333

3.1 Data normalisation and fuzzy sets

Let start with the example data presented in Table 1. The simplest methods of transforming these data are min/max normalisation and standardisation. Table 2 shows the results of the min/max normalisation.

Many researchers (including the author of this work [12, 13]) have tried to treat these vectors as fuzzy sets. This is a convenient approach because there are many useful tools in the field of Fuzzy Logic such as similarity measures or rule based models. This approach leads to the following fuzzy sets in the universal set $U = \{\text{age}, \text{gender}, \text{in family}, \text{WBC}, \text{qualitative}, \text{quantitative}\}$:

$$A_1 = {}^{0.38}/_{\text{age}} + {}^{1}/_{\text{gender}} + {}^{1}/_{\text{in family}} + {}^{0.282}/_{\text{WBC}} + {}^{0}/_{\text{qualitative}} + {}^{0.333}/_{\text{quantitative}} \,. \quad (1)$$

This design is correct (at least formally). In addition, through the use of appropriate methods, it is possible to achieve very good results, for example, in supporting medical diagnosis [12, 14].

But can A_1 be called a (fuzzy) set? What would it contain? It turns out that the interpretation of such fuzzy set is difficult to determine. Of course, according to the characteristics given in the classic work of Dubois and Prade [15], we are dealing here with the "degree of similarity" semantics. However, to be able to interpret the degrees of membership as the similarity, a reference point (prototype) is required. Was it determined in this case? Certainly not explicitly. For some attributes, definition of the "prototype" is a simple task: age – old person; WBC – high serum level; quantitative assessment – very high. The most difficult part is to determine prototype for the qualitative assessment, and it is an attribute type that most often cause problems. Choosing the highest value (4/yellow) as a prototype is as unjustified as any other. There is no naturally defined order between the values. It can not be said that 2/green is somehow

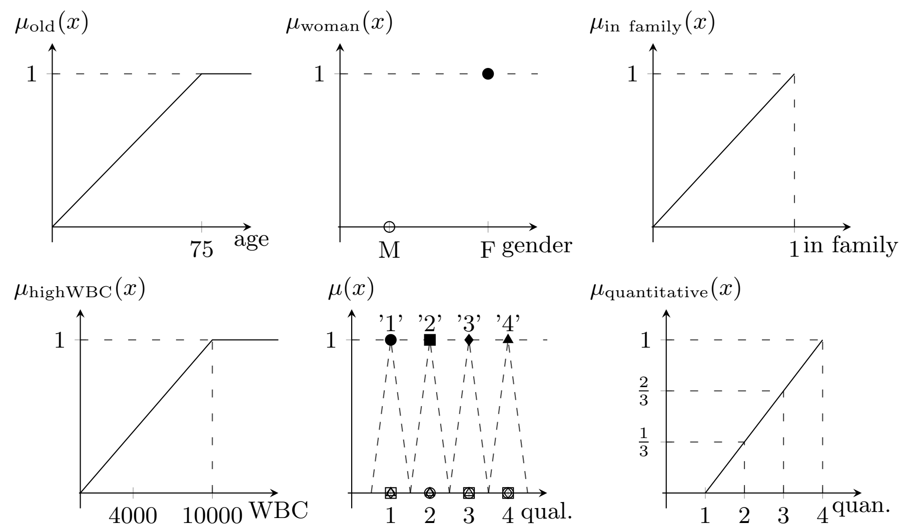

Fig. 1. Linguistic variables used to define feature set for data from previous subsection.

similar to 4/yellow, since they are not related to each other and cannot be compared.

In this section a method of data normalisation will be given, which in particular:

- allows to specify a clear semantics of a (fuzzy) set for a patient representation,
- describes how to deal with different types of data in order to preserve their interpretation,
- can be easily extended to model the uncertainty.

3.2 Proposed approach to normalisation

The basic idea is to represent each instance (the patient) as a fuzzy set. For this purpose, the concept of *feature set* will be introduced. Each instance is described by a fuzzy set of features that it has. The *feature set* itself is defined in terms of the mappings from any number (k) of attributes to a single feature $f_i : \mathrm{R}^k \to [0, 1]$. Each such mapping should also give clear description of the new feature meaning. In this paper we will present simple approach where each term of linguistic variable defines new mapping. Interpretations of terms for linguistic variables for the data from the previous example are shown on Fig. 1.

Particular attention should be paid to the attribute "qualitative assessment". The corresponding linguistic variable has 4 terms, one for each value. In this situation, the fuzzy set representing the patient will contain 4 four features for one attribute. Moreover, each of these features can have a different level of membership, symbolically indicated by the dashed line in the graph. Whether

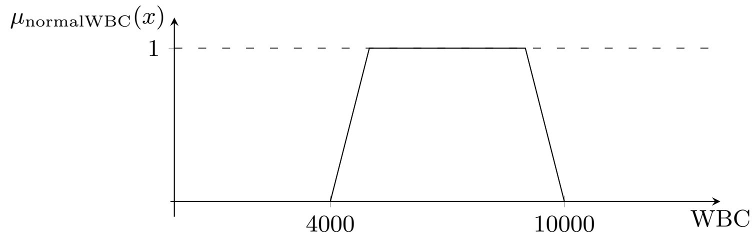

Fig. 2. Alternative definition of linguistic variable for WBC attribute.

a given attribute, and later a feature allows partial membership depends only on the semantics of the attribute (e.g. gender where a partial membership to "women" feature is rather difficult to interpret).

Also for numerical attributes, conversion to features using the linguistic variable brings some certain advantages. In addition to clearly defining the meaning of membership degree, we also get the possibility of a more advanced, but still easy to interpret, transformation of numerical values. For example, for a WBC attribute the medically recognised norm is 4000-10000. Hence, instead of defining a feature as "high", you can specify the "normal" feature, as shown in the Fig. 2. Although it's just a small change, experience shows that it can significantly affect model quality, especially when we are limited to simple models such as regression [16, 17]. In addition, the medical norm, rather than embedded into the decision model, is explicitly included in the data model, which in some applications can be very useful.

What will be the fuzzy representation of patients from the previous subsection? First, the universal set is now a set of all features in feature set. Membership values are now calculated with respect to each term interpretation. The patient record is represented by the following fuzzy set:

$$P_1 = {}^{0.5}/_{\mathrm{old}} + {}^{1}/_{\mathrm{woman}} + {}^{1}/_{\mathrm{in\ family}} + {}^{0.28}/_{\mathrm{high\ WBC}} + {}^{1}/_{1/\mathrm{brown}} + {}^{0.33}/_{\mathrm{quantitative}} \,. \quad (2)$$

Apparently little has changed compared to (1). Referring to the three semantics of a fuzzy set, we are still in the "degree of similarity" semantics. However, now P_1 can certainly be called a fuzzy set. It is a set of features possessed by the first patient. Answer to the question whether $a \in P_1$ has simple interpretation and tells whether and to what extent a feature a describes the specific medical case. All membership degrees, regardless of the data type of the corresponding attribute, now have a uniform and clearly defined interpretation.

However, the greatest advantage of this data model, is the ability to directly model incomplete and, more generally, uncertain data. The proposed data uncertainty model will be formalised in the next section. The aim of the following discussion is to show that an adequate data model can considerably simplify modelling of the uncertainty.

The benefits of using the feature set based model for the representation of incomplete data will be presented on the examples of attributes from the beginning of the section. It will be compared to the classical min/max normalisation. Singleton notation will be used, where membership value depending on the situation may be modelled by number, interval or arbitrary set which corresponds to the normal, interval and hesitant fuzzy sets.

Lack of knowledge of the patient's age may occur for elderly patients whose documentation is missing, and getting information from them is not possible. For example, if a patient was born before 1939, then in the new model, thanks to the use of non-linear normalisation with the use of the linguistic variable we have ${}^{1}/_{\mathrm{old}}$. With min/max normalisation, it would be necessary to use the interval membership degree ${}^{[0.79,1]}/_{\mathrm{age}}$. The uncertainty for the gender attribute may result from the refusal to provide information. In our model, this means that the "woman" feature can simultaneously belong and not belong to the feature set. Hence, we get ${}^{\{0,1\}}/_{\mathrm{woman}}$. Because for this feature there is no sense of partial membership, we use a two-element set instead of interval. In the classic approach, you can use the same model ${}^{\{0,1\}}/_{\mathrm{gender}}$. In the case of cancer in the family, the patient can know only the immediate family, or part of it. For example, if it is known that the cancer did not occurred in parents and the distant family is unknown, then such situation can be modelled in both approaches as ${}^{[0,0.8]}/_{\mathrm{in\ family}}$.

The most interesting, however, is the case of qualitative assessment (in the example of eye colour), where modelling of some uncertainty variants posed problems. Suppose that the only thing that is known about eye colour is that they are not 3/blue. In the original model we get ${}^{\{0,0.33,1\}}/_{\mathrm{qualitative}}$. However, such a degree of membership can not be easily converted to interval, which makes it impossible to use simple modelling using for example IVFS. Thanks to new normalisation scheme we have 4 features which allows us to use following representation ${}^{[0,1]}/_{1/\mathrm{brown}} + {}^{[0,1]}/_{2/\mathrm{grey}} + {}^{[0,1]}/_{4/\mathrm{yellow}}$. Therefore, the same knowledge was successfully modelled using only intervals. What if the doctor evaluating a given parameter is leaning towards option 1/brown but does not exclude 4/yellow? Unfortunately, this information can not be reproduced in the classic model, where at most we can write ${}^{\{0,1\}}/_{\mathrm{qualitative}}$ (solution using T2FS is possible, although is not feasible here). Thanks to the introduction of feature set, we have ${}^{[0,0.75]}/_{1/\mathrm{brown}} + {}^{[0,0.25]}/_{4/\mathrm{yellow}}$

3.3 Evaluation

Evaluation was based on test dataset from recent research on application of aggregation operators to incomplete data classification [18]. Original study group consists of 388 patients diagnosed and treated for ovarian tumor in the Division of Gynecological Surgery, Poznan University of Medical Sciences, between 2005 and 2015. Among them, 61% were diagnosed with a benign tumor and 39% with a malignant one. Moreover, 56% of the patients had no missing values in the attributes required by diagnostic models, 40% had a percentage of missing values in the range $(0\%, 50\%]$, and the remainder had more than 50% missing values. The test set consists of patients with real missing data and some proportion of patients with a complete set of features. As a result, the test set consisted of

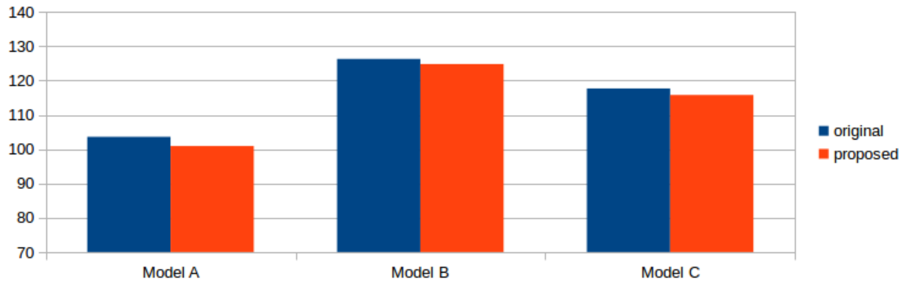

Fig. 3. Total cost of classification based on data normalised using some variant of min/max (original research) and using proposed approach. The lower the value the better classification is. Models A, B and C are selected best models form original research.

175 patients. Patients with more than 50% missing values were excluded from the study. The results from [19] are reproduced and compared with the same classification scheme applied to data normalised using proposed approach. For more information regarding dataset we refer the reader to original papers [18–20].

Performance of classification based on data normalised using some variant of min/max (original research) and using proposed approach are presented in Fig. 3. Total cost method was used to measure model performance since accuracy does not fit the medical diagnosis problem. They show that change of normalisation scheme may give slight improvement in overall classification cost. The main reason of this improvement is proper handling of qualitative attributes.

4 Families of fuzzy membership functions

The previous section shows that the patient's medical record can be reliably and unambiguously modelled while still preserving the semantics of the fuzzy set. It was also presented that it may have positive impact on modelling of incompleteness and uncertainty. Although the examples presented here will relate to medical issues, the model and results of this section are generic to any fuzzy set, regardless of interpretation.

4.1 Motivation

In this subsection, two example problems will be presented. They aim is to show that classic approaches to uncertainty modelling based on fuzzy sets may not be sufficient to fully take into account all available knowledge. First example shows that interval representation of uncertainty, though effective, is not sufficient to fully reflect the real data. Second one presents situation where some information is lost during conversion to fuzzy based representation.

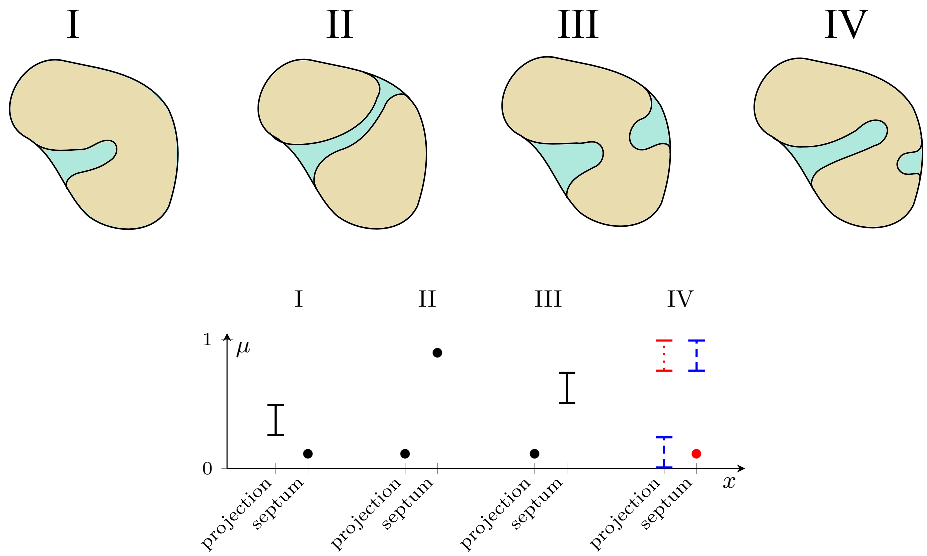

Fig. 4. Drawing of ovaries from Example 1 along with their interval representations.

Example 1. Figure 4 presents drawings of ovaries, which in actual medical practise are obtained with the help of ultrasound. Among the many features evaluated by a physician, very interesting is the case of papillary projections and septum. Drawings I and II show an average size papillary projection and total septum, respectively (along with interval representation below). Sometimes the septum is not yet completely developed which has been depicted in the third drawing. The fourth drawing shows a situation where you can not explicitly qualify whether we are dealing with a large papillary projection and the lack of septum (red, dotted line), or maybe it is a small papillary projection and almost full septum (blue, dashed line). The inability to distinguish between these two situations arises only from lack of knowledge (the impossibility of observing the ovary from the other side). As can be seen interval representation is not sufficient to cover both situations in one data description (it is necessary to assign two disjoint intervals). This example shows the limitations of interval representation of incomplete data, indicating a real need for more general formalism with greater power of expression such as HFS.

Example 2. Let assume we have information on the volume of tumor calculated as $V = \gamma abc$. Unfortunately, some diagnostic models use a combination of the two largest dimensions of the tumor $M = \alpha a + \beta b$. For simplicity, we can assume that $\alpha = \beta = \gamma = 1$ and that $a, b, c \geq 1$.

Let for a certain patient $V = 200\text{mm}^2$. Representing the three dimensions of the tumor as intervals we get $a = [1, 200]$, $b = [1, 200]$ and $c = [1, 200]$. Calculating the model using interval arithmetic we get $M = [2, 400]$.

The same data, however, can be better represented by the use of appropriate set-based approach $\mathcal{A} = \{(a,b,c) : abc = 200, a \geq b \geq c\}$. The M model with such a knowledge representation gives:

$$M' = M(\mathcal{A}) = \{M(a,b) : (a,b,c) \in \mathcal{A}\} \approx [11.696, 201] \,. \tag{3}$$

It is easy to see that the resulting model values are much more specific and $M' \subset M$.

Using interval representation and arithmetic we lose some important information. This knowledge is given in the form of dependencies between several membership degrees, and therefore can not be represented using three independent intervals.

4.2 Proposed approach to uncertainty modelling

This subsection will present a new approach to uncertainty modelling. During its design, it was assumed that it should be as simple as possible and based on fuzzy sets, and most importantly allow for modelling the situations presented in Examples 1 and 2. For the purposes of modelling medical data, we can narrow the discussion to the situation in which the universal set U is at most countable. Then, $\mathcal{FS}(U)$ can be treated as a subset of $\mathbb{R}^n$ or l^∞.

Any closed subset $\mathcal{A}$ of $\mathcal{FS}(U)$ will be called Fuzzy Membership Function Family (FMFF). Most known extensions of fuzzy sets can be fully accurately represented using FMFF. For example, interval–valued fuzzy set $\hat{A}$ corresponds to the following FMFF

$$\mathcal{A} = \left\{A \in \mathcal{FS}(U) : \forall_{x \in U} \quad \underline{A}(x) \leq A(x) \leq \overline{A}(x)\right\} \,. \tag{4}$$

This approach is largely inspired by the Mendel representation theorem and his Wavy-Slice representation [21, 22]. Referring to this theory for the above IVFS we have $\mathcal{A} = \mathrm{FOU}(\hat{A})$.

Returning to the examples from the previous subsection, one can see that both situations can be directly modelled using FMFF.

Example 3 (Solution to Example 1). The situation shown in Fig. 4 IV can be represented by the following FMFF

$$\mathcal{A}_{\mathrm{IV}} = \Big\{ {}^{\alpha}/_{\mathrm{pap}} + {}^{\beta}/_{\mathrm{septum}} \in \mathcal{FS}\big(\{\mathrm{pap}, \mathrm{septum}\}\big) : \\ (\alpha, \beta) \in [0, 0.25] \times \{1\} \cup [0, 0.25] \times [0.75, 1] \Big\} \,. \tag{5}$$

Example 4 (Solution to Example 2). The solution to this problem has already been implicitly given. Now it will be shown in the formalism of fuzzy sets. The first step is normalisation of input data. Because they are numeric attributes, simple approach is sufficient. For all three tumor dimensions, we define the same mapping shown in Fig. 5. Thus, we obtain following FMFF representing patient condition

$$\mathcal{A} = \Big\{ {}^{\mu_{\mathrm{big}}(a)}/_{\mathrm{first}} + {}^{\mu_{\mathrm{big}}(b)}/_{\mathrm{second}} + {}^{\mu_{\mathrm{big}}(b)}/_{\mathrm{third}} : \; abc = 200 \Big\} \,. \tag{6}$$

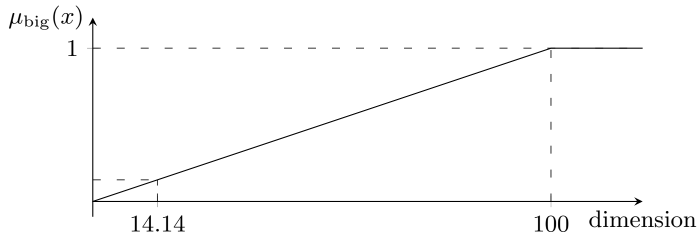

Fig. 5. Linguistic variable used to convert tumor dimensions from Example 4.

Diagnostic model needs to be adapted to handle normalised data

$$M_{\mathrm{norm}}\left({}^{\mu_{\mathrm{big}}(a)}/_{\mathrm{first}} + {}^{\mu_{\mathrm{big}}(b)}/_{\mathrm{second}} + {}^{\mu_{\mathrm{big}}(b)}/_{\mathrm{third}}\right) = a + b\,. \tag{7}$$

With this model we obtain $M_{\mathrm{norm}}\left(\mathcal{A}\right) = [11.696, 201]$ which is an optimal solution.

Definition 1. *Let $\sim$ be an equivalence relation between FMFF such that*

$$\mathcal{A} \sim \mathcal{B} \iff \forall_{x\in U}\left\{\mu_A(x) : A \in \mathcal{A}\right\} = \left\{\mu_B(x) : B \in \mathcal{B}\right\}\,. \tag{8}$$

Two FMFF are $\sim$–equivalent if they have exactly the same membership values for the same elements of the universal set. In accordance with this observation it should not be surprising that the equivalence classes $[\mathcal{A}]_\sim$ can be identified with Hesitant Fuzzy Sets.

Remark 1. Hesitant Fuzzy Sets are the smallest extension of fuzzy sets containing all FMFF in which there are no dependencies between the membership degrees of different elements of U.

Since HFS accurately describes the values of membership, ignoring the dependencies, they can be extended with the description of the dependency, gaining much wider data modelling capabilities.

Consider now the situation in which individual membership functions belonging to the FMFF have the weights assigned to them. Weighted FMFF (WFMFF) is defined as

$$\mathcal{A}^* \subseteq \mathcal{FS}(U) \times [0, 1]\,. \tag{9}$$

Similarly as before, equivalence relation can be defined.

Definition 2. *Let $\sim^*$ be an equivalence relation between WFMFF such that*

$$\mathcal{A}^* \sim^* \mathcal{B}^* \iff \forall_{w\in[0,1]}\forall_{x\in U}\left\{\mu_A(x) : (A, w) \in \mathcal{A}^*\right\} = \left\{\mu_B(x) : (B, w) \in \mathcal{B}^*\right\}\,. \tag{10}$$

In this case, the equivalence classes $[\mathcal{A}^*]_{\sim^*}$ can be identified with Type–2 Fuzzy Sets.

Remark 2. Type–2 Fuzzy Sets are the smallest extension of fuzzy sets containing all WFMFF in which there are no dependencies between the membership degrees of different elements of U.

4.3 Computational issues

In order for the proposed model to be used in practice, it is necessary to solve computational problems. The model in the proposed form is computationally inefficient at least using classical computation methods. Of course, by applying appropriate restrictions to HFS and representation of dependencies, this task can be reduced to various known computational problems, for example to linear programming, for which there is an effective calculation method. The second approach is to use highly parallel and increasingly accessible computing platforms such as GPGPU (general-purpose computing on graphics processing units) to effectively find satisfactory solutions. The author of this work believes that the combination of these two approaches will allow the construction of computationally efficient models based on FMFF.

5 Conclusions and Further Research

The presented results are the starting point for a comprehensive approach to modelling medical data. By defining an unambiguous and uniform interpretation of the patient record as a fuzzy set, further efforts can be focused on the problems of modelling data uncertainty and developing methods of supporting medical diagnostics.

The most important direction of further research, is the research on the effective use of this model in practical computations. However, for this to be possible, at least a few theoretical issues must be solved. It is necessary to examine the theoretical properties of the FMFFs which may be required to simplify computations. Moreover, it is important to solve the problem of additional elements introduced in feature set, which in some cases can increase the computational complexity and even reduce the efficiency [23]. This can be done thought introduction of hierarchical data model. It still needs to be developed formally. There is proposition to use lattice theory and L–Fuzzy Sets to solve this problem. The last step is to reconstruct some operations on FMFFs, important from the practical point of view, such as similarity measures.

References

1. L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning—I, Information sciences 8 (3) (1975) 199–249.
2. K. Atanassov, Intuitionistic fuzzy sets, Fuzzy Sets and Systems 20 (1986) 87–96.
3. K. Atanassov, Intuitionistic fuzzy sets: Theory and Applications, Vol. 35 of Studies in Fuzziness and Soft Computing, Physica-Verlag, Heidelberg, 1999.
4. V. Torra, Hesitant fuzzy sets, International Journal of Intelligent Systems 25 (6) (2010) 529–539.
5. R. John, Type 2 Fuzzy Sets: An Appraisal of Theory and Applications, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 6 (6) (1998) 563–576.

6. J. M. Mendel, R. B. John, Type-2 fuzzy sets made simple, IEEE Transactions on Fuzzy Systems 10 (2) (2002) 117–127.
7. D. Dubois, H. Prade, Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets, Fuzzy Sets and Systems 192 (2012) 3–24.
8. I. Couso, L. Sánchez, Machine learning models, epistemic set-valued data and generalized loss functions: An encompassing approach, Information Sciences 358 (2016) 129–150.
9. H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, B. De Baets, A historical account of types of fuzzy sets and their relationships, IEEE Transactions on Fuzzy Systems 24 (1) (2016) 179–194.
10. I. Grattan-Guinness, Fuzzy Membership Mapped onto Intervals and Many-Valued Quantities, Mathematical Logic Quarterly 22 (1) (1976) 149–160.
11. L. A. Zadeh, Quantitative fuzzy semantics, Information sciences 3 (2) (1971) 159–176.
12. P. Żywica, A. Wójtowicz, et al., Improving medical decisions under incomplete data using interval–valued fuzzy aggregation, in: Proceedings of 9th European Society for Fuzzy Logic and Technology (EUSFLAT), Gijón, Spain, 2015, pp. 577–584.
13. A. Stachowiak, P. Żywica, K. Dyczkowski, A. Wójtowicz, An interval-valued fuzzy classifier based on an uncertainty-aware similarity measure, in: Advances in Intelligent Systems and Computing, Vol. 322, Springer, Heidelberg, 2015, pp. 741–751.
14. P. Żywica, K. Dyczkowski, A. Wójtowicz, A. Stachowiak, S. Szubert, R. Moszyński, Development of a fuzzy-driven system for ovarian tumor diagnosis, Biocybernetics and Biomedical Engineering 36 (4) (2016) 632–643.
15. D. Dubois, H. Prade, The three semantics of fuzzy sets, Fuzzy sets and systems 90 (2) (1997) 141–150.
16. R. Moszyński, P. Żywica, et al., Menopausal status strongly influences the utility of predictive models in differential diagnosis of ovarian tumors: An external validation of selected diagnostic tools, Ginekologia Polska 85 (12) (2014) 892–899.
17. A. Wójtowicz, P. Żywica, et al., Dealing with Uncertainty in Ovarian Tumor Diagnosis, in: K. Atanassov, W. Homenda, et al. (Eds.), Modern Approaches in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications, SRI PAS, Warsaw, 2014, pp. 151–158.
18. A. Wójtowicz, P. Żywica, et al., Solving the problem of incomplete data in medical diagnosis via interval modeling, Applied Soft Computing 47 (2016) 424–437.
19. P. Żywica, Similarity Measures of Interval–Valued Fuzzy Sets in Classification of Uncertain Data. Applications in Ovarian Tumor Diagnosis, Ph.D. thesis, Faculty of Mathematics and Computer Science of Adam Mickiewicz University, in Polish (June 2016).
20. K. Dyczkowski, Intelligent Medical Decision Support System Based on Imperfect Information: The Case of Ovarian Tumor Diagnosis, Vol. 735, Springer, 2017.
21. J. M. Mendel, Tutorial on the uses of the interval type-2 fuzzy set's Wavy Slice Representation Theorem, in: Proceedings of Annual Meeting of the North American Fuzzy Information Processing Society, New York City, USA, 2008, pp. 1–6.
22. J. M. Mendel, R. I. John, F. Liu, Interval type-2 fuzzy logic systems made simple, IEEE Transactions on Fuzzy Systems 14 (6) (2006) 808–821.
23. G. Hughes, On the mean accuracy of statistical pattern recognizers, IEEE transactions on information theory 14 (1) (1968) 55–63.