Jerry M. Mendel
University of Southern California

Type-2 Fuzzy Sets:
Some Questions and Answers

Jerry M. Mendel
Department of Electrical Engineering
University of Southern California, Los Angeles, CA 90089-2564

Abstract. **To use type-1 fuzzy sets as models for words is scientifically incorrect. Type-2 fuzzy sets let us model the uncertainties that are inherent in words as well as other uncertainties. This article introduces the reader to type-2 fuzzy sets through a series of questions and answers that will hopefully provide the motivation to learn more about them and to use them.**

1. Introduction

Fuzzy sets have been around for nearly 40 years and have found many applications. However as I will explain they suffer from certain problems. These fuzzy sets are, in fact, type-1 fuzzy sets. Type-2 fuzzy sets are 'fuzzy fuzzy' sets and are more expressive as we shall see in this article.

Recently (Mendel, 2003), I demonstrated [using Popper's (1954) *Falsificationism*] that to use a type-1 fuzzy set (FS) to model a word is scientifically incorrect, because a word is uncertain whereas a type-1 FS is certain. This may come as a shock to many people, because a FS has been proposed as a model for a word from the very beginning of fuzzy sets, e.g., Zadeh's (1965) first example uses a FS to model a word. Just about every textbook does the same. And this is also true about *computing with words* (e.g., Zadeh (1996)).

Fortunately, most applications of type-1 FSs only use the mathematics of such sets and do not focus on them as actual models for words, e.g. rule-based fuzzy systems, in which antecedents and consequents are words, are often used only in the context of some sort of universal function approximation which is mathematics and not linguistics.

There are also applications of type-1 FSs in which a fuzzy system is used to approximate random data, or to model an environment that is changing in an unknown way with time. Even though universal approximation may also be the underlying basis for these applications, a type-1 FS has limited capabilities to directly handle such uncertainties, where by *handle* I mean *to model and minimize the effect of*. That a type-1 FS cannot do this sounds paradoxical because the word *fuzzy* has the connotation of uncertainty. This *paradox* about a type-1 FS has been known for a long time, and yet not much has been done about it. It has been largely ignored.

Zadeh (1975) already recognized there was a problem with a type-1 FS, when he introduced a type-2 (and even higher-types) FS. This occurred almost 10 years after the publication of his first seminal paper. One could ask: "Why did it take so long for this to happen?" and "Why didn't a type-2 FS immediately become popular?" Eventually I will answer these questions, but first there are more fundamental questions that I will pose and answer about uncertainty and a type-2 FS, since most of the readers of this Newsletter will be unfamiliar with such a set.

2. Some Questions and Answers

1) ***Can you be more specific about the "paradox" of a type-1 FS?***
I am not sure who first referred to "fuzzy" being *paradoxical*, meaning that the word fuzzy has the connotation of *uncertainty*, and yet the MF of a FS is completely certain once its parameters are specified. The following quote from Klir and Folger (1988) uses the word paradoxical: "The accuracy of any MF is necessarily limited. In addition, it may seem problematical, if not *paradoxical*, that a representation of fuzziness is made using membership grades that are themselves precise real numbers. Although this does not pose a serious problem for many applications, it is nevertheless possible to extend the concept of a FS to allow for the distinction between grades of membership to become blurred."

2) ***There are different kinds of uncertainty, so which one(s) are you referring to, and where does randomness fit into all of this?***
Indeed, uncertainty comes in many guises and is independent of what kind of FS or any kind of methodology one uses to handle it. One of the best sources for general discussions about uncertainty is Klir and Wierman (1998). Regarding the *nature of uncertainty*, they state (1998, p. 43): "Three types of uncertainty are now recognized... *fuzziness* (or vagueness), which results from the imprecise boundaries of FSs; *nonspecificity* (or information-based imprecision), which is connected with sizes (cardinalities) of relevant sets of alternatives; and *strife* (or discord), which expresses conflicts among the various sets of alternatives." Observe that these three kinds of uncertainties all involve something about sets, and, as we know a FS is characterized by its MF. So, I will interpret any and all kinds of uncertainties as being transferred to the MF of the FS. If a FS is used to model a word, then these kinds of uncertainties could be called *linguistic uncertainties*. A FS may also be used to model random or time-varying signals.

I shall distinguish between two high-level kinds of uncertainties, *random* and *linguistic*. Probability theory is associated with the former, and, as we now know FSs can be associated with the latter. If FSs are used in applications in which randomness is present (as can occur, e.g. in statistical signal processing or digital communications) then *both kinds of uncertainties should be accounted for*. This does not necessarily mean that random uncertainties have to be modeled probabilistically. Bounded uncertainties can be modeled deterministically and this can be done within the framework of a FS. It is also possible to combine fuzzy sets and probability (e.g., Buckley, 2003), but this article is not about doing this. My arguments below about using type-2 fuzzy sets should apply there as well.

3) ***What exactly does "both kinds of uncertainties should be accounted for" mean?***
Within probability theory we begin with a probability density function (pdf) that

embodies total information about random uncertainties. In most practical applications it is impossible to know or determine the pdf; so, we fall back on using the fact that a pdf is completely characterized by all of its moments. For most pdfs, an infinite number of moments are required. Of course, it is not possible, in practice, to determine an infinite number of moments; so, instead, we compute as many moments as are necessary to extract as much information as possible from the data. At the very least, we use two moments-the mean and variance; and, in some cases, we use even higher-than-second-order moments. To just use the first-order moments would not be very useful, because random uncertainty requires an understanding of dispersion about the mean, and this information is provided by the variance. So, our accepted probabilistic modeling of random uncertainty focuses to a large extent on methods that use *at least* the first two moments of a pdf. This is, for example, why designs based on minimizing mean-squared errors are so popular.

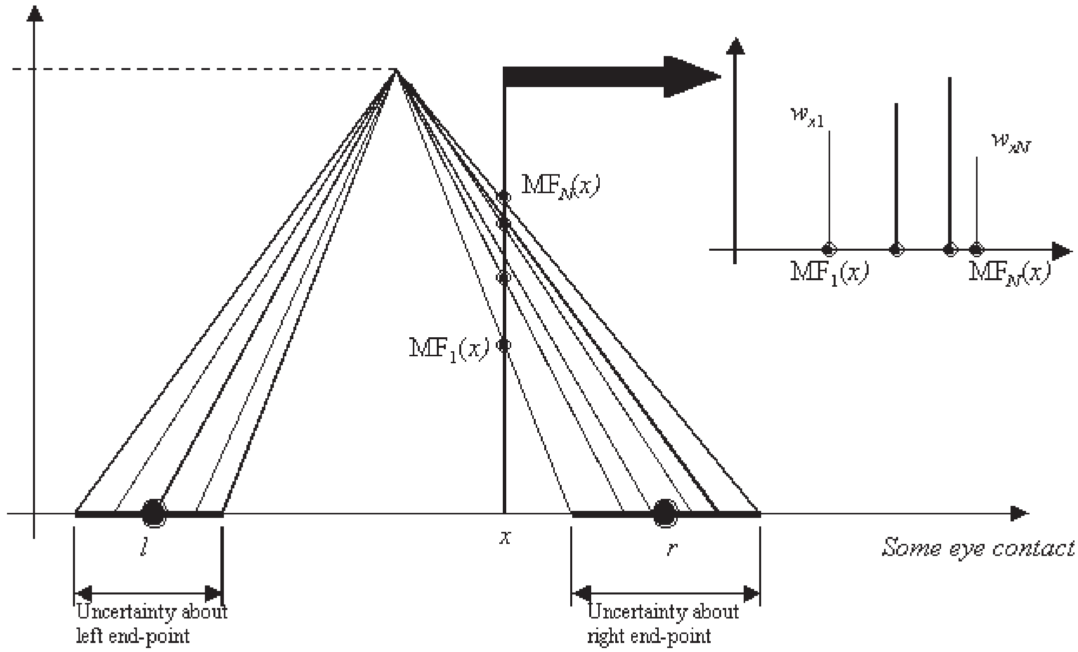

Figure 1. Triangular MFs when base end-points (l and r) have uncertainty intervals associated with them. This is not a unique construction.

Should we expect any less when we use a FS to model linguistic uncertainties? Just as variance provides a measure of dispersion about the mean, and is used to capture more about probabilistic uncertainty in practical statistical-based designs, a FS also needs some measure of dispersion to capture more about linguistic uncertainties than just a single number - which is all that we will get when we use a type-1 FS. A Type-2 FS provides this measure of dispersion.

4) *Where do uncertainties occur in a rule-based fuzzy system?*

Quite often, the knowledge that is used to construct the rules in a rule-based fuzzy system is *uncertain*. Three ways in which such rule uncertainty can occur are: (1) the words that are used in antecedents and consequents of rules can mean different things to different people; (2) consequents obtained by polling a group of experts will often be different for the same rule, because the experts will not necessarily be in agreement; and, (3) only noisy training data is available. Antecedent or consequent uncertainties translate into uncertain antecedent or consequent MFs.

5) *What exactly is a type-2 FS and how is it different from a type-1 FS?*

As already mentioned above, the concept of a type-2 fuzzy set was introduced first by Zadeh (1975) as an extension of the concept of an ordinary fuzzy set, i.e. a type-1 fuzzy set. Type-2 fuzzy sets have grades of membership that are themselves fuzzy. At *each value* of the primary variable (e.g., pressure, temperature), the membership is a function (and not just a point value) -the *secondary MF*-, whose domain -the *primary membership*- is in the interval [0,1], and whose range -*secondary grades*- may also be in [0,1]. Hence, the MF of a type-2 fuzzy set is three-dimensional, and it is the new third dimension that provides new design degrees of freedom for handling uncertainties. Such sets are useful in circumstances where it is difficult to determine the exact MF for a FS, as in modeling a word by a FS.

As an example, suppose the variable of interest is *eye contact*, which we denote as *x*. Let's put *eye contact* on a scale of values 0-10. One of the terms that might characterize the amount of perceived *eye contact* (e.g., during flirtation) is "some eye contact." Suppose that we surveyed 100 men and women, and asked them to locate the ends of an interval for *some eye contact* on the scale 0-10. Surely, we will not get the same results from all of them, because words mean different things to different people.

One approach to using the 100 sets of two end-points is to average the end-point data and to use the average values for the interval associated with *some eye contact*. We could then construct a triangular (other shapes could be used) MF, MF(x), whose base end-points (on the *x*-axis) are at the two average values and whose apex is midway between the two end-points. This type-1 triangular MF can be displayed in two-dimensions. Unfortunately, it has completely ignored the uncertainties associated with the two end-points.

A second approach is to make use of the average values and the standard deviations for the two end-points. By doing this we are blurring the location of the two end-points along the *x*-axis. Now locate triangles so that their base end-points can be anywhere in the intervals along the *x*-axis associated with the blurred average end-points. Doing this leads to a continuum of triangular MFs sitting on the *x*-axis, e.g. picture a whole bunch of triangles all having the same apex point but different base points, as in Fig.1. For purposes of this discussion, suppose there are exactly 100 (N) such triangles. Then at each value of x, there can be up to N MF values, $MF_1(x)$, $MF_2(x)$,…, $MF_N(x)$. Let's assign a weight to each of the possible MF values, say w_{x1}, w_{x2},…, w_{xN} (see Fig.1). We can think of these weights as the *possibilities* associated with each triangle at this value of *x*. At each *x*, the MF is itself a function -the secondary MF- ($MF_i(x)$, w_{xi}), where i = 1, …, N. Consequently, the resulting type-2 MF is three-dimensional.

6) ***If all uncertainty disappears, does a type-2 FS reduce to a type-1 FS?***

Yes, it does. You can already see this in Fig. 1, because if the uncertainties about the left and right end points disappear then only one triangle survives. This is sort of similar to what happens in probability, when randomness degenerates to determinism, in which case the pdf collapses to a single point. So, just as determinism is embedded in randomness, a type-1 FS is embedded in a type-2 FS.

7) ***Why are the pictures in Fig. 1 two-dimensional when the MF of a type-2 FS is three-dimensional?***

It is not as easy to sketch three-dimensional figures of a type-2 MF as it is to sketch two-dimensional figures of a type-1 MF. Another way to visualize a type-2 FS is to sketch (plot) its two-dimensional domain, its *footprint of uncertainty* (FOU), and this is easy to do. The heights of a type-2 MF (its secondary grades) sit atop of its FOU. In Fig. 1, if we filled in the continuum of triangular MFs we would obtain a FOU. Another example of a FOU is shown in Fig. 2. It is for a Gaussian primary MF whose standard deviation is known with perfect certainty, but whose mean, *m*, is uncertain and varies anywhere in the interval from m_1 to m_2. The uniform shading over the entire FOU means that we are assuming uniform weighting (possibilities). Because of the uniform weighting, this type-2 FS is called an *interval type-2 FS*. Such type-2 FSs are today the most widely used ones.

8) ***Is there new terminology for a type-2 FS?***

Yes, there is. The fact that we must now distinguish between a type-1 and type-2 MF is one example of the new terminology. A lot of the new terminology is due to the three-dimensional nature of a type-2 MF. Another term that we have already explained is the FOU. Some other new terms are: primary membership, primary MF, secondary grade, secondary MF, upper and lower MFs, principal MF, embedded type-1 FS and embedded type-2 FS. All of these terms can be defined mathematically and let us communicate effectively about type-2 FSs. A good place to learn about these terms is Mendel and John (2002). And, don't be put off by having to learn new terminology and definitions. We all had to do it when we learned probability, so why should we expect to have to do less when we are now going to use a FS to model linguistic uncertainties?

9) ***How does one choose the MF for a type-2 FS?***

Aha, the $64 question! Just as there is no one answer to this question for a type-1 FS, there is no one answer to this question for a type-2 FS. But, I would like to recommend a simplification. Begin by specifying the FOU. Then, instead of using arbitrary possibilities at each point of the FOU, use the same possibility over the entire FOU. Doing this you will be using an *interval type-2 FS*. Since there is in general no single best choice for a type-1 MF, it seems a bit foolhardy (to me) to believe that at each value of the primary variable, x, there is some optimal secondary MF. In Mendel (2003), I explain why at present the only sensible way to model a word using a type-2 FS is to use an equally weighted FOU. This does not mean though that there couldn't be some very interesting and important theoretical works to be done on more general type-2 FSs.

10) ***Is there an increase in computational complexity using three-dimensional MFs?***

For general type-2 FSs computational complexity is severe. On the other hand, set theoretic and arithmetic computations (yes, you will have to learn how to perform these for type-2 FSs) for interval type-2 FSs are very simple. They all use *interval arithmetic* and many closed-form formulas exist. Computing with interval type-2 FSs is simple. This is another important reason for working with interval type-2 FSs.

11) ***What is the earlier mentioned measure of dispersion for a type-2 FS?***

It should be pretty clear, just by looking at the FOU in Fig. 2 that less (more) uncertainty can be associated with a smaller (larger) FOU, but this is not very quantitative. When we use a type-1 FS, e.g. in a rule-based system, we perform defuzzification in order to obtain a numerical output for that system. Regardless of what kind of defuzzification we choose, we can interpret defuzzification as a mapping of a two-dimensional MF into a one-dimensional MF - a number. When we use a type-2 FS, e.g. in a rule-based system, we ultimately must also obtain a number. This is done in two stages: (1) determine the centroid of the type-2 FS (Karnik and Mendel, 2001)-it will be a type-1 FS; and (2) defuzzify the centroid. Computing the centroid of a general type-2 FS involves an enormous amount of computation; however, computing the centroid of an interval type-2 FS only involves two independent iterative computations that can be performed in parallel. This is because *the centroid of an interval type-2 FS is an interval set*, and such a set is completely characterized by its left- and right end points-yet another reason for using interval type-2 FSs. The larger (smaller) the amount of uncertainty-as reflected by a larger (smaller) FOU-the larger (smaller) will the centroid of the type-2 FS be. So the centroid provides a very useful measure of dispersion for a type-2 FS.

12) ***Other than for modeling words, what are some situations where by using type-2 FSs we may outperform the use of type-1 FSs?***

Some specific situations where we have found that using type-2 FSs will let us outperform type-1 FSs are: (1) *Measurement noise is non-stationary*, but the nature of the non-stationarity cannot be expressed ahead of time mathematically (e.g., variable SNR measurements); (2) A *data-generating mechanism is time-varying*, but the

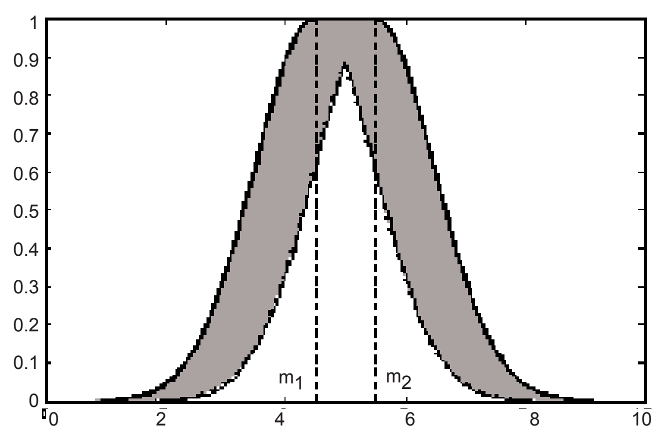

Figure 2. FOU for Gaussian (primary) MF with uncertain mean.

nature of the time-variations cannot be expressed ahead of time mathematically (e.g., equalization of non-linear and time-varying digital communication channels); and, (3) *Features are described by statistical attributes that are non-stationary*, but the nature of the non-stationarity cannot be expressed ahead of time mathematically (e.g., rule-based classification of video traffic).

13) ***Why do we believe that by using type-2 FSs we will outperform the use of type-1 FSs?***

Type-2 FSs are described by MFs that are characterized by more parameters than are MFs for type-1 FSs. Hence, type-2 FSs provide us with more design degrees of freedom; so using type-2 FSs has the *potential* to outperform using type-1 FSs, especially when we are in uncertain environments. Note that, at present, there is no theory that guarantees that a type-2 FS will always do this.

3. Conclusions

We are now ready to answer the two questions posed in the Introduction.

a) ***Why did it take so long for the concept of a type-2 FS to emerge?***

It seems that science moves in progressive ways where one theory is eventually replaced or supplemented by another, and then another. In school we learn about determinism before randomness. Learning about type-1 FSs before type-2 FSs fits a similar learning model. So, from this point of view it was very natural for fuzzyites to develop type-1 FSs as far as possible. Only by doing so was it really possible later to see the shortcomings of such FSs when one tries to use them to model words or to apply them to situations where uncertainties abound.

b) ***Why didn't type-2 fuzzy sets immediately become popular?***

Although Zadeh introduced type-2 FSs in 1975, very little was published about them until the mid-to late nineties. Until then they were studied by only a relatively small number of people, including: Mizumoto and Tanaka (1976, 1981), Nieminen (1977), Dubois and Prade (1978, 1979), Gorzalczany (1987), and, Wagenknecht and Hartmann (1988). Recall that in the 1970's people were first learning what to with type-1 FSs, e.g. fuzzy logic control. Bypassing those experiences would have been unnatural. Once it was clear what could be done with type-1 FSs, it was only natural for people to then look at more challenging problems. This is where we are today.

One last question:

c) ***How can I learn more about type-2 FSs?***

I would start with the article by Mendel and John (2002), and would then read Mendel (2001) (modulo focusing on *interval* type-2 FSs). Doing the latter will save you a lot of time. Oh, and there is lots of free *type-2 software* available at: http://sipi.usc.edu/~mendel/software

Acknowledgements

The author would like to thank Profs. Robert John and Qilian Liang as well as Ms. Honqwei Wu for their very helpful critiques of the draft of this article.

References

Buckley, J. J. (2003). Fuzzy Probabilities: New Approach and Applications, Physica-Verlag, Heidelberg, Germany.

Dubois, D. and Prade, H. (1978). "Operations on fuzzy numbers," Int. J. Systems Science, vol. 9, pp. 613-626.

Dubois, D. and Prade, H. (1979). "Operations in a fuzzy-valued logic," Information and Control, vol. 43, pp. 224-240.

Gorzalczany, M. B. (1987). "A method of inference in approximate reasoning based on interval-valued fuzzy sets," Fuzzy Sets and Systems, vol. 21, pp. 1-17.

Karnik N. N. and Mendel, J. M. (2001). "Centroid of a type-2 fuzzy set," Information Sciences, vol. 132, pp. 195-220.

Klir, G. J. and T. A. Folger (1988). Fuzzy Sets, Uncertainty, and Information, Prentice-Hall, Englewood Cliffs, NJ.

Klir, G. J. and Wierman, M. J. (1998). Uncertainty-Based Information, Physica-Verlag, Heidelberg, Germany.

Mendel, J. M., (2001). Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice-Hall, Upper Saddle River, NJ.

Mendel, J. M., (2003). "Fuzzy sets for words: a new beginning," Proc. of IEEE Int'l. Conf. on Fuzzy Systems, St. Louis, MO, pp. 37-42.

Mendel, J. M. and John, R. I. (2002). "Type-2 Fuzzy Sets Made Simple," IEEE Trans. on Fuzzy Systems, vol. 10, pp. 117-127.

Mizumoto, M. and Tanaka, K, (1976). "Some properties of fuzzy sets of type-2," Information and Control, vol. 31, pp. 312-340.

Mizumoto, M. and Tanaka, K, (1981). "Fuzzy sets of type-2 under algebraic product and algebraic sum," Fuzzy Sets and Systems, vol. 5, pp. 277-290, 1981.

Nieminen, J. (1977). "On the algebraic structure of fuzzy sets of type-2," Kybernetica, vol. 13, no. 4.

Popper, K. (1959). The Logic of Scientific Discovery (translation of Logik der Forschung) Hutchinson, London.

Wagenknecht, M. and Hartmann, K. (1988). "Application of fuzzy sets of type-2 to the solution of fuzzy equation systems," Fuzzy Sets and Systems, vol. 25, pp. 183-190.

Zadeh, L. A. (1965). "Fuzzy sets," Information and Control, vol. 8, pp. 338-353.

Zadeh, L. A. (1975). "The concept of a linguistic variable and its application to approximate reasoning - 1," Information Sciences, vol. 8, pp. 199-249.

Zadeh, L. A. (1996). "Fuzzy logic=computing with words," IEEE Trans. Fuzzy Systems, vol. 4, pp. 103-111.

Jerry M. Mendel is Professor of Electrical Engineering at the University of Southern California in Los Angeles, USA. He has published over 430 technical papers and is author and/or editor of eight books. His present research interests include: type-2 fuzzy logic systems and their applications to a wide range of problems, including target classification and computing with words, and, spatio-temporal fusion of decisions. He is a Fellow of the IEEE and a Distinguished Member of the IEEE Control Systems Society. He was President of the IEEE Control Systems Society in 1986, and is presently Chairman of the Fuzzy Technical Committee of the IEEE Neural Networks Society. Among his awards are the 1983 Best Transactions Paper Award of the IEEE Geoscience and Remote Sensing Society, the 1992 Signal Processing Society Paper Award, the 2002 Transactions on Fuzzy Systems Outstanding Paper Award, a 1984 IEEE Centennial Medal, and an IEEE Third Millenium Medal.

The editorial board at the IEEE coNNectionS would like to acknowledge the contribution of Andres Perez-Uribe on his graphic art appeared in the cover page of May 2003 issue (a single neuron driving a mobile robot). Additionally, sincere gratitude goes to Thomas Bräunl for his illustration appeared in page 9 of the feature article in the last issue and titled "eyebot robot."