Fuzzy Sets for Words: a New Beginning

Jerry M. Mendel
Signal and Image Processing Institute
Department of Electrical Engineering
University of Southern California
3740 McClintock Ave.
Los Angeles, CA 90089-2564

***Abstract*-This paper begins with a delineation of two approaches to fuzzy sets, *abstract mathematics* and *models for words*. It demonstrates, by using Karl Popper's *Falsificationism*, the present approach to fuzzy sets (FSs) for words is *scientifically incorrect*. A new theory of fuzzy sets is then presented for words that is based on collecting data from people —*person MFs*—that reflect *intra-* and *inter-levels of uncertainties* about a word, and defines a word FS as the union of all such person fuzzy sets. It also demonstrates that *intra-uncertainty* about a word can be modeled using type-2 person fuzzy sets, and that *inter-uncertainty* about a word can be modeled by means of an equally weighted union of each person's type-2 fuzzy set. Finally, it proposes a methodology for obtaining a *parsimonious* parametric type-2 fuzzy set approximation to the aggregated type-2 person FSs. This new theory of fuzzy sets for words is testable and is therefore subject to refutation.**

I. Introduction

Stepping back from all of the work that has been done since Zadeh's seminal paper [13] about fuzzy sets, it appears that there have been two very different approaches to fuzzy sets—*Abstract Mathematics* and *Models for Words*. These two approaches are diagrammed in Fig. 1. Both approaches begin with the definition of a fuzzy set (FS): $A = \{(x, \mu_A(x)) | x \in X\}$, where $\mu_A(x) \in [0,1]$. We shall refer to membership function (MF) $\mu_A(x)$ as the *original (Zadeh) type-1 MF*. Referring to the left-hand path of Fig. 1, we are permitted to only use mathematical *non-word* examples to illustrate A, x or $\mu_A(x)$ (does anyone do this?) which is why this approach is called *abstract mathematics*. If we use a word example at this point to illustrate a fuzzy set, then we are using a fuzzy set to model words, and we find ourselves on the right-hand path in the Fig. 1 diagram, even if all we intended to do with the FS is mathematics. As we explain below, a FS needs to be redefined when it is used for words, which is why dashed path 1 on Fig. 1 is illicit. At the abstract mathematical level we can draw a picture for $\mu_A(x)$ without having to provide a linguistic tag for what A is. Clearly, within this abstract mathematical framework, no one can doubt that a fuzzy set has been defined correctly, so that the mathematics about fuzzy sets can be further developed. In this approach what we then obtain are lots of mathematics for both type-1 and type-2 fuzzy sets. At some point though, when this approach is used to solve real problems, MFs must

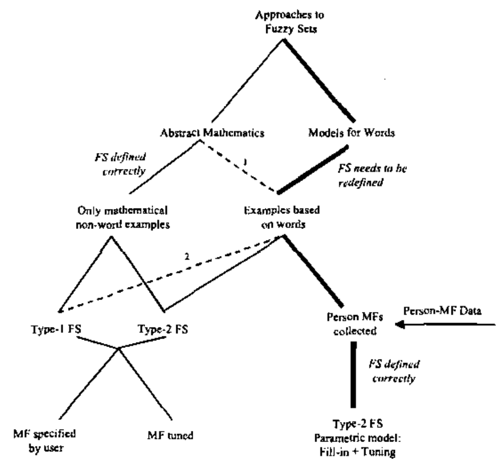

Fig 1: Two approaches to fuzzy sets and the path (heavy) associated with *fuzzy sets for words*. Dashed lines 1 and 2 denote illicit paths.

be completely specified, meaning that MF shapes and parameters must be completely specified, either a priori or by tuning.

Much of fuzzy logic control (FLC) can be presented in this abstract mathematical framework, the major exception being the early works on FLC where rules were stated using words, and human experts provided the rules (e.g., the famous truck backing up and inverted pendulum controllers). When rules are not obtained from human experts, we can interpret the FLC as abstract mathematics; however, if some or all rules are obtained from humans then we run into the models for words approach to fuzzy sets.

Returning to Fig. 1, observe in its right-hand *models for words* path that we are now permitted to provide non-abstract examples of fuzzy sets. All such examples involve *words* (terms), e.g., low pressure, sticky chocolate, etc. In fact, after Zadeh defined a FS [13], his first example used words. Note that this had nothing to do with *computing with words* (e.g., [15], [16], [4], [5], [12]), a phrase that he coined decades later, although today *computing with words* obviously means that we are taking the models for words approach to fuzzy sets.

Using fuzzy sets to model words can be interpreted as a *scientific theory*, and it is therefore legitimate to ask the questions: (1) is the use of fuzzy sets to model words *scientific*? And (2) if so, is it a *correct* scientific theory? The famous 20th century philosopher Karl Popper ([8]-[10]), established that:

> A theory is scientific only if it is refutable by a conceivable event. Every genuine test of a scientific theory, then, is logically an attempt to refute or to falsify it, and one genuine counter instance falsifies the whole theory.

This is called *Falsificationism*. For a theory to be called *scientific* it must therefore be *testable* (e.g., astrology is not testable, hence it is not scientific theory). A scientific theory can be *correct* or *incorrect*. An incorrect scientific theory is still a scientific theory, but is one that must be replaced by another scientific theory that is argued to be correct, but is itself always subject to refutation at a later date.

Let us now return to the important questions raised above. First, is the use of fuzzy sets to model words scientific? No body of *data* collected from people who use a word seems to be available in the FS literature to test (validate) a MF for a word FS. Since the MF for a word is a representation for something for which no data are available, it cannot be subject to scientific scrutiny—testing. So, *at the testing level, using fuzzy sets to model words is not yet a scientific theory.* It behooves us, therefore, to make the use of fuzzy sets that model words scientific. To do this, we shall assume that type-1 MF data can be collected from people (more about this in Section II). We can then turn to the second important question, "Is this scientific theory correct?" That *this theory is an incorrect one* follows from the following line of reasoning: (1) A fuzzy set A for a term (word, phrase) is a well-defined type-1 FS $\mu_A(x)$ ($x \in X$) that is *totally certain* once all of its parameters are specified, (2) words mean different things to different people, and so are *uncertain*, and therefore, (3) it is a contradiction to say that something certain can model something that is uncertain[1]. In the words of Popper, associating the original type-1 FS with a word is a "conceivable event" that has provided a "counter-instance" that falsifies this approach to fuzzy sets as models for words.

Note that the abstract mathematical approach to fuzzy sets cannot be falsified, because it is mathematics and is not a scientific theory.

Referring again to the right-hand path in Fig 1, we see that two situations are possible: (1) We revert to type-2 FS models (reverting to a type-1 FS—dashed path 2—is illicit, as we explain below) that are associated with the abstract mathematical approach to fuzzy sets, or (2) we collect data from people—person-MFs—that reflect intra and inter-levels of uncertainties about a word and then define a FS in a new way as the collection of all such person MFs. In this paper we adopt the second approach, because it lets us directly include the uncertainties about words.

[1] Random uncertainty modeled using a probability model that is characterized by a probability density function can also have certain parameters. However, the results obtained from a probability model—realizations—are never certain. On the other hand, once the parameters of the MF of a type-1 fuzzy set are fixed, the activation of that set by a fixed value of x always gives the same result for that value of x.

In the past when we have applied type-1 (or type-2) fuzzy sets to specific applications we have also had access to "data," but such data has been *application-dependent data*, and is not person-MF data associated with the underlying word fuzzy sets. In these applications, we have used the *architecture*, e.g., of a rule-based fuzzy system, to solve a specific problem, and we almost always have begun by specifying the shapes of the MFs. Frequently, the parameters of the MFs are tuned using the application-dependent data. Application-dependent data, while very important, is quite different from person-MF data.

For a long time people (e.g., [1], [2]) have noted that when a FS is completely determined by specifying its shape and parameters, there is nothing uncertain about the FS, and that this is *paradoxical* with the word "fuzzy," which has the connotation of uncertainty. Anytime a FS is associated with a word this paradox will occur. As early as 1975, this paradox was recognized and Zadeh [14] proposed a type-2 FS as a way to model the uncertainties about a type-1 FS. In Section II we will demonstrate that type-2 FSs are essential to our new theory of fuzzy sets for words.

II. A New Beginning

This section presents a collection of premises and explorations into what they mean for the definition of a fuzzy set that describes a word (i.e., single word, phrase, term).

Premise 1: Words mean different things to different people, and are therefore uncertain. Uncertainty about a word is of two kinds: (1) *intra-uncertainty*, which is the uncertainty a person has about the word, and (2) *inter-uncertainty* which is the uncertainty that a group of people have about the word. ■

Premise 2: *Intra-uncertainty* about a word can be modeled using a *type-2 person fuzzy set*, $\tilde{A}(p_j)$, where $j = 1,...,n_{\tilde{A}}$. ■

The notation $\tilde{A}$ denotes a type-2 FS. Recall, (e.g., [5], [6]) that a type-2 FS is described by a 3-D MF. The first dimension is that of the *primary variable*, x, the second dimension is that of the *primary membership*, u, and the third dimension is that of the *secondary grade*. In a type-2 FS the primary membership is a finite range of values ($u \in [a,b] \subseteq [0,1]$), each element of which may be weighted differently by means of an assigned secondary grade. The secondary grades sit atop a region of the $x-u$ plane that, for continuous variables, is called the *footprint of uncertainty* (FOU). The FOU is very useful because it lets us visualize the 3-D type-2 FS MF in two-dimensions.

Premise 3: For person MF data collected from people we assume that all secondary grades equal 1. As a result, it is not necessary to explicitly carry around the third dimension of the secondary grades. We therefore use the terms MF and FOU interchangeably, with the implicit understanding that the actual person MF has a third dimension that sits atop of the person FOU where the height of the third dimension is equal to 1. ■

It will be very challenging to extract FOU information about a word from a person. To ask them to assign anything other than a uniform weighting to their entire FOU would, in the opinion of this author, be too difficult. We categorize the uncertainty that exists about the person FOU as [7] a *first-order kind of uncertainty*, and the uncertainty that exists about the weight that might be assigned to each element of the person FOU as a *second-order kind of uncertainty*. When a person provides their FOU for a word the first-order uncertainty is across their entire FOU. Clearly, weight information is itself uncertain leading to even higher (and never-ending) kinds of uncertainty. It will be difficult to impossible to *test the validity* of second-order uncertainty by collecting data; hence, in this paper, we focus exclusively on the first-order uncertainty of a person FOU.

Example 1: Let us assume that it is possible to collect person FOUs from individuals for the word *some*. Fig. 2 depicts three typical person FOUs. The constraints that each person must adhere to when sketching their FOU is that the upper bound cannot exceed 1 and the lower bound must not be less than 0. There may be a region for the primary variable x for which there is no uncertainty about the word, as is indicated by a flat spot in each person's FOU at which $u = 1$. ■

We use the notation $\mu_{\tilde{A}}(x \mid p_j)$ to denote a *type-2 person MF (FOU)* of word $\tilde{A}$ for person-j, where $j = 1,\ldots,n_{\tilde{A}}$; it is associated with the *person fuzzy set* $\tilde{A}(p_j)$, i.e.

$$\tilde{A}(p_j) = \int_{x\in X} \mu_{\tilde{A}}(x \mid p_j)/x \tag{1}$$

$$\mu_{\tilde{A}}(x \mid p_j) = \left[a_{\tilde{A}}(x \mid p_j), b_{\tilde{A}}(x \mid p_j)\right] \subseteq [0,1] \tag{2}$$

$a_{\tilde{A}}(x \mid p_j)$ and $b_{\tilde{A}}(x \mid p_j)$ denote lower and upper bounds, respectively of the person MF and [,] denotes an *interval set*, i.e. the set of all real numbers from $a_{\tilde{A}}(x \mid p_j)$ to $b_{\tilde{A}}(x \mid p_j)$.

Comment 1. For people who are familiar with type-2 fuzzy sets (e.g., $\tilde{A}(p_j)$), we are calling the union of *primary membership* ($J_x(p_j)$) of that set the type-2 person MF (FOU). ■

Premise 4: *Inter-uncertainty* about a word can be modeled by means of equally weighted aggregation of each person's word FS $\tilde{A}(p_j)$, $j = 1,\ldots,n_{\tilde{A}}$. This weight can be normalized out of the aggregation, so we choose it to be 1. ■

Space limitations do not let us describe two other aggregation possibilities: (1) people are treated differently, in which case we can assign a different weight to each person, and (2) the same as Case 1, except now we can assign a different weight *function* to each person. It is only equally weighted aggregation that lets us focus on first-order uncertainties.

Premise 5: The natural way to aggregate people's equally weighted word fuzzy sets is by the mathematical operation of the union, because it preserves the commonalities as well as the differences across person FSs. ■

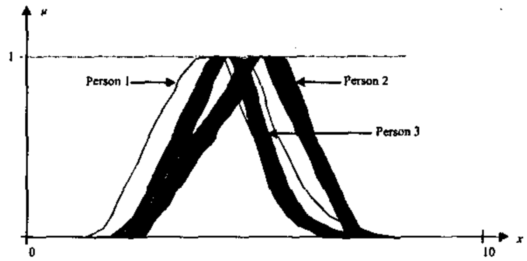

Fig. 2: FOUs from three people for the word *some*. Uniform shading indicates that all of the secondary grades equal 1.

The journal version of this paper will explain why the intersection and addition operations are unacceptable ways to aggregate person MFs.

Person-MF Representations of a Fuzzy Set: From Premise 5, (1) and (2), we obtain the following *new representations for a fuzzy set of a word*:

$$\tilde{A} = \bigcup_{j=1}^{n_{\tilde{A}}} \tilde{A}(p_j) = \bigcup_{j=1}^{n_{\tilde{A}}} \int_{x\in X} \mu_{\tilde{A}}(x \mid p_j)/x = \bigcup_{j=1}^{n_{\tilde{A}}} \int_{x\in X} \left[a_{\tilde{A}}(x \mid p_j), b_{\tilde{A}}(x \mid p_j)\right]/x$$

$$= \int_{x\in X} \left[\bigcup_{j=1}^{n_{\tilde{A}}} \mu_{\tilde{A}}(x \mid p_j)\right]/x = \int_{x\in X} \left[\bigcup_{j=1}^{n_{\tilde{A}}} \left[a_{\tilde{A}}(x \mid p_j), b_{\tilde{A}}(x \mid p_j)\right]\right]/x \tag{3}$$

From the second line of (3), which is called the **vertical-slice representation** of $\tilde{A}$ (it focuses on the primary memberships at each value of primary variable x), it is clear

$$\mu_{\tilde{A}}(x) = \bigcup_{j=1}^{n_{\tilde{A}}} \mu_{\tilde{A}}(x \mid p_j) \quad \forall x \in X \tag{4}$$

This new description of a fuzzy set for a word permits us to include both the intra- and inter-uncertainties that people have about the word. It also lets us easily add or remove person fuzzy sets, as desired.

Examining the three person FOUs in Fig. 2, it is clear that they are each upper and lower bounded. Let ($x \in X$)

$$\underline{\mu}_{\tilde{A}}(x) \equiv \min_{j=1,2,\ldots,n_{\tilde{A}}} a_{\tilde{A}}(x \mid p_j) \tag{5}$$

$$\bar{\mu}_{\tilde{A}}(x) \equiv \max_{j=1,2,\ldots,n_{\tilde{A}}} b_{\tilde{A}}(x \mid p_j) \tag{6}$$

We refer to $\underline{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{A}}(x)$ as the *lower* (bound) and *upper* (bound) MF values of $\tilde{A}$ at *any* $x \in X$, respectively. For $\forall x \in X$, we let $\underline{\tilde{A}}$ [or $LMF(\tilde{A})$] and $\bar{\tilde{A}}$ [or $UMF(\tilde{A})$] denote the lower and upper MFs for $\tilde{A}$, i.e.

$$\underline{\tilde{A}} \equiv \int_{x\in X} \underline{\mu}_{\tilde{A}}(x)/x \equiv LMF(\tilde{A}) \tag{7}$$

$$\bar{\tilde{A}} \equiv \int_{x\in X} \bar{\mu}_{\tilde{A}}(x)/x \equiv UMF(\tilde{A}) \tag{8}$$

Note that the lower and upper MFs are *type-1* (bounding) FSs.

Example 1 (Continued): Beginning with the three person MFs for the word *some*, we can easily establish the lower and upper MFs. They are depicted in Fig. 3. ■

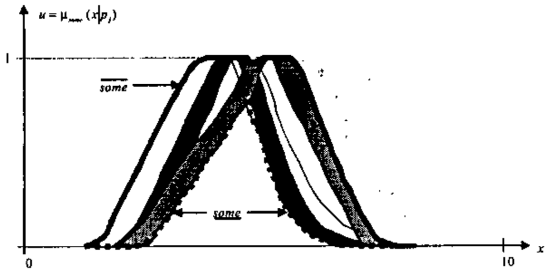

Fig. 3: Lower and upper (bound) MFs for the three person-MFs depicted in Fig. 2.

Our model in (3) contains all of the information in the constituent person-MFs. Although it is possible to develop formulas for performing set theoretic operations for fuzzy sets modeled by (3) we shall not do that, because this model is not a *parsimonious* one. It can result in huge numbers of computations and enormous memory requirements. Instead, we shall develop parametric models for word fuzzy sets.

III. Parametric Models

In the theory of modeling, e.g., system identification (e.g., [3], [11]) a tradeoff is always made between preserving all of the data (information) and achieving a useful and parsimonious model. We have now reached that same fork in the road for the fuzzy set of a word. We shall follow the widely used approach taken in modeling theory of approximating the data by means of *parametric models*. Parsimony is achieved by choosing a model with the smallest number of parameters that best approximates the data (in some sense).

Fig. 2 already suggests that by including more and more person MFs a region will become *filled in* within which all person MFs may lay. Let $\tilde{A}_{FI}$ denote the *filled-in word fuzzy set*. Its associated FOU—the *filled-in FOU*—is denoted $FOU(\tilde{A}_{FI})$. A filled-in FOU for the example in Fig. 3—$FOU(some_{FI})$—is depicted in Fig. 4. Observe the filled-in FOU is bounded by $\overline{some}$ and $\underline{some}$, as in Fig. 3.

Let $J_x(\tilde{A}_{FI})$ denote the **primary membership** of x for filled-in word fuzzy set $\tilde{A}_{FI}$. Because of fill-in, $J_x(\tilde{A}_{FI})$ is now the continuous interval

$$J_x(\tilde{A}_{FI}) = \left[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)\right] \quad \forall x \in X \tag{9}$$

Using (9), we can express $FOU(\tilde{A}_{FI})$ mathematically as

$$FOU(\tilde{A}_{FI}) = \int_{x \in X} J_x(\tilde{A}_{FI}) \big/ x \tag{10}$$

Using (10) and (9), we postulate the following representation for a *filled-in word fuzzy set*:

$$\tilde{A}_{FI} = \int_{x \in X} J_x(\tilde{A}_{FI}) \big/ x = \int_{x \in X} \left[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)\right] \big/ x \tag{11}$$

It is important to understand that the upper and lower MF values used in (11) are obtained directly from the person MFs through a bounding procedure. It should be clear, by comparing (11) with (7) and (8), that $LMF(\tilde{A}_{FI}) = LMF(\tilde{A})$ and $UMF(\tilde{A}_{FI}) = UMF(\tilde{A})$.

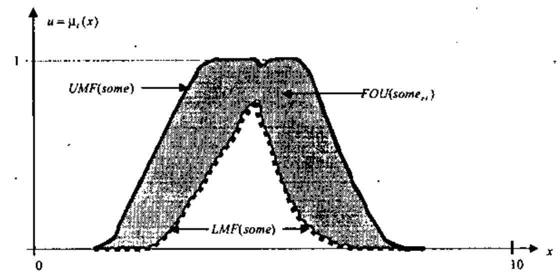

Fig. 4: Filled-in FOU for *some* ($\tilde{s}$)

Premise 6: A *filled-in parametric model*, $\hat{\tilde{A}}$ ($\hat{A}$ for short), for $\tilde{A}_{FI}$ in (11) is one that is described by two functions, $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$. These functions have shapes chosen ahead of time (e.g., triangle, Gaussian, trapezoidal, etc.) that are characterized by a small number of parameters, some or all of which may be shared by both $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$. These parameters are fixed during some sort of design procedure. ■

$\underline{\mu}_{\hat{A}}(x)$ is an approximation to $LMF(\tilde{A}_{FI})$, and $\bar{\mu}_{\hat{A}}(x)$ is an approximation to $UMF(\tilde{A}_{FI})$. If no parameters are shared by $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$, then two independent least-squares approximation problems can easily be established for determining the parameters of $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$. The approximation problem for $\underline{\mu}_{\hat{A}}(x)$ would only use the data in $LMF(\tilde{A}_{FI})$ and is unconstrained, whereas the approximation problem for $\bar{\mu}_{\hat{A}}(x)$ would only use the data in $UMF(\tilde{A}_{FI})$ but is constrained so that $\bar{\mu}_{\hat{A}}(x) = 1$ at least at one value of x where $\bar{\mu}_{\tilde{A}}(x) = 1$, so that the resulting type-1 upper-bound FS is a *normal* FS. This constraint can automatically be satisfied by an appropriate choice for the shape of $\bar{\mu}_{\hat{A}}(x)$, and so *one does not have to solve a constrained optimization problem.*

Example 2: Examining Fig. 4, we could use a trapezoidal function (not necessarily symmetrical, and characterized by four parameters) to approximate $UMF(\tilde{A}_{FI})$, and a triangular function (characterized by three parameters) to approximate $LMF(\tilde{A}_{FI})$. Because of its built-in flat top, the parameters of this upper MF approximation can be obtained by solving an unconstrained optimization problem. Seven parameters are needed in this parametric model. ■

If even one parameter is shared by $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$ then one least-squares approximation problem can be established for *simultaneously* determining the parameters of $\underline{\mu}_{\hat{A}}(x)$ and

$\bar{\mu}_{\hat{A}}(x)$, one that uses the data in both $LMF(\tilde{A}_{FI})$ and $UMF(\tilde{A}_{FI})$.

Example 3: Examining Fig. 4 again, we could also approximate its FOU using a Gaussian *primary* MF having an uncertain mean with values in $[m_1, m_2]$ and an uncertain standard deviation with values in $[\sigma_1, \sigma_2]$, i.e.

$$\mu_{\hat{A}}^{p}(x) = \exp\left[-\tfrac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \quad m \in [m_1, m_2] \text{ and } \sigma \in [\sigma_1, \sigma_2] \tag{12}$$

Equations for $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$ can be found in [5, pp. 94-95]. Note that $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$ share all four of the parameters in (12). It is easy to show that $\bar{\mu}_{\hat{A}}(x)$ has a built-in flat top between m_1 to m_2, so that all four of the parameters in $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$ can be obtained by solving an unconstrained optimization problem. Observe, also, that this parametric model is characterized by four parameters, whereas the one in Example 2 is characterized by seven parameters; hence, the present parametric model is a more parsimonious one. ■

Just as the filled-in word fuzzy set $\tilde{A}_{FI}$ has a FOU associated with it, namely $FOU(\tilde{A}_{FI})$ [see (10) and (11)], $\hat{A}$ also has a FOU associated with it, namely $FOU(\hat{A})$, i.e.

$$FOU(\hat{A}) = \int_{x \in X} \left[\underline{\mu}_{\hat{A}}(x), \bar{\mu}_{\hat{A}}(x)\right] \Big/ x \tag{13}$$

The closer that $\underline{\mu}_{\hat{A}}(x)$ and $\bar{\mu}_{\hat{A}}(x)$ approximate $\underline{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{A}}(x)$ over $\forall x \in X$, the closer $FOU(\hat{A})$ approximates $FOU(\tilde{A}_{FI})$.

Adding or removing person MFs (i.e., increasing or decreasing $n_{\tilde{A}}$) will very likely cause $\underline{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{A}}(x)$ to change, so if this is done $FOU(\hat{A})$ would have to be recomputed.

In summary, what we have done is to go from $\tilde{A}$, the union of person FSs, to $\tilde{A}_{FI}$, the fill-in of all points between the lower and upper bounds of the FOU of $\tilde{A}$, and finally to $\hat{A}$ a parsimonious approximation to the word fuzzy set. This approximation uses data that are obtained from people.

IV. Conclusions

We began this paper with a careful delineation of two approaches to fuzzy sets, *abstract mathematics* and *models for words*. We then demonstrated, by using Karl Popper's Falsificationism, that the present approach to fuzzy sets for words (e.g., as in *computing with words*) is not scientific because its type-1 FSs are not testable, since word MF data is not available from people. In order to make fuzzy sets for words scientific (testable) we assumed that Type-1 MF data could be collected from people. We then demonstrated that such testable word MFs are scientifically incorrect, because words are uncertain whereas type-1 FSs are certain.

Our new theory of fuzzy sets for words is based on collecting data from people—person MFs—that reflect *intra-* and *inter-levels of uncertainties* about a word, and then defining a FS in a new way, as the union of all such person FSs. Our new theory is built on six premises the most important ones of which are: *Intra-uncertainty* about a word can be modeled using an *interval type-2 person fuzzy set*, $\tilde{A}(p_j)$, where $j = 1,...,n_{\tilde{A}}$, and *inter-uncertainty* about a word can be modeled by means of an equally weighted aggregation of each person's FS, $\tilde{A}(p_j)$, $j = 1,..., n_{\tilde{A}}$.

We then argued why aggregation by the union operator is most appropriate. This then led to the new person MF representation of a word FS, given in (3).

Because the new person MF representation of a word FS preserves all of the MF data that is collected from people it is *not* a parsimonious model for word fuzzy sets. This led us to go from $\tilde{A}$, the union of person FSs, to $\tilde{A}_{FI}$, the fill-in of all points between the lower and upper bounds of the FOU of $\tilde{A}$, and finally to $\hat{A}$, a parsimonious parametric approximation for the word fuzzy set. It is the upper and lower bounds of the FOU of $\tilde{A}$, which are determined by the upper and lower bounds of person MF FOUs, that play the central role in obtaining $\hat{A}$.

Returning to Fig.1, let us consider three additional points:

1. What if it is not possible to collect person MFs, or doing so is considered to be too large an effort, *can we still perform computing with words*? Fig. 1 indicates that we could move from its right-hand path to the left-hand abstract mathematical path but only by using type-2 fuzzy sets, because they will let us model uncertainties about a word, whereas type-1 fuzzy sets will not. So, even if person MFs are unavailable we can conceptually think about performing computing with words but we must use type-2 fuzzy sets. However, according to Popper, to do so without data is not scientific. So, if person MF data is unavailable some other kind of data must be available or else this approach to computing with words is not testable, and is therefore not scientific.

   An analogy with adaptive control is appropriate at this point. One kind of adaptive control relies on an *explicit identification* of the plant being controlled. To do this, data about the plant must be available, which is analogous to obtaining person MFs. In another kind of adaptive control—*performance* adaptive control—no explicit identification is necessary. Instead, control parameters are adjusted so that the system achieves a desirable measurable output performance. This latter situation is very much like what we do in the training of a feed-forward neural network or in the tuning of the MF parameters of a rule-based FLS, when training and testing data are available.

   So, if person MF data are unavailable then (computing with words) training and testing data will have to be available so that the parameters of the type-2 fuzzy sets that are used in a system that computes with words can be optimized in order that such a system performs at an acceptable level. To the best knowledge of this author even such data has not been reported on in the literature.

2. Clearly, computing with words must be done using the mathematics of type-2 fuzzy sets. Because we have argued that a word FS is an interval type-2 FS, the mathematics of such fuzzy sets (e.g., union, intersection, complement) are quite simple (e.g., [6]).
3. Type-1 fuzzy sets are conspicuously absent from the right-hand path of Fig. 1; so, if students are taught about fuzzy sets by using word examples—in which case they must traverse this path—where would they learn about type-1 FSs? Mendel and John [6] have recently developed a new representation of a type-2 FS as the union of its embedded type-2 fuzzy sets. The domain of an embedded type-2 FS is a type-1 FS. It is easy to derive formulas for union, intersection and complement of type-2 fuzzy sets using this new representation. The derivation is even simpler for interval type-2 fuzzy sets. But, even for them, we still need the concept of a type-1 FS. So, perhaps an appropriate time to focus on type-1 fuzzy sets would be when deriving set theoretic operations for type-2 FSs.

   Another approach would be to consider the very special case when all intra-uncertainties disappear, for which each person MF reduces to a function—a type-1 FS —and the fuzzy set for a word becomes the union of type-1 fuzzy sets[2]. Although it is easy to rationalize this approach, in order to reach the mathematics of type-1 fuzzy sets quickly, it is not an approach that is supported by the uncertain data, which again explains why type-1 fuzzy sets now only appear in the left-hand abstract mathematical path of Fig. 1.

Our new theory of fuzzy sets for words is testable and is therefore subject to refutation. The weak point of this new theory is the assumption of person MFs, but this seems to be as fundamental an assumption as a molecule is to matter.

Our new definition of a word fuzzy set suggests the creation of a new field for fuzzy sets—*experimental fuzzy sets*[3]—one in which data are collected about person MFs and related issues are formulated and tested (e.g., [2, Ch. 10]). For example, can context-independent results be applied when words (e.g., *some*, *a large amount*, etc.) are used within a specific context? More specifically, does the uncertainty captured by collecting person MFs from a large group of people permit us to use context-independent MFs in different contexts, because people provide their MFs based on their *own personal contextual interpretation* of each word, so that context is actually implicitly contained across the collection of person MFs? Finally, is there a way to map from context-independent MFs to context-dependent MFs (i.e., when context is revealed to people, our conjecture is that there is less uncertainty about a word than if context is not revealed to them, because context causes people to provide their MF based on a specific contextual interpretation of a word, so that their MF becomes more focused, i.e. less uncertain)?

We do not claim that the material in this paper provides a complete solution to the important problem of computing with words. We do, however, claim that such a solution needs to account for the uncertainties about the words that are used in computing with words, and that our new approach to fuzzy sets provides a scientific (in the sense of Popper) framework to do this. Many issues remain to be resolved before computing with words can become a reality. Hopefully, this paper will provide a starting point for that work.

[2] When this paper was originally submitted for review, it did not include intra-uncertainties, and only focused on inter-uncertainties. Many of its conclusions have changed as a result of immediately starting with both intra- and inter-uncertainties.

[3] Analogizing with probability and its experimental counterpart, statistics, we might call this *fuzzistics*.

ACKNOWLEDGEMENTS

The author would like to acknowledge the many discussions (by electronic mail) he has had about the ideas in this paper with Prof. Robert I. John. They helped to make very dramatic changes in this version of the paper from its first version.

REFERENCES

[1] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[2] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall, Upper Saddle River, NJ, 1995.

[3] L. Lung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

[4] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proc. of Third International ICSC Symposium on Fuzzy Logic and Applications*, Rochester Univ., Rochester, NY, June 1999.

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

[6] J. M. Mendel and R. I. Bob John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 117-127, April 2002.

[7] J. M. Mendel and R. I. Bob John, "Footprint of uncertainty and its importance to type-2 fuzzy sets," *Proc. 6th IASTED Int'l. Conf. on Artificial Intelligence and Soft Computing*, Banff, Canada, July 2002, pp. 587-592.

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

[9] K. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, London, 1963.

[10] K. Popper, in *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/popper.

[11] T. Soderstrom and P. Stoica, *System Identification*, Prentice-Hall International, Ltd., New York, 1989.

[12] P. P. Wang (Editor), *Computing With Words*, John Wiley & Sons, Inc., New York, 2001.

[13] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[14] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1," *Info. Sciences*, vol. 8, pp. 199-249, 1975.

[15] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. on Fuzzy Systems*, vol. 4, pp. 103-111, 1996.

[16] L. A. Zadeh, "From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions," *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 4, pp. 105-119, 1999.