Jerry M. Mendel
University of Southern California, USA

Computing with Words:
Zadeh, Turing, Popper and Occam

Abstract: In this article, after explaining Zadeh's computing with words (CWW) paradigm, I argue that for this paradigm to be embraced, it must be validated using a Turing-like test, use a scientifically correct fuzzy set model for words, namely interval type-2 fuzzy sets (IT2 FSs), and be simple, meaning that fuzzy set operations should be as simple as possible. These conclusions are drawn using the ideas of Turing, Popper and Occam. Short descriptions are provided for a Perceptual Computer (Per-C), which is an architecture for CWW for making subjective judgments, IT2 FSs, IT2 FS models for words, and why an IT2 FS model captures first-order uncertainties about a word. Short biographies of Zadeh, Turing, Popper and Occam are also provided.

Digital Object Identifier 10.1109/MCI.2007.9066897

 1556-603X/07/$25.00©2007IEEE

1. Introduction

In 1996 Zadeh [45] published a paper that first equated fuzzy logic (FL) with *computing with words* (CWW or CW). Of course, he did not mean that computers would actually compute using words, a single word or phrase, rather than numbers. He meant that computers would be activated by words, which would be converted into a mathematical representation using fuzzy sets (FSs) (e.g., [43], [44]), and that these FSs would be mapped by means of a CWW engine into some other FS, after which the latter would be converted back into a word (Figure 1).

Since the publication of Zadeh's CWW paper, there have been many articles and even books bearing the phrase "computing with words." A small sampling of these are [12], [13], [15]–[18], [21], [39], [46], and [47]. A related reference is [38]. FL is viewed in these publications as the machinery that will let "input" words, which are provided by a human, be transformed within the computer to "output" words, which are provided back to that human, or to other humans. One instantiation of the CWW Paradigm, a *Perceptual Computer*, is described in Box 1. Potential applications for CWW are many and include Web-based searches, summarizations, subjective judgments, subjective decisions, etc. Box 2 describes how the social judgment of flirtation can be addressed using a Perceptual Computer. It is easy to extend this light-hearted application to many other judgment situations.

A person interacting with such a CWW computer interface would not be concerned with the CWW engine, but would only be interested in knowing the output word for their input word. On the other hand, the designer of the CWW engine would address some very fundamental questions, including:

1. How can a CWW engine be validated?
2. What FS models should be used?
3. What choices should be made to keep the design of the CWW engine as simple as possible?

The purpose of this article is to provide answers to these questions, answers that are built upon the shoulders of three critical thinkers, Alan Turing, Karl Popper and William of Occam, and that no doubt will provoke readers into arguing against my answers or suggesting other answers. Biographies of Zadeh, Turing, Popper and Occam appear at the end of the article.

2. How Can A CWW Engine be Validated?

It is my belief that for the CWW Paradigm to be successful it must provide end-users with results that are equivalent to those from a human. This is in spirit with what the great computer scientist and philosopher Alan Turing [5], [33] proposed as a test, the *Turing Test*, for machine intelligence. I believe that this test is as applicable to CWW as it is to machine intelligence, because CWW is a form of artificial intelligence.

Paraphrasing [5], consider an "imitation game" played with three players, a human being, a machine and an interrogator. The interrogator stays in a room apart from the others. The object is for the interrogator to determine which of the others is the human being or the machine. If the machine cannot be distinguished from the human being under these conditions, then we must credit it with human intelligence. This is the essence of a Turing Test.

According to the marvelous article by Saygin, et al. [31], "The Turing Test is one of the most disputed topics in artificial intelligence and cognitive sciences," because it can be interpreted in many different ways, e.g., by philosophers, sociologists, psychologists, religionists, computer scientists, etc. I am not interested in using a Turing Test to establish whether the Perceptual Computer can think, to demonstrate that it is intelligent. I am interested in using a Turing Test, as explained in [31, p. 467]: "Alternatively, the TT [Turing Test] for machine intelligence can be re-interpreted as a test to assess a machine's ability to pass for a human being."

In order to implement the CWW Paradigm, data will be needed. This data must be collected from people who are similar to those who will ultimately be interacting with the CWW engine. If such data collection is feasible, then the design of a CWW engine can proceed using the data in the usual way, i.e., by using some of it for training[1], some for testing, and the rest for validation. The validation of the designed CWW computer using some of the collected data (the validation set) can be interpreted as a Turing Test.

If, on the other hand, such data collection is not feasible, then the designer of the CWW engine must fabricate it, or even worse, design the CWW engine using no data at all. After such a design, the CWW engine will have to be validated by testing it on a group of subjects, and such a test will again constitute a Turing Test.

Hence, one way or another, validation of a CWW design is by a Turing Test.

3. What Fuzzy Set Models Should be Used?

Because words can mean different things to different people, it is important to use an FS model that lets us capture word uncertainties. At present, there are two possible choices, a type-1 (T1) FS or an interval type-2 (IT2) FS[2] [16], [20], [22]

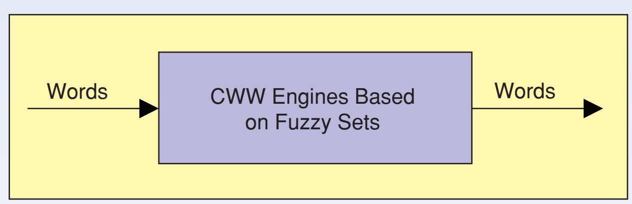


FIGURE 1 The CWW Paradigm.

[1] When the CWW Engine (Figure 2) is a set of IF-THEN rules, then a *training set* can be used to optimize the parameters of antecedent and consequent membership functions, to establish the presence or absence of antecedent terms, and to even determine the number of significant rules, after which the optimized rules can be tested using a *testing set*.

[2] General type-2 FSs are presently excluded, because they model higher degrees of uncertainty, and how to do this is totally unknown.

For the CWW Paradigm to be successful it must provide the end-users with results that cannot be distinguished from those they would have received from a human.

(see Box 3). In order to decide which model to use, two different approaches can be taken:

1. Use a T1 FS, design a CWW engine, and see if it passes a Turing Test. If it does, then it is okay to use such an FS model.
2. Choose between a T1 FS model and an IT2 FS model before designing the CWW engine. Then design the CWW engine and see if it passes a Turing Test.

The first approach needs no further discussion. Regarding the second approach, in order to choose between a T1 FS and an IT2 FS model, we shall rely on the great 20th century scientific philosopher, Sir Karl Popper, who proposed *falsificationism* [29], [30], [32] as a way to establish if a theory is or is not scientific. Falsificationism states: "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."

According to [32], by falsifiability, Popper meant "if a theory is incompatible with possible empirical observations it is scientific; conversely, a theory which is compatible with all such observations, either because, as in the case of Marxism, it has been modified solely to accommodate such observations, or because, as in the case of psychoanalytic theories, it is consistent with all possible observations, is unscientific."

For a theory to be called *scientific* it must be *testable*. This means that it must be possible to make measurements that are related to the 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 itself subject to refutation at a later date.

We suggest that using either a T1 FS or an IT2 FS as a word model can be interpreted as a scientific theory.[3] We must therefore question whether each FS word model qualifies as a scientific theory, and then if each is a correct or incorrect scientific theory.

Data collection and mapping into the parameters of a T1 MF has been reported on by a number of authors (e.g., [11]) but has only started to be researched for a T2 MF [14], [24]–[26] (see Box 4). Names for the different T1 methods include: polling ([4], [12]), direct rating ([11], [27], [28]), reverse rating ([28], [34]–[37]), interval estimation ([1]–[3], [48]), and transition interval estimation [1]. The term *fuzzistics* has been coined [19], [21] for doing this, and represents an amalgamation of the words *fuzzy* and *statistics*.

That using a T1 FS model for a word is an incorrect scientific theory follows from the following line of reasoning [19]: (1) A T1 fuzzy set A for a word is well-defined by its MF $\mu_A(x)(x \in X)$ 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

[3]Note that this is very different from T1 FSs and IT2 FSs as mathematics, which are not scientific theories, and about which we should not take any issues.

Box 1. Perceptual Computer

A specific architecture for CWW using interval type-2 fuzzy sets (IT2 FSs) (see Box 3 for a brief description of such fuzzy sets) was proposed in [18], called a *Perceptual Computer*–Per-C for short. The Per-C consists of three components (Figure 2): encoder, CWW engine and decoder. Perceptions (i.e., granulated terms, words) activate the Per-C and are the Per-C output; so, it is possible for a human to interact with the Per-C using just a vocabulary.

For each application a vocabulary must first be established, one that lets the end-user interact with the Per-C in a user-friendly manner. The encoder transforms linguistic perceptions into IT2 FSs and leads to a codebook of words with their associated IT2 FS models. The outputs of the encoder activate a CWW engine, whose output is another IT2 FS, which is then mapped by the decoder into a word most similar to a vocabulary word.

Box 4 describes the transformation of linguistic perceptions into IT2 FSs (the encoding problem).

The CWW Engine may take the form of IF-THEN rules, e.g. [11], [16], a linguistic weighted average [40], linguistic summarizations [6]–[10], [42], etc., for which the established mathematics of fuzzy sets provides the transformation from the input FSs to the output FSs.

[41] discusses mapping an IT2 FS into a word (the decoding problem).

Box 2 explains how the Per-C can be used in an application of making subjective judgments.

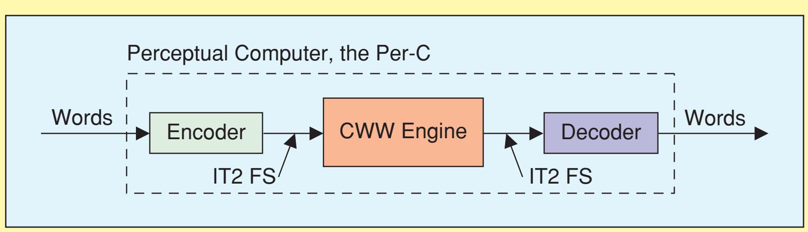


FIGURE 2 Specific architecture for CWW—the Perceptual Computer.

contradiction to say that something certain can model something that is uncertain. In the words of Popper, associating the original T1 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.

An IT2 FS is a scientifically correct first-order uncertainty model for a word.

An IT2 FS model for a word only lets us model *first-order uncertainties* (see Box 5) as is clearly visible from its FOU (Figure 3); hence, an IT2 FS is a scientifically correct first-order uncertainty model for a word; and, in the future the scientific correctness of an IT2 FS model may be falsified by a more complete T2 FS model because measurements can be made about words.

An objection may be raised that a fixed T1 MF also applies to an IT2 FS model, i.e. once the parameters of an IT2 FS model are specified there no longer is anything uncertain about the IT2 FS. This objection is incorrect because the IT2 FS is a first-order uncertainty model, i.e. at each value of the primary variable (at x' in Figure 3) the MF is an interval of values. For a T1 FS the MF is a point value, and it is the interval nature of the MF that provides uncertainty to the IT2 FS model. This argument is similar to one that can be given for a probability distribution function. Once we agree that such a function does indeed model unpredictable (random) uncertainties then fixing its parameters does not cause us to conclude that it no longer is a probability model.

One may argue that a T1 FS model for a word is a model for a *prototypical word*; however, if one also believes that *words mean different things to different people*, then this calls into question the concept of a prototypical word.

When random uncertainties are present, most of us have no problem with using probability models and analyses from the very beginning; hence, when linguistic uncertainties are present, I suggest that we must have no problem with using IT2 FS models and analyses from the very beginning. Some may ask the question "How much linguistic uncertainty must be present before I need to use an IT2 FS?" Maybe, in the very early days of probability a similar question was asked; however, it no longer seems to be asked. When randomness is suspected we use probability. So, I propose that when linguistic uncertainties are suspected we use IT2 FSs.

Box 2. The Per-C for Making Social Judgments

In everyday social interaction, each of us makes judgments about the meaning of another's behavior (e.g., kindness, generosity, flirtation, harassment, etc.). Such judgments are far from trivial, because they often affect the nature and direction of the subsequent social interaction and communications. Although a variety of factors may enter into our decision, behavior (e.g., touching, eye contact) plays a critical role in assessing the level of the variable of interest.

Suppose the variable of interest is flirtation and the only indicator of importance is eye contact. The following *user friendly vocabulary* is established for both eye contact and flirtation: none to very little, a small amount, a little bit, some, a moderate amount, a good amount, a considerable amount, a large amount, a lot, and an extreme amount. Surveyed respondents are asked a question such as: "On a scale of zero to ten where would you locate the end-points of an interval for this word?" These data are then mapped by means of the encoder into a footprint of uncertainty, FOU (see Box 3), for each word. The ten words and their FOUs constitute the *codebook* for the subjective judgment of flirtation and for eye contact.

A small set of five rules (the CWW engine) is then established, using a subset of five of the ten words, e.g. none to very little, some, a moderate amount, a large amount, and an extreme amount. One such rule is "IF eye contact is a moderate amount, THEN the level of flirtation is ________?" Another survey is conducted in which respondents choose one of these five flirtation terms for each rule (i.e., for the rule's consequent). Because all respondents do not agree on the choice of the consequent, this introduces uncertainties into the CWW engine.

An end user can interact with the *Flirtation Adviser Per-C* by inputting any one of the ten words from the codebook for a specific flirtation scenario. Rules are fired using the mathematics of IT2 FSs [22], the result being fired-rule IT2 FSs, i.e. an FOU for each fired rule. These FOUs are aggregated into a composite FOU that is then compared to the word-FOUs in the codebook. This comparison is done using fuzzy set similarity computations [41], the result being the word that best describes the flirtation state to the end user.

Such a flirtation adviser can be used to train a person to better understand the relationship between eye contact and flirtation, so that they reach correct conclusions about such a social situation.

Of course, in this very simple example of only one flirtation indicator not much confusion can occur; however, when more indicators are used (e.g., eye contact and touching) then in an actual social situation it is possible to get "mixed signals," i.e. a certain level of touching may indicate a large amount of flirtation, whereas a certain level of eye contact may indicate none to very little flirtation. So which is it? In this case, more than one rule will fire and the totality of fired rule FOUs is an indicator of what is meant by "mixed signals." By aggregating the fired rule FOUs and comparing the resulting FOU to the word-FOUs in the codebook the result will again be the word that best describes the flirtation state to the end user. In this way, the flirtation adviser can be used to train a person to reach correct conclusions about social situations when he or she is receiving mixed signals.

Finally, even a CWW engine that is designed using IT2 FSs needs to be validated by a Turing Test. The difference in this second approach is that we begin the design using an FS word model that is scientifically correct. This, in itself, does not mean that the resulting CWW engine will pass a Turing Test, because that test is applied to the outputs of the CWW engine, and it is (Figure 2) the combination of a scientifically correct FS input word model, the CWW engine, and a good decoder that leads to the word-output.

4. What Choices Should be Made to Keep the Design of the CWW Engine as Simple as Possible?

Many choices have to be made when designing the CWW engine. For example, if the CWW engine is a set of IF-THEN rules, then choices must be made about:

❑ Shapes of lower and upper MFs for each FOU (e.g., triangles, trapezoids, Gaussians).
❑ Mathematical operators used to model the antecedent connector words *and* and *or*. Such operators are called t-norms

Box 3. Interval Type-2 Fuzzy Sets (The material in this Box is taken from [22])

A T1 FS has a grade of membership that is crisp; whereas, a T2 FS has grades of membership that are fuzzy, so it could be called a "fuzzy-fuzzy set." Such a set is useful in circumstances where it is difficult to determine the exact membership function (MF) for an FS, as in modeling a word by an FS.

As an example [20], suppose the variable of interest is *eye contact*, denoted x, where $x \in [0, 10]$ and this is an intensity range in which 0 denotes no eye contact and 10 denotes maximum amount of *eye contact.* One of the terms that might characterize the amount of perceived eye contact [e.g., during flirtation (see Box 2)] is "some eye contact." Suppose that 50 men and women are surveyed, and are asked to locate the ends of an interval for *some eye contact* on the scale of 0–10. Surely, the same results will not be obtained from all of them because words mean different things to different people.

One approach for using the 50 sets of two end-points is to average the end-point data and to then use the average values to construct an interval associated with *some eye contact*. A triangular (other shapes could be used) MF, $MF(x)$, could then be constructed, one whose base end-points (on the x-axis) are at the two end-point average values and whose apex is midway between the two end-points. This T1 triangular MF can be displayed in two dimensions (Figure 3). Unfortunately, it has completely ignored the uncertainties associated with the two end-points.

A second approach is to make use of the average end-point values and the standard deviation of each end-point to establish an uncertainty interval about each average end-point value. By doing this, we can think of the locations of the two end-points along the x-axis as blurred. Triangles can then be located 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, as in Figure 3. For purposes of this discussion, suppose there are exactly N such triangles. Then at each value of x, there can be up to N MF values (grades), $MF_1(x), MF_2(x), \ldots, MF_N(x)$. Each of the possible MF grades has a weight assigned to it, say $w_{x1}, w_{x2}, \ldots, w_{xN}$ (see the insert in Figure 3). These weights can be thought of as the *possibilities* associated with each triangle's grade at this value of x. Consequently, at each x, the collection of grades, called *secondary grades*, is a function $\{(MF_i(x), w_{xi}), i = 1, \ldots, N\}$ (called *secondary MF*). The resulting T2 MF is 3-D.

It is not as easy to sketch 3-D figures of a T2 MF as it is to sketch 2-D figures of a T1 MF. Another way to visualize a T2 FS is to sketch (plot) its FOU on the 2-D domain of the T2 FS. The heights of a T2 MF (its *secondary grades*) sit atop its FOU. In Figure 3, if the continuum of triangular MFs is filled in (as implied by the shading), then the FOU is obtained. The uniform shading over the entire FOU means that uniform weighting (possibilities) is assumed. Because of the uniform weighting, this T2 FS is called IT2 FS. Observe that the FOU is completely specified by two type-1 MFs, the *lower and upper MFs*, $LMF(\tilde{A})$ and $UMF(\tilde{A})$, also shown on Figure 3. The FOU is the 2-D region between $LMF(\tilde{A})$ and $UMF(\tilde{A})$.

For more information about IT2 FSs, see [22].

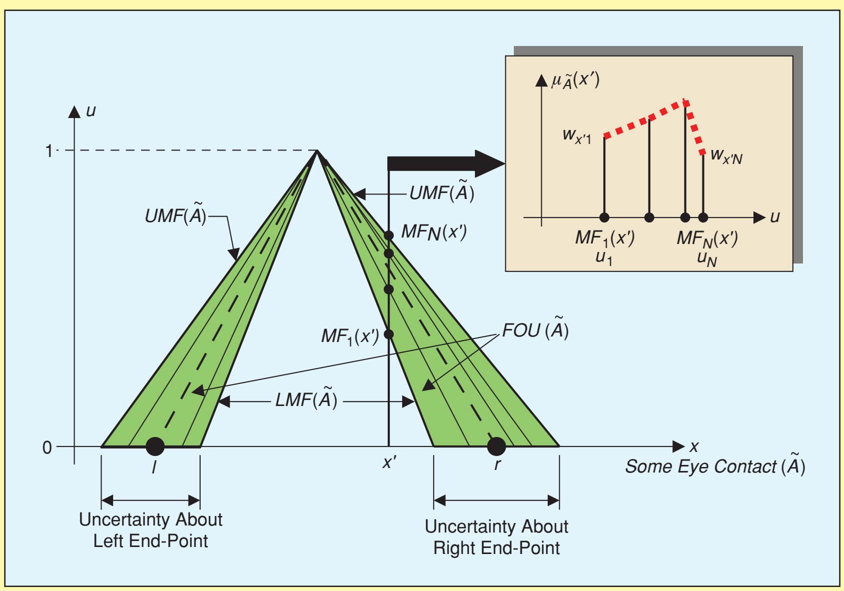


FIGURE 3 Triangular MFs when base end points (l and r) have uncertainty intervals associated with them. The insert depicts the secondary MF at x'. The shaded region is the FOU of T2 FS $\tilde{A}$.

Box 4. IT2 FS Models for Words

In order to establish IT2 FS models for words, one needs to collect data about words from a group of subjects and to then map that data into an IT2 FS. Different approaches for doing this are described below. They map data collected from subjects into a parsimonious parametric model of an FOU, and illustrate the combining of fuzzy sets and statistics—*type-2 fuzzistics*.

In the *person MF approach* [21]:

1. Person MF data [a person MF is an FOU that a person provides on a prescribed scale for a primary-variable (e.g., pressure, temperature)] are collected from a group of subjects. This data reflects both the intra- and inter-levels of uncertainties of the group about a word;
2. An IT2 FS model for a word is defined as a specific aggregation of all such person MFs (e.g., their union); and,
3. This aggregation is mathematically modeled and approximated.

Observe that mathematical IT2 FS models are only used at the very end of this approach.

Person MFs can only be collected from people who are already very knowledgeable about an FS, and may therefore be quite limited.

In the *interval end points* [21], [25] and *interval* [14] *approaches*:

1. Interval end-point data about a word are collected from a group of subjects;
2. Statistics (mean and standard deviation) are established for the data; and,
3. Those statistics are mapped into a pre-specified parametric FS model.

These methods are analogous in statistical modeling to first choosing the underlying probability distribution (i.e., data-generating model) and then fitting the parameters of that model using data and a meaningful design method, e.g., the method of maximum-likelihood.

and t-conorms, respectively, and there are many t-norms and t-conorms to choose from (e.g., [11]).

❑ Implication operators (an IF-THEN rule is mathematically modeled using an implication operator), and there are many such operators (e.g., [11]).

❑ How to aggregate fired rules, i.e., when more than one rule is fired, rule outputs must be combined (aggregated), and there are many different ways to do this (e.g., [11]). The result is an aggregated FOU.

❑ How to go from the aggregated FOU to a word, i.e., the decoder design in which a similarity measure is used, and there are many kinds of similarity measures (e.g., [41]).

Other examples of a CWW engine are given in Box 1, and no doubt many new engines will be developed in the future.

On the one hand, it is the multitude of choices that provide FL with versatility and flexibility. On the other hand, having so many choices, with none to very few guidelines on how to make them, is confusing. I used to believe that a strong point of FL was its many choices. I still do for certain applications of FL; however, for CWW I no longer believe this to be so.

How does one make the choices needed to implement a CWW engine?

According to [52]–[54] Occam's (or Ockham's) Razor is a principle attributed to the 14th century logician and Franciscan friar, William of Occam. The most useful statement of the principle is "when you have two competing theories which make exactly the same predictions, the one that is simpler is the better." This principle is sometimes misstated as "keep it as simple as possible." We can have two (or more) competing theories that lead to different predictions. Occam's Razor does not apply in that case, because the results that are obtained from the competing theories are different.

Box 5. Why an IT2 FS Model Captures First-Order Uncertainties About a Word

It is this author's experience from collecting survey information from subjects that they do not like to answer a lot of questions, and they like the questions to be simple. Asking subjects to assign a weight or a weighting function to their word data (reflecting their degree of certainty about the interval) is much too difficult.

Asking a subject to provide even just a crisp weight (i.e., a number) for a word, about which we have already argued there is much uncertainty, is contradictory. How can a subject be absolutely sure about that number? Instead, perhaps a subject might be able to assign linguistic terms (e.g., *pretty sure, very sure*, etc.) to different regions of their word data indicating their confidence about that data in different regions. But such terms are words about which there will be additional uncertainties.

We therefore categorize the uncertainty that exists about each person's word data as [23] a *first-order kind of uncertainty*, and the uncertainty that exists about the weight that might be assigned to each element of that data as a *second-order kind of uncertainty*.

When a subject provides their data for a word, the first-order uncertainty is across all of that data. Clearly, weight information is itself uncertain leading to even higher-order (and never ending) kinds of uncertainty. While second-order uncertainty may be interesting from a theoretical point of view, in that it lets us use the complete three-dimensional machinery of a type-2 FS, this author believes that it will be difficult-to-impossible to test the validity of second-order uncertainty by collecting data. Consequently, we should focus exclusively on the first-order uncertainty of each person's word data.

In summary, an IT2 FS captures "first-order uncertainties" about a word whereas a more general T2 FS that has non uniform secondary grades captures first- and second-order uncertainties.

Using Occam's Razor, choices about operators should be made in the crisp domain where different choices all lead to the same results.

What I am about to propose may seem like heresy to many in the FL community. In this article it applies only to the CWW Paradigm. All of our FS and FL operators originate from crisp sets and crisp logic. In the crisp domain, although there can be many different operators, they all give the same results; hence, I propose that for CWW Occam's Razor should be applied to the multitude of t-norm, t-conorms, implication operators, etc., in the crisp domain. It should not be applied after the operators have been fuzzified, because then it is too late–they give different results. By this argument, one would choose minimum or product t-norm, maximum t-co-norm, etc.

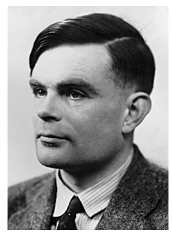
Finally, note, that even a CWW engine that is designed using IT2 FSs and the "simplest" operators needs to be validated by a Turing Test. If, for example, a Per-C that uses the simplest operators does not pass a Turing Test, then more complicated operators should be used.

5. Conclusions

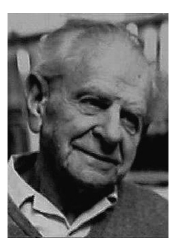
In this article I have argued that Zadeh's CWW Paradigm must be validated, must use scientifically correct FS models for words, and must be simple. I have also demonstrated that the ideas of Turing, Popper and Occam can help us resolve each of these issues. Based on Turing's works, I conclude that validation will require a Turing Test. Based on Popper's Falsificationism, I conclude that one should use interval type-2 FS models in order to model first-order word uncertainties and, using Occam's Razor, I conclude that choices about operators such as t-norms, t-conorms, implications, etc, should be made back in the crisp domain where different choices all lead to the same results. Hence, it is in that domain where it is possible to choose the simplest operators.

Appendix: Biographies of Zadeh, Turing, Popper and Occam

Lotfi A. Zadeh, born in Baku, Azerbaijan on February 4, 1921, and educated at Alborz College in Tehran, the University of Tehran, M. I. T. and Columbia University, spent most of his career at the University of California at Berkeley, after ten years at Columbia University. He was already a famous system theorist when in 1965 he published what has now become the seminal paper on fuzzy sets. This paper marked the beginning of a new direction; by introducing the concept of a fuzzy set, that is a class with unsharp boundaries, he provided a basis for a qualitative approach to the analysis of complex systems in which linguistic rather than numerical variables are employed to describe system behavior and performance. In this way, a much better understanding of how to deal with uncertainty may be achieved, and better models of human reasoning may be constructed. Although his unorthodox ideas were initially met with some skepticism, they have gained wide acceptance in recent years and have found application in just about every field imaginable. He is now acknowledged to be the "Father of Fuzzy Logic." For very detailed information about his technical contributions and awards, see [49]. His wife, Fay Zadeh, has written the 1998 book *My Life and Travels with the Father of Fuzzy Logic*, which provides up-close and personal impressions about Zadeh, his travels, colleagues and friends.



Alan M. Turing was born in London on June 23, 1912, and educated at King's College, Cambridge and Princeton University. His photo is © National Portrait Gallery, London. According to Hodges [5], "He was a mathematician who in 1937 suggested a theoretical machine, since called a Turing Machine that became the basis of modern computing. In 1950 he suggested what has become known as a 'Turing's test,' still the criterion for recognizing intelligence in a machine. During World War II, Turing led the team that succeeded in breaking German high-level secret codes (Enigma), using the first practical programmed computer, called Colossus. … the most lucid and far-reaching expression of Turing's philosophy of machine and mind, the paper *Computing Machinery and Intelligence* [33], appeared in the philosophical journal Mind in 1950. … The wit and drama of the Turing Test has proved a lasting stimulus to later thinkers, and the paper a classic contribution to the philosophy and practice of Artificial Intelligence research. … Alan Turing appears now as the founder of computer science, but, [unlike Zadeh who has been acknowledged to be the Father of Fuzzy Logic during his lifetime] these words were not spoken in his own lifetime. … He died June 7, 1954 of cyanide poisoning, a half-eaten apple beside his bed. The coroner's verdict was suicide." The Alan Turing Web-site, maintained by Andrew Hodges, from which most of this short biography was taken, is [50].



Sir Karl R. Popper was born July 28, 1902, in Vienna, and was schooled at Vienna University. He spent his career at the University of Canterbury in New Zealand (1937–1946), London School of Economics (1946–1949) and University of London (1949–1966), where he became professor of logic and scientific method. According to [51], "[Popper] was an Austrian and British philosopher … counted among the most influential philosophers of science of the 20th century … perhaps best known for repudiating the classical observationalist-inductivist account of scientific method by advancing empirical falsifiability as the criterion for distinguishing scientific theory from non-science … Popper envisioned science as evolving by the successive rejection of falsified theories, rather than falsified statements. Falsified theories are to be replaced by theories which can account for the phenomena which falsified the prior theory, that is, with greater explanatory power." Popper's book *The Logic of Scientific Discovery* (1959) is universally recognized as a classic in the field. He was knighted in 1965 and died September 17, 1994. For additional information, see [32].

William of Ockham (also Occam or any of several other spellings) (c. 1288–1348) was, according to [52], "… an English Franciscan friar and scholastic philosopher, from Ockham, a small village in Surrey, near East Horsley. As a Franciscan, William was devoted to a life of extreme poverty. One important contribution that he made to modern science and modern intellectual culture was through the principle of parsimony in explanation and theory building that came to be known as Ockham's razor. This maxim, as interpreted by Bertrand Russell, states that if one can explain a phenomenon without assuming this or that hypothetical entity, there is no ground for assuming it. That is, one should always opt for an explanation in terms of the fewest possible number of causes, factors, or variables. … Ockham was excommunicated for heresy but his philosophy was never officially condemned. He died on April 9, 1348 in the Franciscan convent in Munich, Germany.… He was posthumously rehabilitated by the official Church in 1359."

References

[1] A.M.G. Cornelissen, *The Two Faces of Sustainability: Fuzzy Evaluation of Sustainable Development*, Ph. D. Thesis, Wageningen Univ., The Netherlands, 2003.

[2] M.R. Civanlar and H.J. Trussel, "Constructing membership functions using statistical data," *Fuzzy Sets and Systems*, vol. 18, pp. 1–14, 1986.

[3] D. Dubois and H. Prade, "Fuzzy sets and statistical data," *European J. of Operational Research*, vol. 25, pp. 345–356, 1986.

[4] H.M. Hersch and A. Caramazza, "A fuzzy set approach to modifiers and vagueness in natural language," *J. of Experimental Psychology*, vol. 105, no. 3, pp. 254–276, 1976.

[5] A. Hodges, *Turing: a Natural Philosopher*, Phoenix, Great Britain, 1997.

[6] J. Kacprzyk and R.R. Yager, "Linguistic summaries of data using fuzzy logic," *Int. J. of General Systems*, vol. 30, pp. 33–154, 2001.

[7] J. Kacprzyk and S. Zadrozny, "Linguistic database summaries and their protoforms: toward natural language based knowledge discovery tools," *Information Sciences*, vol. 173, pp. 281–304, 2005.

[8] J. Kacprzyk and S. Zadrozny, "Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining," in B. Gabrys, K. Leiviska, J. Strackeljan (Eds.), *Do Smart Adaptive Systems Exist?*, pp. 321–339, Springer, Berlin Heidelberg New York, 2005.

[9] J. Kacprzyk, A. Wilbik and S. Zadrozny, "Linquistic summaries of time series via a quantifier based aggregation using the Sugeno integral," *Proc. FUZZ-IEEE 2006*, pp. 3610–3616, Vancouver, BC, Canada, July 2006.

[10] J. Kacprzyk, R.R. Yager, and S. Zadrozny, "A fuzzy logic based approach to linguistic summaries of databases," *Int'l. J. of Applied Mathematics and Computer Science*, vol. 10, pp. 813–834, 2000.

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

[12] J. Lawry, "An alternative to computing with words," *Int'l. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, Suppl., pp. 3–16, 2001.

[13] J. Lawry, J. Shanahan, and A. Ralescu (Eds.), *Modeling With Words*, Lecture Notes in Artificial Intelligence 2873, Springer, New York, 2003.

[14] F. Liu and J.M. Mendel, "An interval approach to fuzzistics for interval type-2 fuzzy sets," *Proc. FUZZ-IEEE 2007*, London, England.

[15] 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.

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

[17] J.M. Mendel, "The perceptual computer: an architecture for computing with words," *Proc. of Modeling With Words Workshop in the Proc. of FUZZ-IEEE 2001*, pp. 35–38, Melbourne, Australia, Dec. 2001.

[18] J.M. Mendel, "An architecture for making judgments using computing with words," *Int. J. Appl. Math. Comput. Sci.*, vol. 12, no. 3, pp. 325–335, 2002.

[19] J.M. Mendel, "Fuzzy sets for words: a new beginning," *Proc. FUZZ-IEEE 2003*, St. Louis, MO, pp. 37–42, 2003.

[20] J.M. Mendel, "Type-2 fuzzy sets: some questions and answers," *IEEE Connections*, Newsletter of the IEEE Neural Networks Society, vol. 1, pp. 10–13, 2003.

[21] J.M. Mendel, "Computing with words and its relationships with fuzzistics," *Information Sciences*, vol. 177, pp. 988–1006, 2007.

[22] J.M. Mendel, "Type-2 fuzzy sets and systems: an overview," *IEEE CI Magazine*, vol. 2, pp. 20–29, Feb. 2007.

[23] J.M. Mendel and R.I. 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, pp. 587–592, July 2002.

[24] J.M. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 1, forward problems," *IEEE Trans. on Fuzzy Systems*, vol. 14, pp. 781–792, Dec. 2006.

[25] J.M. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 2, inverse problems," *IEEE Trans. on Fuzzy Systems*, vol. 15, pp. 301–308, 2007.

[26] J.M. Mendel and H. Wu, "Type-2 fuzzistics for non-symmetric interval type-2 fuzzy sets: Forward problems," *IEEE Trans. on Fuzzy Systems*, vol. 15, 2007.

[27] A.M. Norwich and I.B. Türksen, "The construction of membership functions," in *Fuzzy Sets and Possibility Theory: Recent Developments*, (R, R. Yager, ed.), Pergamon Press, Oxford, pp. 49–60, 1982.

[28] A.M. Norwich and I.B. Türksen, "A model for the measurement of membership and the consequences of its empirical implementation," *Fuzzy Sets and Systems*, vol. 12, pp. 1–25, 1984.

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

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

[31] A.P. Saygin, I. Cicekli, and V. Akman, "Turing test: 50 years later," *Minds and Machines*, vol. 10, pp. 463–518, 2000.

[32] S. Thornton, "Karl Popper," *The Stanford Encyclopedia of Philosophy (Summer 2005 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2005/entries/popper/

[33] A.M. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, pp. 433–460, 1950.

[34] I.B. Türksen, "Measurement of membership functions," in *Applications of Fuzzy Set Theory in Human Factors* (W. Karwowski and A. Mital, Eds.), Elsevier Science Publishers, Amsterdam, The Netherlands, pp. 55–67, 1986.

[35] I.B. Türksen, "Stochastic fuzzy sets: A survey," in *Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making*, (J. Kacprzyk and M. Fedrizzi, eds.), Springer-Verlag, New York, pp. 168–183, 1988.

[36] I.B. Türksen, "Measurement of membership functions and their acquisition," *Fuzzy Sets and Systems*, vol. 40, pp. 5–38, 1991 (especially Section 5).

[37] I.B. Türksen and I.A. Wilson, "A fuzzy set preference model for consumer choice," *Fuzzy Sets and Systems*, vol. 68, pp. 253–266, 1994.

[38] P. Walley and G. de Cooman, "A behavioral model for linguistic uncertainty," *Information Sciences*, vol. 134, pp. 1–37, 2001.

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

[40] D. Wu and J.M. Mendel, "The linguistic weighted average," *Proc. FUZZ-IEEE 2006*, pp. 3030–3037, Vancouver, CA, July 2006.

[41] D. Wu and J.M. Mendel, "A vector similarity measure for interval type-2 fuzzy sets and type-1 fuzzy sets," *Information Sciences*, in press, 2007.

[42] R.R. Yager, "A new approach to the summarization of data," *Information Sciences*, vol. 28, pp. 69–86, 1982.

[43] L.A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[44] L.A. Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning–1," *Information Sciences*, vol. 8, pp. 199–249, 1975.

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

[46] 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.

[47] L.A. Zadeh and J. Kacprzyk (Eds.), *Computing With Words in Information/Intelligent Systems 1 & 2*, Physica-Verlag, New York, 1999.

[48] R. Zwick, "A note on random sets and the Thurstonian scaling methods," *Fuzzy Sets and Systems*, vol. 21, pp. 351–356, 1987.

[49] The Berkeley Initiative on Soft Computing web-site: http://www.cs.berkeley.edu/~zadeh/awfeho.html

[50] The Alan Turing web-site: http://www.turing.org.uk

[51] Wikipedia, the free encyclopedia: http://en.wikipedia.org/wiki/Karl_Popper

[52] Wikipedia, the free encyclopedia: http://en.wikipedia.org/wiki/William_of_Ockham

[53] "Occam's Razor," http://en.wikipedia.org/wiki/Occam%27s_razor

[54] "What is Occam's Razor," http://www.weburbia.com/physics/occam.html