

“How Many Unemployed Are There?”
—When Fuzzy Logic met Linguistics and Semantics

Rudolf Seising
Fundamentals of Soft Computing
European Centre for Soft Computing
Mieres (Asturias), Spain
rudolf.seising@softcomputing.es

Abstract—This paper illuminates the first 10 years of Fuzzy Sets and Systems (FSS) where nobody could know that this theory would be a big success in the applied sciences and technology. Starting with the content of an (perhaps) unpublished text of Lotfi A. Zadeh, we show that, the founder of FSS expected in the period of the years 1965-1972, before Fuzzy control appeared, that this new theory would play an important role in humanities and social sciences. Moreover he tried to link his theory to the fields of linguistics and semantics.

Keywords—Fuzzy Set Theory; history of science, linguistics, semantics, computer science

I. INTRODUCTION

“How many Unemployed Are There?” is the title of a script that I found in the last year in the private archive of Lotfi A. Zadeh. Perhaps this two-pages-paper that was written at “July 5, 1971” was never published [1]. The text is of historical interest because it consolidates my view that in this time Zadeh was intended to establish FSS in humanities, arts and social sciences, notably in linguistics and semantics. [2, 3].

Already in 1969 Zadeh had notified that he did not expect the incorporation of FSS into the fields of sciences and engineering: “What we still lack, and lack rather acutely, are methods for dealing with systems which are too complex or too ill-defined to admit of precise analysis. Such systems pervade life sciences, social sciences, philosophy, economics, psychology and many other ‘soft’ fields.” [4] Much later, in an interview that he gave in 1994, he mentioned his surprise on the fact that Fuzzy Logic (FL) was “embraced by engineers” and “used in industrial process controls and in ‘smart’ consumer products such as hand-held camcorders that cancel out jittering and microwaves that cook your food perfectly at the touch of a single button.” ([5], see also [1, 6])

In that interview he also said that he had “expected people in the social sciences-economics, psychology, philosophy, linguistics, politics, sociology, religion and numerous other areas to pick up on it [FL]. It’s been somewhat of a mystery to me, why even to this day, so few social scientists have discovered how useful it could be.” ([5] but see [7])

In July 1971, when Zadeh wrote “How Many Unemployed Are There?” [1], he was intended to open the field of the new theory’s applications to humanities and social sciences. In section II we go through Zadeh’s claims in this manuscript and in section III we refer to some other papers that he had recited, written, published or supervised in the decade between 1965 and 1975 and which show his intention. The paper concludes with some remarks on the actuality of this question until today.

II. A (PERHAPS) UNPUBLISHED TEXT

The paper “How Many Unemployed Are There?” starts with a version of the Sorites paradox (see fig. 1) that is asking, “How many bald men are there?” Zadeh claims: “In both cases, the answer is not unique because the class of the unemployed, like the class of the bald, does not have a sharply defined natural boundary which separates those who are members of the class from those who are not.”

The Sorites that can be traced back to the old Greek word σορως (sorós for “heap”) used by Eubulid of Alexandria (4th century BC). Already when the mathematician and philosopher Bertrand Russell (1872-1970) quoted the Sorites he discussed the “bald men” (Greek: falakros, English: fallacy, false conclusion) and he discussed colours: “Let us consider the various ways in which common words are vague, and let us begin with such a word as ‘red’. It is perfectly obvious, since colours form a continuum, that there are shades of colour concerning which we shall be in doubt whether to call them red or not, not because we are ignorant of the meaning of the word ‘red’, but because it is a word the extent of whose application is essentially doubtful. This, of course, is the answer to the old puzzle about the man who went bald. It is supposed that at first he was not bald, that he lost his hairs one by one, and that in the end he was bald; therefore, it is argued, there must have been one hair the loss of which converted him into a bald man. This, of course, is absurd. Baldness is a vague
conception; some men are certainly bald, some are certainly not bald, while between them there are men of whom it is not true to say they must either be bald or not bald.” ([8], p. 85). Russell showed in this article from 1923 that concepts are vague even though there have been and continue to be many attempts to define them precisely: “The metre, for example, is defined as the distance between two marks on a certain rod in Paris, when that rod is at a certain temperature. Now, the marks are not points, but patches of a finite size, so that the distance between them is not a precise conception. Moreover, temperature cannot be measured with more than a certain degree of accuracy, and the temperature of a rod is never quite uniform. For all these reasons the conception of a metre is lacking in precision.” ([8], p. 86)

In [1] Zadeh linked the problems of the Sorites and “how many unemployed are there?” with his then five years old theory of Fuzzy Sets (see fig. 2).

He wrote that we can identify “the class of the unemployed” with a fuzzy set (or a “fuzzy class”) and therefore the number of the unemployed equals the cardinality of that fuzzy set. – But how can we compute this cardinality? – Zadeh wrote that to this end “a two-stage process” is required:

First, a formula must be devised under which each individual in a group is assigned a grade of membership in the class of the unemployed. For example, someone who has lost a job and is actively looking for one would be assigned the grade of membership 1. A housewife who takes a half-time job to supplement her family’s income would be assigned the grade of membership 0. On the other hand, someone who is working half-time but would work full-time if he could get a full-time job would be assigned the grade of membership 0.5. And so forth.

Second, the grades of membership of the members of the group are added together. The result is the “number” or, more precisely, the cardinality of the class of the unemployed within the group in question.

However, in these 1960s the requirement of big amounts of data could be difficult but Zadeh wrote that it would have been “not a very big problem” to require the “lot of data” “to compute the grades of membership”. Therefore he concluded that “the approach sketched above has the potential of providing a much truer measure of unemployment that is possible with the conventional dichotomizing techniques which divide the population into two complementary classes – the employed and the unemployed – with no provision for in-between categories. Needless to say, there is bound to be some arbitrariness in the formula for computing the grades of membership in the fuzzy class of the unemployed, but it should not be difficult to devise one which would represent a broad consensus.” [1] Zadeh also emphasized that the cardinality of the fuzzy class of unemployed “corresponds to the number of »full-time equivalent unemployed« in the group, and as such has an intuitively appealing and simple interpretation.

This short manuscript shows Zadeh’s intention in the early years of his new mathematical theory FSS to present it as a proper tool in social sciences and humanities. Already in these late 1960s it is obvious that FSS depends on the used language and beyond that on the meanings of its words. In the next sections we will show these dependencies in more details.

III. FUZZY LOGIC, -LANGUAGES, -SYNTAX, -SEMANTICS

A. Joseph Goguen and “A Logic of Inexact Concepts”

When Zadeh looked for applying FSS in different academic disciplines he kept interdisciplinary company with colleagues on the campus of Berkeley: Joseph Goguen (1941-2006) – who was a Ph. D. student of the mathematician Hans-Joachim Bremermann (1926-1996), the psychologist Eleanor Rosch (Heider) (born 1938) and the linguist George Lakoff (born 1941). Zadeh and Bremermann served as first and second moderators in 1949 a debate meeting about digital computers in Berkeley and he also was a well-known protagonist of System Theory, but then, Zadeh recalled in one of my interviews, “System Theory came grown up but then computers and computers then took over. In other words: the center of attention shifted. [...] So, before that, there were some universities that started departments of system sciences, departments of system engineering, something like that, but then they all went down. They all went down because computer science took over.” [12] Zadeh was interested in computers since the end of the 1940s when he had received the Ph. D from New York’s Columbia University and started working in this new field of research. Here, organized and moderated in 1949 a debate meeting about digital computers in which Claude E. Shannon (1916-2001), Edmund C. Berkeley

---

Illustration and text excerpt from Zadeh’s manuscript showing the concept of a fuzzy set and the formula for computing the cardinality of that fuzzy set.
(1923-1988) and Francis J. Murray (1911-1996) took part. It was probably the first public debate on this subject ever! [12]

When Zadeh had become chairman of the department of electrical engineering (EE) at Berkeley in 1963 (fig. 3), he experienced these shifts very intensively, for it was during his five-year tenure in this position that his department was renamed the Department of Electrical Engineering and Computer Science (EECS) [10]. Another decision was also very important, as Zadeh told me in an interview: “When I was chair of the department in ’63, ’64, the question was what kind of computer should we get and at that time people wanted to get analog computers – at that time! ... Huge installations, you know ... They wanted to spend billions of dollars getting an analog computer. I said no, let’s go digital!” [13]

C. Fuzzy Sets

In the 1960s Zadeh had extensively criticized the relationship between mathematics and his own discipline of electrical engineering (see e.g. chapter 4, 5 in [2]). The mathematical tools offered by classical set theory were not appropriate to the problems that needed to be handled in the engineering sciences. Information and communications technology had led to the construction and design of systems that were so complex that it took much more effort to measure and analyze these systems than had been the case just a few years before. Much more exact methods were now required to identify, classify or characterize such systems or to evaluate and compare them in terms of their performance or adaptivity (see particular chapter 4, in [2]).

In order to provide a mathematically exact expression of experimental research with real systems, it was necessary to employ meticulous case differentiations, differentiated terminology and definitions that were adapted to the actual circumstances, things for which the language normally used in mathematics could not account. The circumstances observed in reality could no longer simply be described using the available mathematical means.

In the summer of 1964 Zadeh was thinking about pattern recognition problems and grades of membership of an object to be an element of a class as he returned to mind almost 50 years later: “While I was serving as chair, I continued to do a lot of thinking about basic issues in systems analysis, especially the issue of unsharpness of class boundaries. In July 1964, I was attending a conference in New York and was staying at the home of my parents. They were away. I had a dinner engagement but it had to be canceled. I was alone in the apartment. My thoughts turned to the unsharpness of class boundaries. It was at that point that the simple concept of a fuzzy set occurred to me. It did not take me long to put my thoughts together and write a paper on the subject. This was the genesis of fuzzy set theory.” ([14], p. 7 see also [2]).

Since that time he often compared the strategies of problem solving by computers on the one hand and by humans on the other hand. In a conference paper in 1970 (published in 1972), he called it a paradox that the human brain is always solving problems by manipulating “fuzzy concepts” and “multidimensional fuzzy sensory inputs” whereas “the computing power of the most powerful, the most sophisticated digital computer in existence” is not able to do this. Therefore, he stated that “in many instances, the solution to a problem need not be exact”, so that a considerable measure of fuzziness in its formulation and results may be tolerable. The human brain is designed to take advantage of this tolerance for imprecision whereas a digital computer, with its need for precise data and instructions, is not.” ([15], p. 132) He continued: “Although present-day computers are not designed to accept fuzzy data or execute fuzzy instructions, they can be programmed to do so indirectly by treating a fuzzy set as a data-type which can be encoded as an array […]”. Granted that this is not a fully satisfactory approach to the endowment of a computer with an ability to manipulate fuzzy concepts, it is at least a step in the direction of enhancing the ability of machines to emulate human thought processes. It is quite possible, however, that truly significant advances in artificial intelligence will have to await the development of machines that can reason in fuzzy and non-quantitative terms in much the same manner as a human being.” ([15], p. 132)
D. From Language to Fuzzy Language

Real world phenomena are very complex and rich of members. To characterize or picture these phenomena in terms of our natural languages we use our vocabulary and because this set of words is restricted, Zadeh argued (in an another probably unpublished text) that this process leads to fuzziness: “Consequently, when we are presented with a class of very high cardinality, we tend to group its elements together into subclasses in such a way as to reduce the complexity of the information processing task involved. When a point is reached where the cardinality of the class of subclasses exceeds the information handling capacity of the human brain, the boundaries of the subclasses are forced to become imprecise and fuzziness becomes a manifestation of this imprecision. This is the reason why the limited vocabulary we have for the description of colors makes it necessary that the names of colors such as red, green, bleu [sic.], purple, etc. be, in effect, names of fuzzy rather than non-fuzzy sets. This is why natural languages, which are much higher in level than programming languages, are fuzzy whereas programming languages are not.” ([16], p. 10) Here, Zadeh argued explicitly for programming languages that are – because of their missing rigidity and preciseness and because of their fuzziness – more like natural languages. He presented a short sketch of his program to extend non-fuzzy formal languages to fuzzy languages which he published in elaborated form co-authored by his Ph. D. student Edward T.-Z. Lee in “Note on Fuzzy Languages” [17].

His definition in these early papers was given in the terminology of the American computer scientists John Edward Hopcroft (born 1939) and Jeffrey David Ullman (born 1942) that was published in the same year [18].

L is a fuzzy language if it is a fuzzy set in the set \( V_T^* \) (the so-called “Kleene closure” of VT, named after the American mathematician Stephen Kleene (1909-1994)) of all finite strings composed of elements of the finite set of terminals \( V_T \), e.g. \( V_T = \{ a, b, c, \ldots, z \} \). The membership function \( \mu_L(x) : V_T^* \rightarrow [0,1] \) associates with each finite string \( x \), composed of elements in \( V_T \), its grade of membership in \( L \). Here is one of the simple examples that he gave in the early paper ([16], p. 16): “Assume that \( V_T = \{ 0, 1 \} \), and take \( L \) to be the fuzzy set \( L = \{ (0,0.9), (1,0.2), (00, 0.8), (01, 0.6), (10,0.7), (11, 0.3) \} \) with the understanding that all the other strings in \( V_T^* \) do not belong to \( L \) (i.e., have grade of membership equal to zero).” ([16], p. 16).

In general the language \( L \) has high cardinality and therefore it is not usual to define it by a finite set of its elements but by a finite set of rules for generating them. Thus, in analogy to the case of non-fuzzy languages Zadeh defined a fuzzy grammar as “a quadruple \( G = (V_N, V_T, P, S) \), where \( V_N \) is a set of variables (non-terminals) from \( V_T \), \( P \) is a set of [fuzzy] productions and \( S \) is an element of \( V_N \). The elements of \( V_N \) (called [fuzzy] syntactic categories) and \( S \) is an abbreviation for the syntactic category »sentence«. The elements of \( P \) define conditioned fuzzy sets \((V_T \cup V_N)^*\).” ([16], p. 16)

Turning to the 1970s Zadeh worked out the basic framework of his FST and FL that gave him the opportunity to characterize fuzzy languages broader than before.

E. From Semantics to Fuzzy Semantics

Zadeh’s occupation with natural and artificial languages gave rise to his studies in semantics. His question was: “Can the fuzziness of meaning be treated quantitatively, at least in principle?” ([19], p. 160). His 1971 article “Quantitative Fuzzy Semantics” [19] starts with a hint to this studies: “Few concepts are as basic to human thinking and yet as elusive of precise definition as the concept of »meaning«. Innumerable papers and books in the fields of philosophy, psychology, and linguistics have dealt at length with the question of what is the meaning of »meaning« without coming up with any definitive answers.” ([19], p. 159)

In this paper Zadeh started a new field of research “to point to the possibility of treating the fuzziness of meaning in a quantitative way and suggest a basis for what might be called quantitative fuzzy semantics” combining his results on Fuzzy languages and Fuzzy relations. In the section “Meaning” of this paper he set up the basics: “Consider two spaces: (a) a universe of discourse, \( U \), and (b) a set of terms, \( T \), which play the roles of names of subsets of \( U \). Let the generic elements of \( T \) and \( U \) be denoted by \( x \) and \( y \), respectively. Then he started to define the meaning \( M(x) \) of a term \( x \) as a fuzzy subset of \( U \) characterized by a membership function \( \mu_L(x) \) which is conditioned on \( x \). One of his examples was: “Let \( U \) be the universe of objects which we can see. Let \( T \) be the set of terms white, grey, green, blue, yellow, red, black. Then each of these terms, e.g., red, may be regarded as a name for a fuzzy subset of elements of \( U \) which are red in color. Thus, the meaning of red, \( M(\text{red}) \), is a specified fuzzy subset of \( U \).”

F. Meaning

In the section “Language” in [15] Zadeh regarded a language \( L \) as a “fuzzy correspondence”, more explicit, a fuzzy binary relation, from the term set \( T = \{ x \} \) to the universe of discourse \( U = \{ y \} \) that is characterized by the membership function \( \mu_L : T \times U \rightarrow [0,1] \).

If a term \( x \) of \( T \) is given, then the membership function \( \mu_L(x, y) \) defines a set \( M(x) \) in \( U \) with the following membership function: \( \mu_{M(x)}(y) = \mu_L(x, y) \). Zadeh called the fuzzy set \( M(x) \) the meaning of the term \( x \); \( x \) is thus the name of \( M(x) \).

---

1 In a footnote he named the works of the philosophers, linguists or cognitive scientists Samuel Abraham and Ferenc Kiefer, Yehoshua Bar Hillel, Max Black, Rudolf Carnap, Noam Chomsky, Alan Fodor and Jerrold J. Katz, Leonard Linsky, Sir John Lyons, Shimon Ullman, Willard Van Orman Quine, Sebastian K. Shaumyan, Zellig Harris.
With this framework Zadeh continued in his 1972 article [15] to establish the basic aspects of a theory of fuzzy languages that is “much broader and more general than that of a formal language in its conventional sense.” ([15], p. 134) In the following we quote his definitions of the concepts fuzzy language, structured fuzzy language and meaning:

**Definition 1:** A fuzzy language $L$ is a quadruple $L = (U, T, E, N)$, in which $U$ is a non-fuzzy universe of discourse; $T$ (called the term set) is a fuzzy set of terms which serve as names of fuzzy subsets of $U$; $E$ (called an embedding set for $T$) is a collection of symbols and their combinations from which the terms are drawn, i.e., $T$ is a fuzzy subset of $E$; and $N$ is a fuzzy relation from $E$ (or more specifically, the support of $T$) to $U$ which will be referred to as a naming relation.

![Figure 6. The components of a fuzzy language: $U =$ universe of discourse; $T =$ term set; $E =$ embedding set for $T$; $N =$ naming relation from $E$ to $U$; $s =$ term; $y =$ object in $U$; $\mu(x, y) =$ strength of the relation between $x$ and $y$; $\mu(x) =$ grade of membership of $x$ in $T$. ([15], p. 136).](Image)

In the case that $U$ and $T$ are infinite large sets, there is no table of membership values for $\mu(x)$ and $\mu_y(x, y)$ and therefore the values of these membership functions have to be computed. To this end, universe of discourse $U$ and term set $T$ have to be endowed with a structure and therefore Zadeh defined the concept of a structured fuzzy language.

**Definition 2:** A structured fuzzy language $L$ is a quadruple $L = (U, S_T, E, S_N)$, in which $U$ is a universe of discourse; $E$ is an embedding set for term set $T$; $S_T$ is a set of rules, called syntactic rules of $L$, which collectively provide an algorithm for computing the membership function, $\mu_T$, of the term set $T$; and $S_N$ is a set of rules, called the semantic rules of $L$, which collectively provide an algorithm for computing the membership function, $\mu_N$, of the fuzzy naming relation $N$. The collection of syntactic and semantic rules of $L$ constitute, respectively, the syntax and semantics of $L$.

To define the concept of meaning, Zadeh characterized the membership function $\mu_T$: $\mu_T: \text{supp}(T) \times U \rightarrow [0,1]$ representing the strength of the relation between a term $x$ in $T$ and an object $y$ in $U$.

However, he clarified: “A language, whether structured or unstructured, will be said to be fuzzy if [term set] $T$ or [naming relation] $N$ or both are fuzzy. Consequently, an non-fuzzy language is one in which both $T$ and $N$ are non-fuzzy. In particular, a non-fuzzy structured language is a language with both non-fuzzy syntax and non-fuzzy semantics.” ([15], p. 138)

With these definitions it is clear that natural languages have fuzzy syntax and fuzzy semantics whereas programming languages, as they were usual in the early 1970s, were non-fuzzy structured languages. The membership functions $\mu_T$ and $\mu_N$ for term set and naming relation, respectively, were two-valued and the compiler used the rules to compute these values 0 or 1. This means that the compiler decides deterministically by using the syntactic rules whether a string $x$ is a term in $T$ or not and it also determines by using the semantic rules whether a term $x$ hits an object $y$ or not. On the other hand we have natural languages, e.g. English, and it is possible that we use sentences that are not completely correct but also not completely incorrect. These sentences have a degree of grammaticality between 0 and 1. Of course, at least native speakers use with high frequency correct sentences. “In most cases, however, the degree of grammaticality of a sentence is either zero or one, so that the set of terms in a natural language has a fairly sharply defined boundary between grammatical and ungrammatical sentences”, Zadeh wrote ([15], p. 138).

Much more fuzziness we find in semantics of natural languages: Zadeh gave the example “if the universe of discourse is identified with the set of ages from 1 to 100, then the atomic terms young and old do not correspond to sharply defined subsets of $U$. The same applies to composite terms such as not very young, not very young and not very old, etc. In effect, most of the terms in a natural language correspond to fuzzy rather than non-fuzzy subsets of the universe of discourse.” ([15], p. 139)

![Figure 7. Membership functions of the fuzzy sets $M(\text{young})$, $M(\text{middle-aged})$ and $M(\text{old})$. ([15], p. 140).](Image)

Zadeh now identified these fuzzy subsets of the universe of discourse that correspond to terms in natural languages with its “meaning”:

**Definition 3:** The meaning of a term $x$ in $T$ is a fuzzy subset $M(x)$ of $U$ in which the grade of membership of an element $y$ of $U$ is given by $\mu_{M(x)}(y) = \mu_T(x, y)$.

Thus, $M(x)$ is a fuzzy subset of $U$ which is conditioned on $x$ as a parameter and which is a section of $N$ in the sense that its

---

3 The support of a fuzzy subset $A$ of $X$ is a non-fuzzy subset $\text{supp}(A)$ defined by $\text{supp}(A) = \{x | \mu_A(x) > 0\}$.

4 Zadeh mentioned that this observation “that natural languages are generally characterized by slightly fuzzy syntax and rather fuzzy semantics does not necessarily hold true when $T$ is associated with an infinite rather than finite alphabet. Thus, when the terms of a language have the form of sounds, pictures, handwritten characters, etc., the fuzziness of its syntax may be quite pronounced. For example, the class of handwritten characters (or sounds) which correspond to a single letter, say $R$, is rather fuzzy, and this is even more true of concatenation of handwritten characters (or sounds).” ([15], p. 139)
membership function, $\mu_{M(x)}: U \rightarrow [0,1]$, is obtained by assigning a particular value, $x$, to the first argument in the membership function of $N$.

Zadeh concluded this paper mentioning that “the theory of fuzzy languages is in an embryonic stage” but he expressed his hope that basing on this framework better models for natural languages will be developed than the models of the “restricted framework of the classical theory of formal languages.” ([15], p. 163)

### IV. Future Work

In the 1970s Zadeh published papers summarizing and developing the concepts we presented above: it appeared in 1973 “Outline of a new approach to the analysis of complex systems and decision processes” [20], in 1975 the three-part article “The concept of a Linguistic Variable and its Application to Approximate Reasoning” [21], in the same year “Fuzzy Logic and Approximate Reasoning” [22] and in 1978 “PRUF – a meaning representation language for natural languages” [23]. In these years psychologist Eleanor Rosch developed her prototype theory [24] she met linguist George Lakoff who was then working at the Center for Advanced Study in Behavioral Sciences at Stanford. During their discussions someone there mentioned Zadeh’s name and his idea of linking English words to membership functions and establishing fuzzy categories in this way. Lakoff wrote an article where he employed “hedges” (meaning barriers) to categorize linguistic expressions and he invented the term “fuzzy logic” [25]. But based on his later research, however, Lakoff came to find that FL was not an appropriate logic for linguistics: Later, he said: “It doesn’t work for real natural languages, in traditional computer systems it works that way.” [26] Also Zadeh wrote an article concerning the meaning of “hedges” as he called them “linguistic operators”. Thus he advocated FL as a suitable tool for linguistics: “A basic idea suggested in this paper in that a linguistic hedge such as very, more, more or less, much, slightly etc. may be viewed as an operator which acts on the fuzzy set representing the meaning of its operand [27].

### V. Outlook

We used some of Zadeh’s unknown or generally not well-known archival papers to show that in the first period after the foundation of FSS, Zadeh expected that this new theory would be useful in humanities and social sciences. Then, we gave a presentation of his own works to linking his theory to the fields of natural and computer languages and semantics. In a final step we show that Zadeh started interdisciplinary research work with scientists in Linguistics and Cognitive Sciences. This historical research work is to be continued.

### Acknowledgment

I would like to thank Lotfi A. Zadeh for his offer to let me browse in his private archives for my historical research. Work leading to this paper was partially supported by the Foundation for the Advancement of Soft Computing Mieres, Asturias Spain.

---

5 PRUF is an acronym for “Possibilistic Relational Universal Fuzzy”.

### References

13. R. S. Interview with L. A. Zadeh on July, 26, 2000, University of California, Berkeley, Soda Hall, unpublished, see [2].
26. R. Seising: Interview with George Lakoff (6.8. 2002), University of California, Berkeley, unpublished, see [2].