Linguistic Hedges: a Quantifier Based Approach

Martine DE COCK
Dept. of Applied Mathematics and Computer Science
Ghent University, Krijgslaan 281 (S9), B-9000 Gent, Belgium
Martine.DeCock@rug.ac.be, http://fuzzy.rug.ac.be

Abstract. We present an entirely new approach for the representation of intensifying and weakening linguistic hedges in fuzzy set theory, which is primarily based on a crisp ordering relation associated with the term that is modified, as well as on a fuzzy quantifier. With this technique we can generate membership functions for both atomic and modified linguistic terms. We prove that our model respects semantic entailment and we show that it surpasses traditional approaches, such as powering and shifting modifiers, on the intuitive level and on the level of applicability.

1 Introduction

The success of fuzzy expert systems is greatly due to their ability to represent and handle vague information expressed by means of linguistic terms in facts and (IF-THEN) rules. Indeed, this tends to make such systems very compact and tolerant for the imprecision and incompleteness that often afflicts our knowledge of the real world. Furthermore, their input, output and inference mechanism can be more easily understood by humans, who's most popular daily means of communication and reasoning is, after all, natural language. In fuzzy systems, each linguistic term is represented by a fuzzy set on a suitable universe X [14]. A fuzzy set A on X is a $X - [0,1]$ mapping, also called the membership function of A, such that for all x in X, $A(x)$ is the membership degree of x in A. In this paper we will use the same notation to denote a fuzzy set A and the term A represented by it.

One of the most crucial and often difficult tasks in developing a fuzzy expert system is the construction of the membership functions for the linguistic terms involved. Fortunately these terms have a specific structure [15] which allows to partially automate this task by computing the membership functions of composed linguistic terms from those of atomic ones. In this context, an atomic linguistic term is an adjective (e.g. expensive). Composed terms are generated by applying either a linguistic modifier to a term (e.g. very expensive), negating a term (e.g. not very expensive) or by combining terms by means of a connective (e.g. not very expensive or sophisticated, nice and easy). If A and B are fuzzy sets on X, the terms A and B, A or B and not A can respectively be modelled by the Zadeh-intersection, the Zadeh-union and the complement, defined as follows

$$\begin{aligned} A \cap B(x) &= \min(A(x), B(x)) \\ A \cup B(x) &= \max(A(x), B(x)) \\ co\ A(x) &= 1 - A(x) \end{aligned}$$

for all x in X. It is even more common to model them by $\mathcal{T}$-intersection, $\mathcal{S}$-union and $\mathcal{N}$-complement based on a triangular norm $\mathcal{T}$, a triangular conorm $\mathcal{S}$ and a negator $\mathcal{N}$; the above mentioned operations are just special cases of this more general approach. In this paper however we will focus on linguistic modifiers (also called linguistic hedges). Since the origin of fuzzy set theory, many researchers have paid attention to the representation of these adverbs, probably because they allow for the generation of many modified terms from existing ones. The first serious attempt was sketched by Zadeh, as early as 1972 [15]. For the representation of very A and more or less A, he defined the concentration and dilation operator (also called powering modifiers) based on a simple involution, i.e. very $A(x) = A(x)^2$ and more or less $A(x) = A(x)^{0.5}$, for all x in X. One can easily verify that in this representation

$$\text{very } A(x) \leq A(x) \leq \text{more or less } A(x)$$

for all x in X. The satisfaction of this so-called semantical entailment [11] (very A is a subset of A, which in turn is a subset of more or less A) can be considered as a prerequisite for any representation, since very is an intensifying modifier, while more or less has a weakening effect. The best known shortcomings of Zadeh's approach, pointed out in e.g. [8, 9, 11], are that they keep the kernel and the support, which are defined as

$$\begin{aligned} \ker A &= \{x | x \in X \wedge A(x) = 1\} \\ \text{supp } A &= \{x | x \in X \wedge A(x) > 0\} \end{aligned}$$

As a consequence, this representation cannot distinguish between being A to degree 1 and being very A to degree 1. One might feel however that a person of 80 years is old to degree 1 but very old only to a lower degree (e.g. 0.7), but this cannot be modelled by means of powering modifiers. The shifting modifiers, informally suggested by Lakoff [11] and more formally developed later on [3, 8, 9], do not have this shortcoming, but it is a serious drawback that they cannot be applied straightforwardly to all kinds of membership functions in the same way, and hence sometimes require artificial tricks. Many representations developed in the same period are afflicted with very similar disadvantages as the powering and shifting modifiers (we refer to [10] for an overview). We believe these kinds of shortcomings on the level of intuition and the level of applicability are due to the fact that these modifiers are only technical tools, lacking inherent meaning.

In fact it was not until the second half of the 1990's that new models with a clear semantics started to surface. In the horizon approach [12] the semantics is derived from the concepts of horizon and visibility [5]. It is one of the few techniques in which the representation for both atomic and modified terms is generated from within the model (which can be a constraint if one prefers to modify membership functions for atomic terms obtained elsewhere). In the fuzzy relational based model [6], the semantics is retrieved from the context. A characteristic of the traditional approaches is that they do not really look at the context: to determine the degree to which x is very A, the concentration operator for instance only looks at the degree to which x is A and ignores the objects in the context of x. By bringing a fuzzy relation into action however, the membership degrees of the elements related to x can also be taken into account to some extent. Depending on the nature of the fuzzy relation, different kind of linguistic modifiers can be modelled within the same framework: an ordering relation gives rise to ordering-based modifiers such as at most and at least [2], while a resemblance relation can be used to represent both weakening (such as roughly and more or less) and intensifying hedges (such as very and extremely) [4]. Informal starting points for the latter two are observations such as "a person is more or less old if he resembles an old person" and "a person is very old if everybody he resembles is old". Translating all components of these observations into their fuzzy set theoretical counterparts, results on the formal level in the use of the direct image (related to the compositional rule of inference) and superdirect image of the fuzzy sets being modified under the fuzzy relations that describe the context.

In this paper we will once again attempt the technique of transforming natural language statements, describing the meaning of (modified) terms, into their mathematical counterparts. This time however the starting point will not be our intuition, but an analysis made by Wheeler in 1972 [13]. Putting all pieces of the puzzle together in the right way gives rise to a stunningly elegant, computationally efficient and semantically very comprehensable representation of both (!) atomic terms and modified terms. The key notion is that of a quantifier, which is not surprising since Wheeler's goal was to reveal that English is a first-order language at some level of analysis. Besides this, relations are once again brought into action to model the context. Most remarkable (compared to the fuzzy relational based approach) is however that crisp ordering relations prove to be very useful to model intensifying and weakening hedges. Since Wheeler describes the meaning of rather and very, we will study the representation of precisely these two hedges. However before presenting the model for modified terms (Section 4), we go into the representation of atomic terms (Section 3) after the necessary preliminaries (Section 2).

2 Preliminaries

Throughout this paper, let X denote a finite universe of discourse (i.e. a non-empty set containing a finite number of objects we want to say something about). In the following, the class of all fuzzy sets on X will be denoted by $\mathcal{F}(X)$. A fuzzy set A takes membership values in the real unit interval $[0,1]$. If all membership values of A belong to $\{0,1\}$, A is called a crisp set. In this case, the notation $A(x) = 1$ corresponds to $x \in A$, while $A(x) = 0$ is the same as $x \notin A$. A fuzzy relation R on X is a fuzzy set on $X \times X$. If the relation R is crisp, we denote $(x,y) \in R$ also by xRy. A useful concept concerning fuzzy relations is that of foreset.

Definition 1 (Foreset). [1] Let R be a fuzzy relation on X and y in X. The R-foreset of y is the fuzzy set Ry defined by $Ry(x) = R(x,y)$, for all x in X.

In other words the R-foreset of y is the fuzzy set of objects related to y. Furthermore we need the concept of inclusion of fuzzy set, as a means to express the mathematical couterpart of semantic entailment.

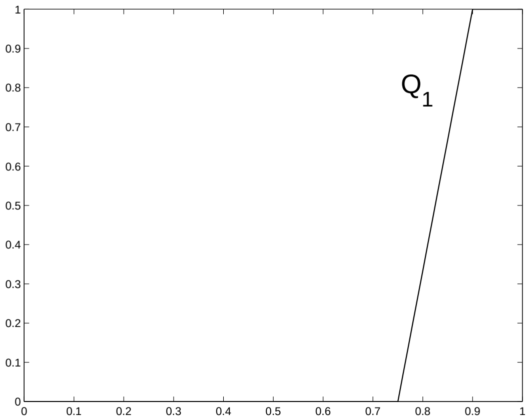

Figure 1: Fuzzy quantifier Q_1

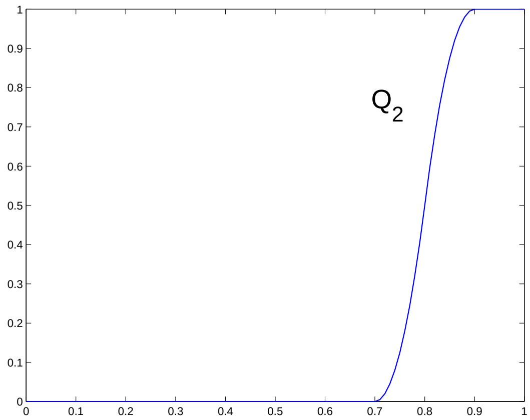

Figure 2: Fuzzy quantifier Q_1

Definition 2 (Inclusion). For A and B in $\mathcal{F}(X)$ one says that A is included in B, denoted $A \subseteq B$, if and only if $A(x) \leq B(x)$, for all x in X.

As already mentioned in the introduction a quantifier will play a crucial role in our model. Intuitively, quantifiers relate to the concept of cardinality of sets. One can easily verify that the power of a fuzzy set is a generalization of the classical concept of cardinality of a crisp set.

Definition 3 (Power). [7] The power of a fuzzy set A on X is defined as

$$|A| = \sum_{x \in X} A(x)$$

More specifically we need a kind of relative cardinality, called proportion.

Definition 4 (Proportion). [17] For A and B fuzzy sets on X, the (size of the) relative proportion of A in B is given by

$$\frac{|A \cap B|}{|A|}$$

Definition 5 (Fuzzy quantifier). [17] A fuzzy quantifier Q is a $[0,1] - [0,1]$ mapping. Q is called regular increasing if it is increasing (i.e. $x \leq y \Rightarrow Q(x) \leq Q(y)$ for all x and y in $[0,1]$) and if it satisfies the boundary conditions $Q(0) = 0$ and $Q(1) = 1$.

Figures 1 and 2 depict two regular increasing fuzzy quantifiers. Finally we recall that Zadeh [18] suggested to calculate the truth value of

$$Q\ A'\text{s are } B'\text{s} \tag{1}$$

as

$$Q\left(\frac{|A \cap B|}{|A|}\right)$$

3 Representation of atomic terms

Wheeler [13] suggests to view adjectives as relations between an individual and a class of individuals. A possibility is to analyze "John is a tall man", which means "John is tall for a man", as "John is taller than most men". In this way the meaning of the vague linguistic term tall is distributed over a crisp ordering relation (is taller than) and a quantifier (most) that helps to preserve the vagueness of the original expression. To obtain something of the form (1) we suggest to further analyze this as "most men are shorter than John".

More formally, to compute the membership degrees for a term A, we suggest to associate with it a crisp ordering relation R (i.e. a reflexive, anti-symmetrical and transitive relation) on the universe X which ranks all elements of X in descending order of satisfying A. From now on we will refer to this relation as "the relation associated with the term". E.g. when representing the term young we will use the relation "is older than or equal to", for expensive we can use "is cheaper than or equal to" etc. For y in X, the foreset Ry then denotes the set of objects related to y, e.g. the set of ages older than y, the set of prices cheaper than y etc. The membership degree of y in A can now be computed as the degree to which most elements of the universe are related to y, i.e. the truth degree of "most X's are Ry", which leads us to the following representational scheme:

Scheme 1. Let Q be a regular increasing quantifier representing most. If R is the ordering relation associated with a term, this term can be represented by the fuzzy set A defined as

$$A(y) = Q\left(\frac{|Ry|}{|X|}\right)$$

for all y in X.

Proposition 1. If A is constructed as described in Scheme 1, then A is increasing w.r.t. the ordering R, i.e.

$$xRy \Rightarrow A(x) \leq A(y)$$

for all x and y in X.

Proof. For all u in X it holds: if $u \in Rx$ then uRx. Because of the transitivity of R and the assumption xRy also uRy or $u \in Ry$. Hence $Rx \subseteq Ry$ and $|Rx| \leq |Ry|$. The proposition now follows from Scheme 1 and the increasing nature of Q. □

Example 1. The membership functions for the atomic terms old, young and middle − aged in the universe of ages $X = \{0, 1, ..., 120\}$ as depicted in Figures 3, 4 and 5 were generated using the fuzzy quantifier Q_2 of Figure 2. Furthermore we respectively used the ordering relations R_1, R_2 and R_3 defined by

$$\begin{array}{lll}(x, y) \in R_1 & \text{if and only if} & x \leq y \\ (x, y) \in R_2 & \text{if and only if} & y \leq x \\ (x, y) \in R_3 & \text{if and only if} & |y - 60| \leq |x - 60|\end{array}$$

for all x and y in X, with $\leq$ being the usual ordering on natural numbers.

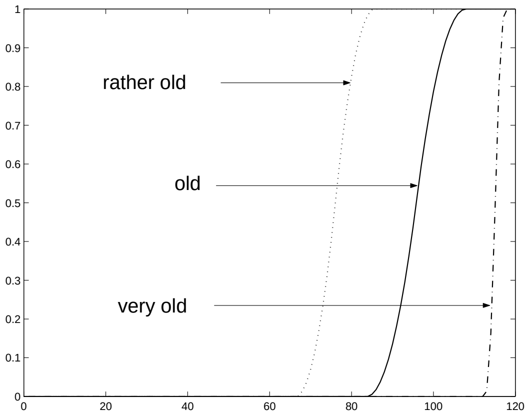

Figure 3: Old and modified terms

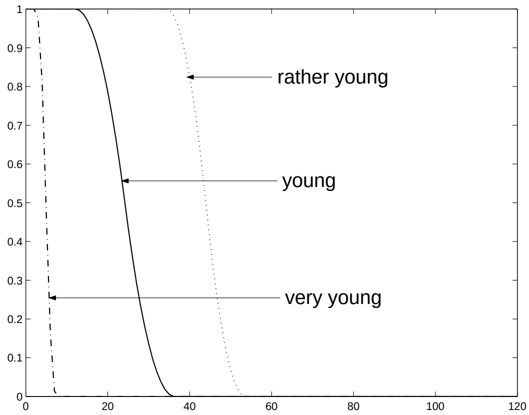

Figure 4: Young and modified terms

4 Representation of modified terms

Wheeler[13] analyzes "John is a very tall man" as "John is tall for a tall man". Furthermore he suggests that "a rather tall man is a man who bears the "tall" relation to men who are not tall but who may or may not be one of the men who are not tall". In other words we could informally say that "John is very tall" if "most tall men are shorter than John" and that "John is rather tall" if "most not tall men are shorter than John", leading to the following representational scheme:

Scheme 2. Let Q be a regular increasing quantifier representing most. If A is the representation of a term and R is the ordering relation associated with this term, the modified terms can be represented by

$$\text{very } A(y) = Q\left(\frac{|A \cap Ry|}{|A|}\right)$$

$$\text{rather } A(y) = Q\left(\frac{|co\ A \cap Ry|}{|co\ A|}\right)$$

for all y in X.

The following proposition reveals that these representations are surprisingly efficient from the computational point of view .

Proposition 2. If the representations for very A and very B are constructed according to Scheme 2 then

$$\text{very } A(y) = Q\left(\frac{\sum\limits_{x \in Ry} A(x)}{\sum\limits_{x \in X} A(x)}\right)$$

$$\text{rather } A(y) = Q\left(\frac{\sum\limits_{x \in Ry} (1 - A(x))}{\sum\limits_{x \in X} (1 - A(x))}\right)$$

Proof. The first equality is an immediate result of

$$\frac{|A \cap Ry|}{|A|} = \frac{\sum\limits_{x \in X} \min(A(x), Ry(x))}{\sum\limits_{x \in X} A(x)} = \frac{\sum\limits_{x \in Ry} A(x)}{\sum\limits_{x \in X} A(x)} \tag{2}$$

The second equality can be proven analogously. □

Note that the denominators in both expressions do not depend on y and therefore have to be computed only once. Furthermore if the membership degrees are computed in the ordering implied by R, the numerator in the expression for an element y can be obtained simply by adding one term to the numerator of the expression for the preceding element. This new term is based on the membership degree of y in A (in the case of very it is precisely $A(y)$, in the case of rather it is $1 - A(y)$), which illustrates very nicely how the context is expanded gradually with every new element taken into account.

In the next theorem we show that the representations of Scheme 1 and 2 satisfy semantic entailment. For better understanding of the proof, note that for all strictly positive real numbers a, b, c and d such that $a \leq d$ and $b \leq c$, one can easily verify that

$$\frac{a+b}{a+c} \leq \frac{d+b}{d+c}$$

I.e. if the denominator of a fraction is dominating, replacement of identical terms in numerator and denominator by larger ones, gives rise to an enlargement of the fraction as a whole.

Theorem 1 (Inclusiveness). If R is a total ordering then

$$\mathsf{very}\ A \subseteq A \subseteq \mathsf{rather}\ A$$

Proof. Replacing $A(x)$ by its definition in (2) we obtain

$$\frac{|A \cap Ry|}{|A|} = \frac{\sum\limits_{x \in Ry} Q(\frac{|Rx|}{|X|})}{\sum\limits_{x \in X} Q(\frac{|Rx|}{|X|})} = \frac{\sum\limits_{x \in Ry} Q(\frac{|Rx|}{|X|})}{\sum\limits_{x \in Ry} Q(\frac{|Rx|}{|X|}) + \sum\limits_{x \notin Ry} Q(\frac{|Rx|}{|X|})}$$

If $x \in Ry$, the transitivity of R implies $Rx \subseteq Ry$. (Indeed, for every $u \in Rx$ it holds that $R(u, x) = 1$. Furthermore because of the assumption that $x \in Ry$ also $R(x, y) = 1$ holds and hence by the transitivity of R we obtain $R(u, y) = 1$, in other words $u \in Ry$.) Therefore, if $x \in Ry$, $|Rx| \leq |Ry|$ holds. Since Q is increasing we obtain

$$\frac{|A \cap Ry|}{|A|} \leq \frac{|Ry| \cdot Q(\frac{|Ry|}{|X|})}{|Ry| \cdot Q(\frac{|Ry|}{|X|}) + \sum\limits_{x \notin Ry} Q(\frac{|Rx|}{|X|})}$$

If $x \notin Ry$ by a similar reasoning as above (assuming that R is a total ordering) we obtain $|Ry| \leq |Rx|$ and hence

$$\frac{|A \cap Ry|}{|A|} \leq \frac{|Ry| \cdot Q(\frac{|Ry|}{|X|})}{|Ry| \cdot Q(\frac{|Ry|}{|X|}) + (|X| - |Ry|)Q(\frac{|Ry|}{|X|})} = \frac{|Ry|}{|X|}$$

The first inclusion now follows immediately from Scheme 1 and 2 and the fact that Q is increasing. The second inclusion can be proven analogously. □

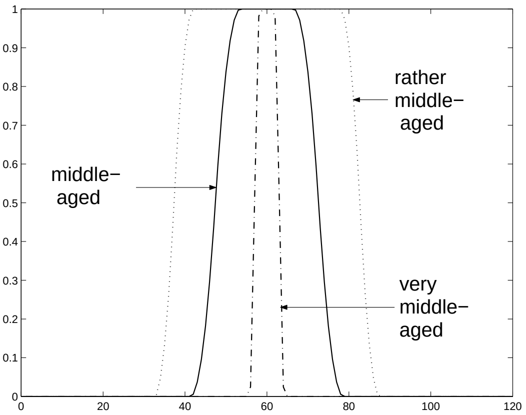

Figure 5: Middle-aged and modified terms

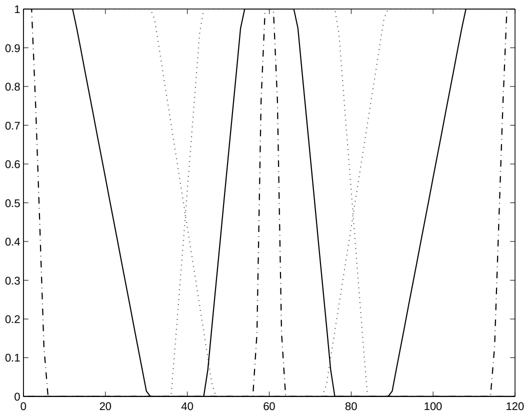

Figure 6: Membership functions in the universe of ages

Example 2. Figures 3, 4 and 5 show membership functions for modified terms generated using the same fuzzy quantifier and the same ordering relations as in Example 1. In Figure 6 the same exercise was done using fuzzy quantifier Q_1 of Figure 1. It is clear from all these figures that the proposed representational schemes respect semantic entailment as shown in Theorem 1, and that the kernel and the support are changed in the modification process. Furthermore the figures illustrate that the same technique can be straightforwardly applied for increasing, decreasing and partially increasing and decreasing membership functions.

5 Conclusion

We have presented a new approach for the fuzzy set theoretical representation of both atomic terms and terms modified by the intensifying hedge very and the weakening hedge rather. Our model is based on a crisp ordering relation that is closely related to the meaning of the term being represented, and that helps to take the context into account. The vagueness of the term is preserved by the use of a fuzzy quantifier. The model turns out to be efficient from the computational point of view, and it respects semantic entailment ($\mathsf{very}\ A \subseteq A \subseteq \mathsf{rather}\ A$). Finally it clearly surpasses traditional approaches, such as powering and shifting modifiers, on the intuitive level and the level of applicability, which is probably due to its clear inherent semantics.

Acknowledgement

The author would like to thank the Fund for Scientific Research - Flanders (FWO) for funding the research reported on in this paper.

References

[1] W. Bandler and L. J. Kohout, Fuzzy Relational Products as a Tool for Analysis and Synthesis of the Behaviour of Complex Natural and Artificial Systems, in: S. K. Wang and P. P. Chang (Ed.), Fuzzy Sets: Theory and Application to Policy Analysis and Information Systems, Plenum Press, New York and London (1980) 341–367.

[2] U. Bodenhofer, A Similarity-Based Generalization of Fuzzy Orderings, Schriftenreihe der Johannes-Kepler-Universität Linz, C 26, Universitätsverlag Rudolf Trauner (1999).

[3] B. Bouchon-Meunier, La Logique Floue, series: Que sais-je?, Vol. 2702, Paris (1993).

[4] M. De Cock and E. E. Kerre, A New Class of Fuzzy Modifiers, in: Proceedings ISMVL 2000 (30th IEEE International Symposium on Multiple-Valued Logic, IEEE Computer Society, (2000) 121–126.

[5] M. De Cock, Fuzzy Hedges: a Next Generation, in: (Kristina Striegnitz, ed.) Proceedings of the ESSLLI2001 Student Session. (2001) 59–70.

[6] M. De Cock, E. E. Kerre, Fuzzy Modifiers Based on Fuzzy Relations, to appear in Information Sciences (2002).

[7] A. De Luca, S. Termini, A Definition of Non-probabilistic entropy in the setting of fuzzy sets, Inform. Control 20 (1972) 301–312.

[8] H. Hellendoorn, Reasoning with fuzzy logic, Ph. D. thesis, T.U. Delft (1990).

[9] E. E. Kerre, Introduction to the Basic Principles of Fuzzy Set Theory and Some of its Applications, Communication and Cognition, Gent (1993).

[10] E. E. Kerre, M. De Cock, Linguistic Modifiers: an overview, in: G. Chen, M. Ying and K.-Y. Cai (Ed.), Fuzzy Logic and Soft Computing. Kluwer Academic Publishers (1999) 69–85.

[11] G. Lakoff, Hedges : a Study in Meaning Criteria and the Logic of Fuzzy Concepts, Journal of Philosophical Logic 2(1973) 458–508.

[12] V. Novák, A Horizon Shifting Model of Linguistic Hedges for Approximate Reasoning, in: Proc. I of the Fifth IEEE International Conference on Fuzzy Systems, New Orleans (1996) 423–427.

[13] S. C. Wheeler III, Attributives and their Modifiers, Noûs VI, 4 (1972) 310–334.

[14] L. A. Zadeh, Fuzzy Sets, Information and Control 8 (1965) 338–353.

[15] L. A. Zadeh, A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, Journal of Cybernetics, 2,3 (1972) 4–34.

[16] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning, part I, II, III, Information Sciences, 8 (1975), 199–249, 301–357; 9 (1975) 43–80.

[17] L. A. Zadeh, A Computational Approach to Fuzzy Quantifiers in Natural Languages, Comput. Math. Appl. 9 (1983) 149–184.

[18] L. A. Zadeh, Test-score Semantics as a Basis for a Computational Approach to the Representation of Meaning, in: (R. R. Yager et. al, Ed.), Fuzzy Sets and Applications: Selected Papers by L. A. Zadeh, Wiley, New York (1987) 655–684.