A SURVEY OF METHODS FOR EVALUATING QUANTIFIED SENTENCES

Miguel Delgado, Daniel Sánchez & Maria Amparo Vila
Department of Computer Science and A.I.
University of Granada, Granada, Spain
e-mail: {delgado, daniel, vila}@decsai.ugr.es

Abstract

The evaluation of quantified sentences is used for solving several problems. Most of the methods proposed are not satisfactory because they do not verify some intuitive properties. In this paper we propose an extension of both possibilistic and probabilistic methods, based on the Sugeno and the Choquet fuzzy integrals respectively, for the evaluation of type II sentences, the most general kind of sentences. These methods have good properties, and they are shown to be better than existing ones.

Keywords: Quantified sentences, quantifiers, fuzzy integrals.

1. Introduction

Quantified sentences are assertions about the number or percentage of objects that verify a certain property. They have applications in many fields such as flexible database querying and expert systems among others (see [8] for a brief review). These sentences are classified into two classes, called type I sentences and type II sentences. A type I sentence is a sentence of the form "Q of X are A", where $X=\{x_1,...,x_n\}$ is a finite set, Q is a linguistic quantifier and A is a fuzzy property defined over X. One example of such a sentence is "Most of the students are young". In this case, X is a finite set of students, the property A is "young" and Q is the quantifier "Most". A type II sentence can be described in general as "Q of D are A", where D is also a fuzzy property over X. One example is the sentence "Most of the young people are tall". Obviously, type I sentences are a special case of type II sentences where D=X.

The evaluation of quantified sentences tries to obtain an accomplishment degree in the real interval [0,1] for the sentence. In [5] we propose a set of intuitive properties that the evaluation must verify. According to those properties, we show that some existing evaluation methods are not appropriate in some cases. The case of the evaluation with non-coherent quantifiers is very significant for both type I and type II sentences. A fuzzy quantifier defines a fuzzy quantity (absolute quantifier) or a fuzzy percentage (relative quantifier). Absolute quantifiers are defined as fuzzy sets over the nonnegative integers, while relative quantifiers are fuzzy sets over the real interval [0,1]. A quantifier Q is said to be coherent if it is non-decreasing and Q(0)=0 and Q(1)=1. The only method that works with non-coherent quantifiers is Zadeh's method, but this method has the drawback that it is very strict when evaluating crisp quantifiers such as "exists" and "forall". These problems are sufficient to justify the search for new methods with better properties. Nevertheless there are other properties that are not verified by some of the existing methods.

In our opinion, the best available methods before our work are based on Sugeno and Choquet fuzzy integrals, although they work only with coherent quantifiers. The latter is equivalent to the method of Yager based on the OWA operator ([9]). They can be interpreted as a possibilistic and a probabilistic approach respectively. In [2] the relation among them is studied, but none of them is preferred. One of our objectives is generalizing these methods for working with coherent quantifiers. Another objective is generalizing these methods for evaluating type II sentences, and studying the relation among them in this case to find some criterion, if possible, to choose one of the methods.

2. Type I sentences

2.1. Previous methods

An early method for evaluating type I sentences was introduced by Zadeh [11] to be

$$Z_Q(A) = Q\left(\frac{P(A)}{|X|}\right)$$

for relative quantifiers, where P(A) is the power of A defined as

$$P(A) = \sum_{i=1}^{n} A(x_i)$$

This method is very strict when evaluating crisp quantifiers. One example is the case of the quantifiers "exists" and "all", that can be defined as

$$\forall(x)=\begin{cases}1 & x=1\\ 0 & x<1\end{cases}\qquad \exists(x)=\begin{cases}0 & x=0\\ 1 & x>0\end{cases}$$

Then, the evaluation of the sentences is

$$Z_{\forall}(A)=\begin{cases}1 & A=X\\ 0 & A\neq X\end{cases}\qquad Z_{\exists}(A)=\begin{cases}0 & A=\varnothing\\ 1 & A\neq\varnothing\end{cases}$$

Other existing methods for evaluating type I sentences are based on the Sugeno and the Choquet integral (see [1]). They are not strict methods, but both of them are restricted by definition to the case of monotonic quantifiers. The method based on the Sugeno integral is defined as

$$S_Q(A)=\max_{1\le i\le n} min(Q(i/n),b_i)$$

where b_i is the i-th greater value of $A(x_i)$, $i\in\{1,...,n\}$. The method based on the Choquet integral is defined as

$$C_Q(A)=\sum_{i=1}^{n} b_i*(Q(i/n)-Q((i-1)/n)$$

Example 1 shows that S_Q is not an appropriate method when dealing with non-coherent quantifiers. Example 2 shows that C_Q has the same problem.

Example 1: Let $X=\{x_1, x_2, x_3\}$ and $A=\{1/x_1, 1/x_2\}$ and let Q(0)=0.5, Q(1/3)=1, Q(2/3)=0.5, Q(1)=0. The quantifier Q can be linguistically interpreted as "approximately 1/3". Then we have the evaluation $S_Q(A)$=max{ min(1,1), min(0.5,1), min(0,0)}=1. But A is crisp, |A|=2 and |X|=3, so the expected result of the evaluation was Q(2/3)=0.5.

Example 2: Let X and Q be as in example 1, and let $A=\{1/x_1\}$. Now we have |A|=1, so the expected result is Q(1/3)=1. Instead, C_Q gives a value $C_Q(A)=0.5$.

2.2. Our methods

In [3] we introduced an extension of the method C_Q, called GD, to be

$$GD_Q(A)=\sum_{i=0}^{n}\left(b_i-b_{i+1}\right)*Q(i/n)$$

In the same paper we show that if Q is coherent, then $GD_Q(A)=C_Q(A)$. This extension can be used with non-monotonic quantifiers and it has good properties.

The results obtained for the examples 1 and 2 are the expected results. In general, it is shown in [5] that if A is crisp then $GD_Q(A)=Q(|A|/|X|)$. Moreover, the evaluation method is not strict in the sense of Zadeh's method, i.e. we can obtain a value in (0,1) even if the quantifier is crisp.

In [4] we introduced an extension called ZS of the method based on the Sugeno integral, to be

$$ZS_Q(A)=\max_{\alpha\in M(A)} min\left(\alpha, Q\left(\frac{|A_\alpha|}{|X|}\right)\right)$$

where

$$M(A)=\left\{\alpha\in\left]0,1\right] \mid \exists x_i\in X \text{ such that } A(x_i)=\alpha\right\}\cup\{1\}$$

We show in [4] that if Q is coherent then $ZS_Q(A)=S_Q(A)$. We also show that this method gives the expected result in the case of A being crisp, and it is not strict.

3. Type II sentences

3.1. Previous methods

Zadeh's method was introduced in [11] to be

$$Z_Q(A/D)=Q\left(\frac{P(A\cap D)}{P(D)}\right)$$

This is also an strict method with crisp quantifiers, for example for the quantifiers "exists" and "forall" we have

$$Z_{\forall}(A/D)=\begin{cases}1 & D\subseteq A\\ 0 & otherwise\end{cases}$$

and

$$Z_{\exists}(A/D)=\begin{cases}0 & A\cap D=\varnothing\\ 1 & otherwise\end{cases}$$

Yager introduced another method in [10] to be

$$Y_Q(A/D)=\sum_{i=1}^{n} w_i c_i$$

where c_i is the i-th largest value of membership to the fuzzy set $D'\cup A$, where D' is the standard complement of D, and

$$w_i=Q(S_i)-Q(S_{i-1})\quad i\in\{1,...,n\}$$

and

$$S_i = \frac{1}{d}\sum_{j=1}^{i} e_i \quad and \quad d = \sum_{k=1}^{n} e_k$$

and e_k is the i-th smallest value of membership to D and S_0=0.

This method has several drawbacks. First, it is defined only for coherent quantifiers. The following examples show two more problems of the method.

Example 3. Let us consider A={1/x_1, 0/x_2} and let D={0/x_1,0.1/x_2}. Also let Q=∃. We have A∩D=∅, so clearly the percentage of objects of D that are in A is 0, and hence the evaluation of the sentence "Q of D are A" should give ∃(0)=0 (see [5]), but the method gives $Y_\exists$(A/D)=0.9 instead.

Example 4. Let us consider A={1/x_1, 0.9/x_2} and D={1/x_1, 0.5/x_2} and Q(x)=x. Clearly D⊆A and D is normalized, so the percentage of objects of D that are in A is 1, and hence the evaluation should be Q(1)=1 (see also [5]). Instead of the expected result, the method gives the result Y_Q(A/D)= 0.93. The same problem arises if Q=∀. The expected result is ∀(1)=1, but the method gives the result $Y_\forall$(A/D)=0.9.

3.2. Our methods

The previous evaluation methods for type II sentences are less satisfactory than the previous methods for type I sentences. We have generalized our methods GD and ZS to the case of type II sentences to obtain methods with good properties.

First we introduced in [4] the method ZS for type II sentences to be

$$ZS_Q(A/D) = \max_{\alpha\in M(A/D)} min\left(\alpha\, , Q\left(\frac{|(A\cap D)_\alpha|}{|D_\alpha|}\right)\right)$$

where M(A/D)=M(A∩D)∪M(D). We assume D is a normal fuzzy set. If it is not, we normalize D in the usual way and we apply the same normalization factor to the fuzzy set A∩D (so, at least conceptually, we are not changing the relative cardinality of D with respect to A).

This method has several good properties as we show in [5]. First, if A and D are crisp, then ZS_Q(A/D) is the expected result Q(|A∩D| / |D|). Second if D⊆A then ZS_Q(A/D)=Q(1). Third, if A∩D=∅ then ZS_Q(A/D)=Q(0). Because of this, the method ZS obtain the expected values for the examples 3 and 4. Moreover this method obtain intuitively good results when dealing with non-coherent quantifiers and it is not an strict method. Finally, its computational complexity is O(nlogn), and the evaluation is faster and easier than Y_Q because Yager's method must obtain the values S_i and w_i previously.

It is clear that when D=X a type II sentence turns to be a type I sentence. We show in [4] that in this case ZS_Q(A/X)=ZS_Q(A).

We have also obtained a generalization of the probabilistic method GD to the case of type II sentences. In [5] we introduce GD_Q(A/D) to be

$$GD_Q(A/D) = \sum_{\alpha_i\in M(A/D)} (\alpha_i - \alpha_{i+1}) * Q\left(\frac{|(A\cap D)_{\alpha_i}|}{|D_{\alpha_i}|}\right)$$

where M(A/D) defined before is described as

$$M(A/D) = \{\alpha_i\} \qquad 1 = \alpha_1 < \alpha_2 < ... < \alpha_m < \alpha_{m+1} = 0$$

This method verify the same properties as method ZS, and hence it gives the expected results for the examples 3 and 4. The computational efficiency of this method is O(nlogn). Also we have that when D=X then GD_Q(A/X)=GD_Q(A).

We have generalized the methods S_Q and C_Q to the case of any quantifier, either coherent or not, and to the case of type II sentences. In [2] Bosc and Lietard compare both methods and conclude that the maximum difference between the results is 0.25. We have also shown that in the case of coherent quantifiers, these methods verify the properties we require for a good method. So S_Q and C_Q are both acceptable. Nevertheless there are more deep differences between their generalizations, GD_Q and ZS_Q , at least in the case of type II sentences, as the following example shows.

Example 5. Let us consider A={1/x_1, 0.01/x_2} and D={1/x_1, 0.99/x_2} and Q=∀. We have that x_2 is very close to be in D, but it is very far from being in A, so intuitively the percentage of elements of D that are in A should not be 1. We obtain a value coherent with that intuition from method GD, $GD_\forall$(A/D)=0.02. But with ZS we obtain $ZS_\forall$(A/D)=1. As we can see, the difference in the case of type II sentences is very high. In fact, we can make this difference as close to 1 as desired by decreasing A(x_2) and increasing D(x_2) without reaching 0 and 1 respectively.

Example 5 gives us a motivation to prefer the method GD to the method ZS. We think that ZS gives a counterintuitive result in this example because the functions maximum and minimum ignore most of the information contained in A and D, while method GD takes into account all the information of A and D. This case also arises in the case of S_Q and C_Q, but the consequences are smoothed because the information about the referential X is fully used by both methods.

4. Conclusions

We have briefly reviewed previously existing evaluation methods for type I and type II sentences. We have discussed these methods from the point of view of some intuitive properties that we think any method must verify. We show by means of some examples that some of the previous methods are not appropriate in some cases, such as the case of sentences with non-coherent quantifiers. Zadeh's method is shownnto be very strict for the evaluation of both type I and type II sentences. We have developed generalizations of the best (to our knowledge) existing methods based on the Sugeno and the Choquet fuzzy integrals. These generalizations have better properties than previous ones. Finally, we have shown with example 5 that the method GD should be preferred in general because it takes into account all the information involved and (to our experiments) the results are always coherent with our intuition.

References

[1] Bosc, P. Lietard, L. Monotonic Quantified Statements and Fuzzy Integrals. *Proc.NAFIPS/IFIS/NASA Conference*, 8-12, 1994.

[2] Bosc, P. Lietard, L. On the Comparison of the Sugeno and the Choquet Fuzzy Integrals for the Evaluation of Quantified Statements. *Proc. of EUFIT '95*, 709-716, 1995.

[3] Delgado, M. Sánchez, D. Vila, M.A. Un Enfoque Lógico para Calcular el Grado de Cumplimiento de Sentencias con Cuantificadores Linguísticos. *Proceedings of Estylf'97*, Tarragona, Spain, 15-20, 1997.

[4] Delgado, M. Sánchez, D. Vila, M.A. Un Método para la Evaluación de Sentencias con Cuantificadores Linguísticos. *Proceedings of Estylf'98*, Pamplona, Spain, 193-198, 1998.

[5] Delgado, M. Sánchez, D. Vila, M.A. Fuzzy Cardinality Based Evaluation of Quantified Sentences. Submitted to Int. Journal of Approximate Reasoning.

[6] Delgado, M. Sánchez, D. Vila, M.A. A Logic Based Definition of Fuzzy Cardinality. Submitted to Fuzzy Sets and Systems.

[7] Liu, Y. Kerre, E. An Overview of Fuzzy Quantifiers. (I). Interpretations. *Fuzzy Sets and Systems* 95, 1-21, 1998.

[8] Liu, Y. Kerre, E. An Overview of Fuzzy Quantifiers. (II). Reasoning and Applications. *Fuzzy Sets and Systems* 95, 135-146, 1998.

[9] Yager, R.R. On Ordered Weighted Averaging Aggregation Operators in Multicriteria DecisionMaking. *IEEE Transactions on Systems, Man and Cybernetics* 18 (1), 183-190, 1988

[10] Yager, R.R. Fuzzy Quotient Operators for Fuzzy Relational Databases. *Proc. of IFES '91*, 289-296, 1991

[11] Zadeh, L.A. A Computational Approach to Fuzzy Quantifiers in Natural Languages. *Computing and Mathematics with Applications* 9 (1), 149-184, 1983.