Symbolically Dealing with Vagueness Expressed in Natural Language

Mazen El-Sayed and Daniel Pacholczyk
University of Angers, 2 Boulevard Lavoisier, 49045 ANGERS Cedex 01, France
{elsayed,pacho}@univ-angers.fr

Abstract. This paper presents a symbolic model for handling *nuanced information* expressed in the affirmative form "x is m_α A". In this model, *nuanced information* are represented in a qualitative way within a symbolic context. For that purpose, vague terms and linguistic modifiers that operate on them are defined. The model presented is based on a symbolic M-valued predicate logic and provides a new deduction rule generalizing the classical Modus Ponens Rule.

Keyword: Knowledge representation and reasoning, Imprecision, Vagueness, Multisets theory, Many-valued logic.

1 Introduction

In this paper, we present a model dealing with *nuanced information* expressed in an affirmative form as they may appear in knowledge bases including, rules like "if the tomato is red then it's ripe" and facts like, "the tomato is very red". The model has been conceived in such a way that the user can deal with statements expressed in natural language, that is to say, referring to a *graduation scale* containing a finite number of *nuances*. The *nuanced statements*, like *"Jo is rather tall"* or *"Jo is really very young"*, can be represented more formally under the form "x is m_α A" where m_α and A are labels denoting respectively a nuance and a vague term. There are two formalisms for handling with nuanced information.

The first one refers to *fuzzy logic* introduced by Zadeh [14, 15] and which is used when the imprecise information is evaluated in a *numerical way*. In this formalism, each vague term, like "red" and "young", is represented by a fuzzy set. This one is defined by a *membership function* that characterizes the gradual membership to the fuzzy set and indicates some properties of the term like *precision, imprecision and fuzziness*. Zadeh [15] uses a fuzzy modifier m_α for representing, from the fuzzy set A, the fuzzy set "m_α A". So, "x is m_α A" is interpreted by Zadeh as "x is $(m_\alpha$ A)" and regarded as many-valued statement. The *fuzzy modifiers* [3, 4, 11, 12] are defined in such a way that operate on fuzzy sets by modifying some of their properties.

The second formalism refers to a symbolic many-valued logic [9, 12] which is used when the imprecise information is evaluated in a *symbolic way*. This logic is the logical counterpart of *multiset theory* introduced by De Glas [9]. In this theory, the term m_α linguistically expresses the degree to which the object x satisfies the term A. So, "x is m_α A" is interpreted by De Glas [9] as "x (is m_α) A", and then regarded as boolean statement. Agreeing on this idea, Pacholczyk [12] considers nevertheless that some nuances of natural language can not be interpreted as satisfaction degrees and must be instead defined such as *linguistic modifiers*. The modifiers have not been studied within a multiset context. The introduction of linguistic modifiers constitutes the main idea of our work. As we noticed previously, the modifiers operate on the term by modifying its meaning. Within a multiset context, there are not concepts used to represent the properties of a term. So, before defining linguistic modifiers we have to propose a

new representation model based on multiset theory and in which we can describe concretely a vague term. This will be our first contribution in this paper. The new model generalizes the results of fuzzy sets theory, namely when the domains are not necessarily numerical scales. Our basic idea has been to associate with each vague term a new *symbolic* concept called "*rule*". This symbolic concept is equivalent to the membership function within a fuzzy context. In other words, its geometry (1) modelizes the gradual membership to the multiset representing the term, and (2) indicates the precision, imprecision and the fuzziness of this term. This new concept allows us to define the *linguistic modifiers* within a multiset context.

Our second contribution in this paper is to propose a deduction rule dealing with nuanced information. For that purpose, we propose a deduction rule generalizing the classical *Modus Ponens* rule in a many-valued logic proposed by Pacholczyk [12]. Note that the first version of this rule has been proposed in a fuzzy context by Zadeh [15] and has been studied later by various authors [1, 3, 5, 10]:

Rule	: if "X is A" then "Y is B"
Fact	: "X is A'"
Conclusion	: "Y is B'"

Where X and Y are variables and A, B, A' and B' are fuzzy concepts.

This paper is organized as follows. In section 2, we present briefly the basic concepts of the M-valued logic which forms the backbone of our work. Section 3 introduces our new approach for the symbolic representation of vague terms. In section 4, we define new linguistic modifiers in a purely symbolic way. In section 5, we propose a new *Generalized Modus Ponens* rule. Section 6 is devoted to some concluding remarks and to further works.

2 M-valued predicate logic

Consider the statement "the tomato is v_α red" where v_α is a nuance of natural language. According to De Glas [9], "x is v_α A" means "x (is v_α) A". Within a multiset context, to a vague term A and a nuance v_α are associated respectively a multiset $\mathbb{A}$ and a symbolic degree τ_α. So, the statement "x is v_α A" means that x belongs to multiset $\mathbb{A}$ with a degree τ_α[1]. The M-valued predicate logic [12] is the logical counterpart of the multiset theory. In this logic, to each multiset $\mathbb{A}$ and a membership degree τ_α are associated a M-valued predicate **A** and a truth degree $\tau_\alpha-$true (some basic elements of this logic are given in appendix B). In this context, the following equivalence holds:

$$\text{x is } v_\alpha \text{ A} \Leftrightarrow x \in_\alpha \mathbb{A} \Leftrightarrow \text{"x is } v_\alpha \text{ \textbf{A}" is true} \Leftrightarrow \text{"x is \textbf{A}" is } \tau_\alpha-\text{true.}$$

2.1 Algebraic structures

One supposes that the membership degrees are symbolic degrees which form an ordered set $\mathcal{L}_M = \{\tau_\alpha, \alpha \in [1, M]\}$. This set is provided with the relation of a total order: $\tau_\alpha \leq \tau_\beta \Leftrightarrow \alpha \leq \beta$, and whose smallest element is τ_1 and the largest element is τ_M. We can then define in $\mathcal{L}_M$ two operators $\wedge$ and $\vee$ and a decreasing involution $\sim$ as follows: $\tau_\alpha \vee \tau_\beta = \tau_{max(\alpha,\beta)}, \tau_\alpha \wedge \tau_\beta = \tau_{min(\alpha,\beta)}$ and $\sim \tau_\alpha = \tau_{M+1-\alpha}$. One obtains then a chain $\{\mathcal{L}_M, \vee, \wedge, \leq\}$ having the structure of De Morgan lattice [12]. On this set, an implication $\rightarrow$ and a T-norm T can be defined respectively as follows: $\tau_\alpha \rightarrow \tau_\beta = \tau_{min(\beta-\alpha+M,M)}$ and $T(\tau_\alpha, \tau_\beta) = \tau_{max(\beta+\alpha-M,1)}$.

---

[1] In the multiset theory, $x \in_\alpha \mathbb{A} \Leftrightarrow$ x belongs to multiset $\mathbb{A}$ with a τ_α degree. It corresponds to $\mu_\mathbb{A}(x) = \tau_\alpha$ within a fuzzy context[14].

Example 1. For example, by choosing M=9, we can introduce: $\mathcal{L}_9=\{$*not at all, little, enough, fairly, moderately, quite, almost, nearly, completely*$\}$.

Remark 1. We use the notation v_α to designate a nuance that must be interpreted as membership degree. From now, we focus to study the statements nuanced with linguistic modifiers which will be denoted as m_α.

In [12], Pacholczyk noticed that some nuances like "*very*" and "*really*", can not be interpreted as membership degrees and they must be defined such as linguistic modifiers. In fact, "m_α A" represent new multisets result from the multiset A. In section 4, we will define these modifiers in multiset theory. They generalize some fuzzy modifiers [3, 4, 15] within a purely symbolic context.

In the following section, we propose a symbolic model for vague terms while inspiring by the representation method within a fuzzy context. More precisely, we associate with each multiset a symbolic concept which has the same role as the membership function associated with a fuzzy set[2].

3 Representation of vague terms

Let us suppose that our knowledge base is characterized by a finite number of concepts $\mathcal{C}_i$. A set of terms P_{ik} is associated with each concept $\mathcal{C}_i$, whose respective domain is denoted as X_i. The terms P_{ik} are said to be the *basic terms* connected with the concept $\mathcal{C}_i$. As an example, basic terms such as *"small", "moderate" and "tall"* are associated with the particular concept "*size of men*". A finite set of *linguistic modifiers* m_α allows us to define *nuanced terms*, denoted as "$m_\alpha P_{ik}$".

Linguists distinguish [6] three signed terms: a negative term like "small", a positive term like "tall" and a neutral term like "moderate". This distinction is important since we define the linguistic modifiers. We suppose that the set of basic terms covers the domain X_i and that each one of these terms is signed. Given that each term P_{ik} is represented by a multiset denoted as P_{ik}, we can propose the following axiom:

Axiom 1. *For each concept $\mathcal{C}_i$, defined on a domain X_i, is associated a family of multisets, denoted as $\mathbb{C}_i = \{P_{i1},...,P_{iN}\}$, which covers the domain X_i. In other words, for all $x \in X_i$ there exists a multiset $P_{ik} \in \mathbb{C}_i$ and $\alpha > 1$ such as $x \in_\alpha P_{ik}$.*

3.1 Representation with *"rules"*

In the following, we propose a symbolic representation to modelize the vague terms which define a "concept". We suppose that a domain of a vague term, denoted by X, is not necessarily a numerical scale. It can for example be "set of men", "set of animals", etc. This domain is simulated by a *"rule"* (cf. Figure 1) representing an arbitrary set of objects. Thus, the set {small, moderate, tall} can be represented as follows:
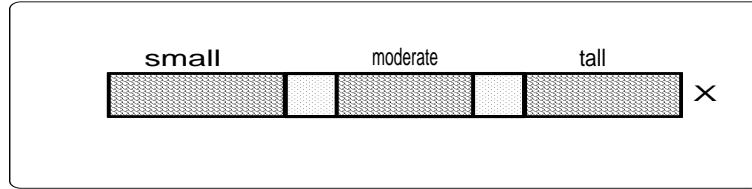

Fig. 1: Representation with *"rule"* of a domain X

---

[2] A short review on the representation of vague terms within a fuzzy context is presented in appendix A.

The basic idea is to associate with each multiset a object which represents a symbolic equivalent to the membership function in the fuzzy set theory. In our work, we focus only on vague terms which can be represented by a membership L-R function. The new object, called "rule", has a geometry similar to a membership L-R function and its role is to illustrate the membership graduality to the multisets. In order to define the geometry of this "rule", we use concepts similar to those defined within a fuzzy context like the core, the support and the fuzzy part of a fuzzy set [15]. The core represents the typical elements of the term, the support contains the elements satisfying at least a little the term, and the fuzzy part represents the atypical elements satisfying partially the term. For example, for the term "tall", the core of the associated multiset represents the perfectly tall men, its support represents the men qualified in the class of tall people, and its fuzzy part represents the more or less tall men.

Definition 1. *The core of a multiset P, denoted as Core(P), is defined by: Core(P) = $\{x \in X \mid x \in_M P\}$.*

Definition 2. *The fuzzy part of a multiset P, denoted as F(P), is defined by: F(P) = $\{x \in X \mid x \in_\alpha P \text{ and } \alpha \in [2, M-1]\}$.*

Definition 3. *The support of a multiset P, denoted as Sp(P), is defined by: Sp(P) = $\{x \in X \mid x \in_\alpha P \text{ and } \tau_\alpha > \tau_1\}$.*

We associate with each multiset a "rule" that contains the elements of its support (cf. Figure 2). Since for any multiset P_i, $\text{Sp}(P_i) = \text{Core}(P_i) \cup \text{F}(P_i)$ and by supposing that the fuzzy part is the union of two disjoined subsets, we can say that the "rule" associated with a multiset is the union of three disjoined subsets: *the left fuzzy part, the right fuzzy part* and *the core*. For a multiset P_i, they are denoted respectively by L_i, R_i and C_i. In order to define formally the concept of directions (left and right) between subsets, we introduce a relation of strict order whose role is to order classical subsets in the universe X.
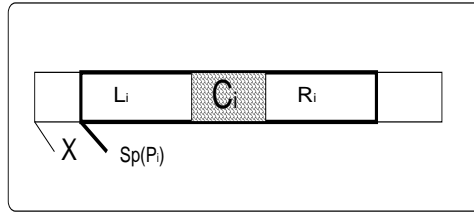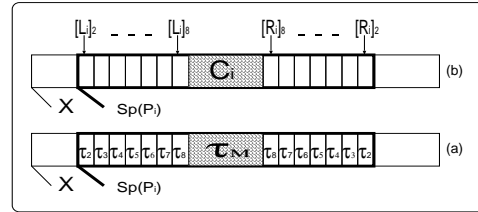

Fig. 2: Parts of a *"rule"*

Fig. 3: Graduality in a *"rule"*

Definition 4. *Let A and B be two disjoined subsets of X. A is said to be on the left compared to B, denoted as $A \prec B$, if and only if, by traversing the "rule" X of left on the right, one meets A before meeting B.*

We use the relation which has been just introduced to define the fuzzy parts L_i and R_i. We want to say by the left fuzzy part of a multiset the subset of $F(P_i)$ located on the left of the core of this multiset. This part is maximal in the meaning of it contains all elements of $F(P_i)$ which are on the left of C_i. In other words, it represents the largest subset of $F(P_i)$ located on the left of C_i. In the same way, we can define the right fuzzy part of P_i as the largest subset of $F(P_i)$ located on the right of C_i. We can thus define them formally as follows.

Definition 5. *Let L_i be a subset of $F(P_i)$. Then, L_i is called left fuzzy part of P_i if and only if: $\forall A$, $A \subset F(P_i)$, if $A \prec C_i$ then $A \subset L_i$.*

Definition 6. *Let R_i be a subset of $F(P_i)$. Then, R_i is called right fuzzy part of P_i if and only if: $\forall A$, $A \subset F(P_i)$, if $C_i \prec A$ then $A \subset R_i$.*

We recall that each fuzzy part contains the elements belonging to P_i with degrees varying from τ_2 to τ_{M-1}. We thus suppose that each fuzzy part L_i and R_i is the union of M-2 subsets which partition it and of which each one contains the elements belonging to P_i with the same degree (cf. Figure 3). These subsets are defined in the following way:

Definition 7. *The set of elements of L_i belonging to P_i with a τ_α degree, denoted as $[L_i]_\alpha$, is defined as follows: $[L_i]_\alpha = \{x \in L_i \mid x \in_\alpha P_i\}$.*

Definition 8. *The set of elements of R_i belonging to P_i with a τ_α degree, denoted as $[R_i]_\alpha$, is defined as follows: $[R_i]_\alpha = \{x \in R_i \mid x \in_\alpha P_i\}$.*

In order to keep a similarity with the fuzzy sets of type L-R, we choose to place, in a "rule" associated with a multiset, the subsets $[L_i]_\alpha$ and $[R_i]_\alpha$ so that the larger α is, the closer the $[L_i]_\alpha$ subsets and $[R_i]_\alpha$ are to the core C_i (cf. Figure 3). That can be interpreted as follows: the elements of the core of a term represent the typical elements of this term, and the more one object moves away from the core, the less it satisfies the term. Finally, we can propose the definition of a multiset represented by a "rule".

Definition 9. *A multiset P_i is defined by (L_i, C_i, R_i), denoted as $P_i = (L_i, C_i, R_i)$, such that:*
- *$\{L_i, C_i, R_i\}$ is totally ordered by the relation $\prec$ and partitions $Sp(P_i)$,*
- *$\{[L_i]_2, ..., [L_i]_{M-1}\}$ is totally ordered by the relation $\prec$ and partitions L_i,*
- *$\{[R_i]_{M-1}, ..., [R_i]_2\}$ is totally ordered by the relation $\prec$ and partitions R_i.*

One supposes that the "rules" associated with multisets have the same geometry but the position of each "rule" and the sizes (or the cardinalities) of its parts depend on the semantics of the term with which it is associated. In the paragraph 3.3, we introduce symbolic parameters to represent these "rules". These parameters are symbolic values using essentially the notion of symbolic cardinality of a subset. This notion is introduced briefly in the next paragraph.

3.2 Symbolic quotient of cardinalities
In order to qualify the cardinality of an ordinary subset, we use a binary predicate called *Rcard* which is defined on boolean formulas of the logic language. This predicate has been defined for handling with symbolic cardinalities of subsets [7, 8]. The definition of this predicate and some properties governing it are presented in appendix B. This predicate allows to define the *Quotient of Cardinalities* of two subsets. More precisely, it allows to qualify the cardinality of a subset compared to another subset with a bigger cardinality. Given two subsets A and B such as the cardinality of B is bigger than the cardinality of A, we can express the cardinality of A compared to the cardinality of B in a qualitative way by using linguistic terms like *"approximately the quarter"*, *"approximately the half"*, etc. Then, we can say "the cardinality of A is *"approximately the half"* among that of B". These terms constitute a set of M symbolic degrees of quotient of cardinalities: $\mathcal{Q}_M = \{Q_\alpha, \alpha \in [1, M]\}$. More generally, we can say "the cardinality of A is Q_α among that of B" which will be denoted as A $\trianglelefteq_\alpha$ B (see appendix B for more details).

Example 2. For M=9, we can introduce the following set: $\mathcal{Q}_9 = \{nothing\ at\ all,\ less\ of$
the quarter, approximately the quarter, approximately the third, approximately the half,
approximately the two thirds, approximately the three quarters, near to equal, equal}.

3.3 Symbolic parameters defining a multiset

Within a symbolic context, we want to define a multiset P_i by a symbolic parameter
set. These parameters are given by experts which describe the "rule" associated with P_i
in a qualitative way. We distinguish two types of parameters: parameters describing the
sign of the term and the internal geometry of the "rule". The description of the internal
geometry concerns the relative sizes of the fuzzy parts and the core compared to the
"rule". The second type of parameters relates to the size of the "rule" and its position in
the domain. The position of a *"rule"* is given compared to another, known as multiset
or "rule" of reference. For example, for the concept "size of men" described by {small,
moderate, tall}, we can say that *"moderate"* is on the right compared to *"small"* and
"tall" on the right compared to *"moderate"*. Equivalently, we can say that *"moderate"*
is on the left compared to *"tall"* and *"small"* on the left compared to *"moderate"*.
So, to locate a "rule" representing a multiset P_i, we introduce two parameters that
indicate: (1) the position of this "rule" compared to a "rule" representing a multiset of
reference P_r and (2) to what degree these "rules" overlap.
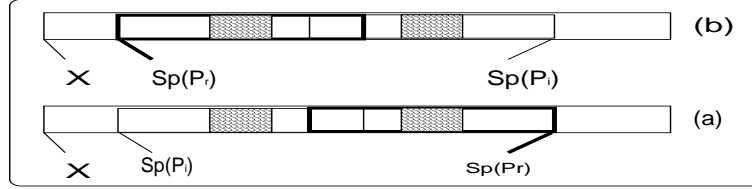

Fig. 4: Relative position for a *"rule"*

We can introduce the following parameters to represent a multiset:

- $Q_{l_i} \in \mathcal{Q}_M$: indicates the relative size of the left fuzzy part compared to the "rule".
  In other words, we have: $L_i \trianglelefteq_{l_i} Sp(P_i)$.
- $Q_{r_i} \in \mathcal{Q}_M$: indicates the relative size of the right fuzzy part compared to the "rule".
  In other words, we have: $R_i \trianglelefteq_{r_i} Sp(P_i)$. Let us announce that the degrees Q_{l_i} and Q_{r_i}
  determine the fuzzy character of the term. The larger these degrees are, the more
  the term is fuzzy.
- $Q_{c_i} \in \mathcal{Q}_M$: indicates the relative size of the core compared to the "rule". In other
  words, we have: $C_i \trianglelefteq_{c_i} Sp(P_i)$.
- $Q_{\epsilon_i} \in \mathcal{Q}_M$: indicates the relative size of the *"rule"* compared to the domain X. In
  other words, we have: $Sp(P_i) \trianglelefteq_{\epsilon_i} X$. This parameter determines the precision of the
  term. More this degree is bigger, more the term is imprecise.
- $\sigma_i \in \{-1, 0, +1\}$: indicates the sign of P_i. $\sigma_i = -1$ if P_i is negative, $\sigma_i = 0$ if P_i is
  neutral and $\sigma_i = +1$ if P_i is positive.
- $s_i \in \{"l", "r"\}$: indicates the position of P_i compared to P_r. $s_i = "l"$ if P_i is on the
  left compared to P_r and $s_i = "r"$ if P_i on the right compared to P_r.
- $Q_{\rho_i} \in \mathcal{Q}_M$: indicates the relative size of the intersection between the "rules" repre-
  senting respectively P_i and P_r. More precisely, we have: $Sp(P_i \cap P_r) \trianglelefteq_{\rho_i} Sp(P_r)$.

Finally, we will represent a multiset P_i by: $P_i = \{< P_r, Q_{\rho_i}, s_i, Q_{\epsilon_i} >, < \sigma_i, Q_{l_i}, Q_{c_i}, Q_{r_i} >\}$.

Remark 2. In this paper, we suppose that, for any multiset P_i, the subsets $[L_i]_\alpha$ (resp. $[R_i]_\alpha$) where α varying from 2 to M-1 have the same cardinality.

Example 3. We consider the concept "size of men" which is described by the three following basic terms: $\mathbb{C} = \{P_i | i \in [1..3]\} = \{$small, moderate, tall$\}$ which are considered respectively as negative, neutral and positive terms. We can show, as an example, how to define the "rule" representing the term "tall". Its position is defined as on the right to "moderate". The relative size of the "rule" and the sizes of the parts forming it can be given as follows. The size of the "rule" is "approximately the third" of that of X which means that "approximately the third" of men are tall. So, we have $Q_{\epsilon_3} = Q_4$. The size of the core is "approximately the quarter" of that of the "rule" which means that "approximately the quarter" of tall men are perfectly tall. Thus, we have $Q_{c_3} = Q_3$. We can define the other parameters as $Q_{l_3} = Q_4$ and $Q_{r_3} = Q_4$ which mean respectively that the size of the left (resp. right) fuzzy part is "approximately the third" of that of the "rule". In the same way, we can define the basic terms as follows:

- *small*: $P_1 = \{< P_1, Q_M, "r", Q_4 >, < -1, Q_1, Q_6, Q_4 >\}$
- *moderate*: $P_2 = \{< P_1, Q_4, "r", Q_4 >, < 0, Q_4, Q_3, Q_4 >\}$
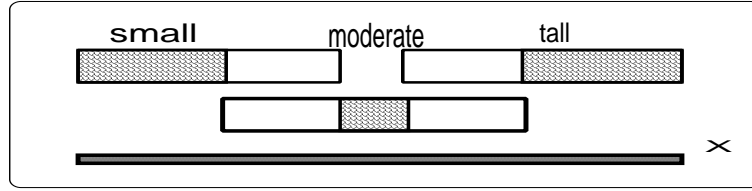- *tall*: $P_3 = \{< P_2, Q_4, "r", Q_4 >, < +1, Q_4, Q_6, Q_1 >\}$.

Fig. 5: Terms associated with "*size of men*"

4 Linguistic modifiers

We have noted previously that some nuances can not be interpreted as symbolic degrees and they must be defined as *linguistic modifiers* [3, 4, 11, 12]. These modifiers provide new vague terms starting from a vague basic term. Thus, from a particular term *"tall"* one obtains new vague terms like *"very tall"* and *"really tall"*. In this section, we define linguistic modifiers in a completely symbolic way and we are interested in some modifiers known as precision modifiers and translation modifiers. We define these modifiers within a symbolic context in which one uses the representation in *"rules"* of vague terms.

4.1 Precision modifiers

The *precision modifiers* make it possible to increase or decrease the precision of the basic term. We distinguish two types of precision modifiers: contraction modifiers and dilation modifiers. A contraction (resp. dilation) modifier m produces nuanced term mP_i more (resp. less) precise than the basic term P_i. In other words, the "rule" associated with mP_i is smaller (resp. bigger) than that associated with P_i. We define these modifiers in a way that the contraction modifiers contract simultaneously the core and the support of a multiset P_i, and the dilation modifiers dilate them. The amplitude of the modification (contraction or dilation) for a precision modifier m is given by a new parameter denoted as τ_γ. The higher τ_γ, the more important the modification is. We give now two definitions for these types of modifiers.

Definition 10. *Let P_i be a multiset. m is said to be a τ_γ-contraction modifier if, and only if it is defined in the following way:*

1. if $P_i = (L_i, C_i, R_i)$ then $mP_i = (L_i', C_i', R_i')$ such as $R_i' \trianglelefteq_M L_i$ and $R_i' \trianglelefteq_M R_i$

2. $\forall x, x \in_\alpha P_i$ with $\tau_\alpha < \tau_M \Rightarrow x \in_\beta mP_i$ such as $\beta = max(1, \alpha - \gamma + 1)$

Definition 11. *Let P_i be a multiset. m is said to be a τ_γ-dilation modifier if, and only if it is defined in the following way:*

1. if $P_i = (L_i, C_i, R_i)$ then $mP_i = (L_i^{'}, C_i^{'}, R_i^{'})$ such as $L_i^{'} \trianglelefteq_M L_i$ et $R_i^{'} \trianglelefteq_M R_i$

2. $\forall x, x \in_\alpha P_i$ with $\tau_\alpha > \tau_1 \Rightarrow x \in_\beta mP_i$ such as $\beta = min(M, \gamma + \alpha - 1)$

In this paper, we use $\mathbb{M}_6 = \{m_k | k \in [1..6]\} = \{$*exactly, really, $\emptyset$, more or less, approximately, vaguely*$\}$ which is totally ordered by $j \leq k \Leftrightarrow m_j \leq m_k$ (cf. Figure 6). $\mathbb{M}_6$ contains a modifier by default, denoted as $\emptyset$, which keeps unchanged the basic term. The modifiers situated after $\emptyset$ are *dilation modifiers* and those preceding it are *contraction modifiers*.
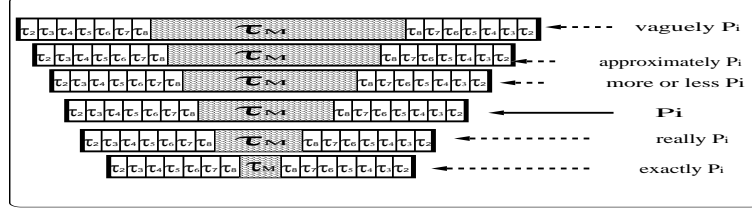

Fig. 6: Precision modifiers

4.2 Translation modifiers

The *translation modifiers* operate both a translation and precision variation on the basic term. We define translation modifiers similar to those defined by Desmontils and Pacholczyk [4] within a fuzzy context. In this work, we use $\mathbb{T}_9 = \{t_k | k \in [1..9]\} = \{$*extremely little, very very little, very little, rather little, $\emptyset$, rather, very, very very, extremely*$\}$ totally ordered by $k \leq l \Longleftrightarrow t_k \leq t_l$ (cf. Figures 7 and 8). $\mathbb{T}_9$ contains a particular modifier, denoted as $\emptyset$, which keeps unchanged the multiset P_i to which it operates. The modifiers preceding $\emptyset$ produce dilations as well as translations on the right if P_i is negative and translations on the left if P_i is positive. The modifiers following $\emptyset$ produce contractions as well as translations on the right if P_i is positive and translations on the left if P_i is negative. The translation amplitudes, the contraction or dilation amplitudes are calculated in such a way that the "rules" associated with nuanced multisets $t_k P_i$ cover the universe X.
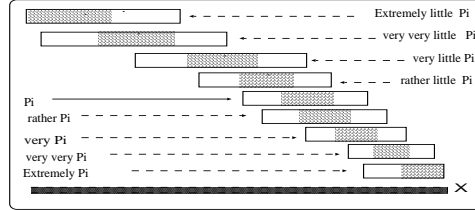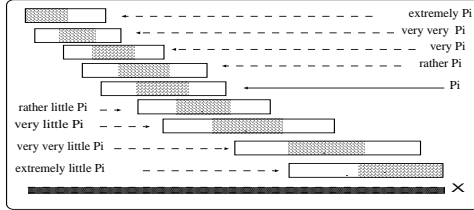

Fig. 7: Translation modifiers (P_i is negative) Fig. 8: Translation modifiers (P_i is positive)

Within a fuzzy context, a modifier operates on the membership function associated to a fuzzy set and modifies the numerical parameters defining it. In a similar way, we define a translation modifier t_k which operates on multisets by modifying their symbolic parameters. So, we can define the translation modifiers as follows:

Definition 12. *Let P_i be a multiset such as $P_i = \{< P_r, Q_{\rho_i}, s_i, Q_{\epsilon_i} >< \sigma_i, Q_{l_i}, Q_{c_i}, Q_{r_i} > \}$. The nuanced multiset $t_k P_i$ is defined in the following way:*

- *$t_k P_i = P_i = \{< P_r, Q_{\rho_i}, s_i, Q_{\epsilon_i} >< \sigma_i, Q_{l_i}, Q_{c_i}, Q_{r_i} >\}$ if $\sigma_i = 0$ or k=5*

- *$t_k P_i = \{< t_{k+1} P_i, Q_{\rho_k}, s_i^{'}, Q_{\epsilon_i^{'}} >< \sigma_i^{'}, Q_{l_i}, Q_{c_i}, Q_{r_i} >\}$ if $1 \leq k < 5$*

- $t_k P_i = \{< t_{k-1} P_i, Q_{\rho_k}, s'_i, Q_{\epsilon'_i} >< \sigma'_i, Q_{l_i}, Q_{c_i}, Q_{r_i} >\}$ if $5 < k \leq 9$

with:

1. $\sigma'_i = $-1 if $\{\sigma_i = -1$ and $k > 5\}$ or $\{\sigma_i = +1$ and $k < 5\}$ and $\sigma'_i = +1$ otherwise,

2. $s'_i = "r"$ if $\{\sigma_i = -1$ and $k > 5\}$ or $\{\sigma_i = -1$ and $k < 5\}$ and $s'_i = "l"$ otherwise,

3. $Q_{\epsilon'_i} \leq Q_{\epsilon_i}$ if $k > 5$ and $Q_{\epsilon'_i} \geq Q_{\epsilon_i}$ otherwise.

In this definition, we choose to place the nuanced multisets one compared to the other. More precisely, one defines the position of $t_k P_i$ compared to $t_{k-1} P_i$ if $k \geq 5$ and $t_k P_i$ compared to $t_{k+1} P_i$ if $k \leq 5$. Let us notice that the parameters Q_{ρ_k} and $Q_{\epsilon'_i}$ are defined in such a way that the "rules" associated with multisets $\{t_k P_i\}_{1 \leq k \leq 9}$ cover the universe X. They indicate respectively the translation amplitudes and the modification (contraction or dilation) amplitudes.

5 Exploitation of vague knowledge

In this section, we treat the exploitation of nuanced information. In particular, we are interested to propose a generalization of the Modus Ponens rule within a many-valued context [12]. The Modus Ponens rule allows, from the rule *If "x is A " then "y is B " is true* and the fact *"x is A " is true*, to conclude *"y is B " is true*. In the M-valued logic on which we work, a generalization of Modus Ponens rule has one of the following forms:

F1- If we know that $\{$*If "x is A " then "y is B " is τ_β-true* and *"x is A' " is τ_ϵ-true*$\}$ and that $\{A'$ is more or less near to A$\}$, what can we conclude for "y is B", in other words, to what degree "y is B" is true?

F2- If we know that $\{$*If "x is A " then "y is B " is τ_β-true* and *"x is A' " is τ_ϵ-true*$\}$ and that $\{A'$ is more or less near to A$\}$, can we find a B' such as $\{B'$ is more or less near to B$\}$ and to what degree "y is B' " is true?

These two forms of the *Generalized Modus Ponens* rule have been studied firstly by Pacholczyk in [12] and later by El-Sayed in [7, 8]. In Pacholczyk's versions, the concept of nearness binding multisets A and A' is modelled by a similarity relation which is defined as follows:

Definition 13. *Let A and B be two multisets. A is said to be τ_α-similar to B, denoted as A $\approx_\alpha B$, if and only if $\{\forall x | x \in_\gamma A$ and $x \in_\beta B \Rightarrow min\{\tau_\gamma \rightarrow \tau_\beta, \tau_\beta \rightarrow \tau_\gamma\} \geq \tau_\alpha\}$.*

This relation generalizes the equivalence relation in a many-valued context as the similarity relation of Zadeh [15] has been in a fuzzy context. It is (1) reflexive: $A \approx_M A$, (2) symmetrical: $A \approx_\alpha B \Leftrightarrow B \approx_\alpha A$, and (3) weakly transitive: $\{A \approx_\alpha B, B \approx_\beta C\} \Rightarrow A \approx_\gamma C$ with $\tau_\gamma \geq T(\tau_\alpha, \tau_\beta)$ where T is a T-norm.

In this paper, we only study the first form (F1) of the *Generalized Modus Ponens* rule. By using the similarity relation to modelize the nearness binding between multisets, the inference rule can be interpreted as: $\{$*more the rule and the fact are true*$\}$ and $\{$*more A' and A are similar*$\}$, *more the conclusion is true*. In particular, when A' is more precise than A $(A' \subset A)$ but they are very weakly similar, any conclusion can be deduced or the conclusion deduced isn't as precise as one can expect. This is due to the fact that the similarity relation isn't able alone to modelize in a satisfactory way the nearness between A' and A. For that, we add to the similarity relation a new relation called *nearness relation* and which has as role to define the nearness of A' to A when $A' \subset A$. In other words, it indicates the degree to which A' is included in A.

Definition 14. *Let A and B be two multisets such as $A \subset B$. A is said to be τ_α-near to B, denoted as $A \sqsubset_\alpha B$, if and only if $\{\forall x \in F(B), x \in_\gamma A$ and $x \in_\beta B \Rightarrow \tau_\alpha \rightarrow \tau_\beta \leq \tau_\gamma\}$.*

Proposition 1. *The nearness relation satisfies the following properties:*

1. *Reflexivity:* $A \sqsubset_M A$
2. *Weak transitivity:* $A \sqsubset_\alpha B$ and $B \sqsubset_\beta C \Rightarrow A \sqsubset_\gamma C$ with $\tau_\gamma \leq min(\tau_\alpha, \tau_\beta)$.

In the relation $A \sqsubset_\alpha B$, the less the value of α is, the more A is included in B. We can notice that the properties satisfied by the nearness relation are similar to those satisfied by the resemblance relation proposed by Bouchon-Meunier and Valverde [2] within a fuzzy context.

Example 4. For a term A, we can obtain: "really A" $\approx_8$ "A", "really A" $\sqsubset_8$ "A", "vaguely A" $\approx_6$ "A", "extremely little small" $\approx_5$ "tall", "very tall" $\sqsubset_2$ "tall", ...

Finally, by using similarity and nearness relations, we can propose our *Generalized Modus Ponens* rule.

Proposition 2. *[Generalized Modus Ponens] Let A and $A^{'}$, B and $B^{'}$ be predicates associated respectively with the concepts $\mathcal{C}_i$ and $\mathcal{C}_e$. Given the following assumptions:*
1. *it is τ_β-true that if "x is A " then "y is B"*
2. *"x is $A^{'}$ " is τ_ϵ-true with $A^{'} \approx_\alpha A$.*

Then, we conclude : "y is B " is τ_δ-true with $\tau_\delta = T(\tau_\beta, T(\tau_\alpha, \tau_\epsilon))$.

If the predicate $A^{'}$ is such as $A^{'} \sqsubset_{\alpha'} A$, then, we conclude: "y is B" is τ_δ-true with $\tau_\delta = T(\tau_\beta, \tau_{\alpha'} \longrightarrow \tau_\epsilon)$.

This inference rule can be interpreted as: {*more the rule and the fact are true*} and {*more $A^{'}$ and A are near*}, *more the conclusion is true*. The nearness concept is represented here by the nearness relation if $A^{'}$ is included in A and by the similarity relation otherwise.

Example 5. From the following rule and fact:
- it is true that if *"the tomato is red"* then *"it is ripe"*
- *"the tomato is really red"* is quite-true,
we can deduce: *"the tomato is ripe"* is almost-true. With the Pacholczyk's inference rule presented in [12] one obtains: *"the tomato is ripe"* is fairly-true. Given that almost-true $>$ fairly-true, we can remark clearly that, when the term in the fact ($A^{'}$) is more precise than the term in the antecedent of the rule (A), our new result is more strong and more precise than that obtained with the old Pacholczyk's inference rule. For the other cases, the two results are identical. Thus, we can conclude that by adding a supplementary assumption for more precising the relation binding $A^{'}$ to A, we can obtain a more precise conclusion from the inference rule.

6 Conclusion

In this paper, we have proposed a model symbolically dealing with nuanced information. In this model, we defined a term by symbolic parameters provided by the expert in a qualitative way. Based on this representation method, we defined some linguistic modifiers in a purely symbolic way. Lastly, we presented a new *Generalized Modus Ponens* rule for exploiting nuanced information. With this rule, we obtain satisfactory results. Finally, we plan to generalize our model of reasoning in order to reason on *gradual rules* like *"the more the tomato is red, the more it is ripe"*.

References

1. J. F. Baldwin. A new approach to approximate reasoning using fuzzy logic. *Fuzzy sets and systems*, 2:309 − 325, 1979.
2. B. Bouchon-Meunier and L. Valverde. A ressemblance approach to analogical reasoning function. *Lecture notes in computer sciences*, 1188:266–272, 1997.

3. B. Bouchon-Meunier and J. Yao. Linguistic modifiers and imprecise categories. *Int. J. of intelligent systems*, 7:25–36, 1992.

4. E. Desmontils and D. Pacholczyk. Towards a linguistic processing of properties in declarative modelling. *Int. J. of CADCAM and Comp. Graphics*, 12:351–371, 1997.

5. D. Dubois and H. Prade. Fuzzy sets in approximate reasoning, part 1 : Inference with possibility distributions. *Fuzzy sets and systems*, 40:143–202, 1991.

6. O. Ducrot and J-M. Schaeffer. Nouveau dictionnaire encyclopédique des sciences du langage. *Editions du Seuil*, 1995.

7. M. El-Sayed. Une approche logico-symbolique des connaissances nuancees via des modificateurs linguistiques. *Proc. RJCIA'00, France, 113-126*, 2000.

8. M. El-Sayed. *Une approche logico-symbolique des connaissances nuancees via des modificateurs linguistiques*. PhD thesis, University of Angers, 2001.

9. M. De glas. Knowladge representation in fuzzy setting. Technical Report 48, LAFORIA, 1989.

10. L. D. Lascio, A. Gisolfi, and U. C. Garcia. Linguistic hedges and the generalized modus ponens. *Int. Journal of intelligent systems*, 14:981–993, 1999.

11. L. D. Lascio, A. Gisolfi, and V. Loia. A new model for linguistic modifiers. *Int. Journal of approximate reasoning*, 15:25–47, 1996.

12. D. Pacholczyk. *Contribution au traitement logico-symbolique de la connaissance.* PhD thesis, University of Paris VI, 1992.

13. D. Pacholczyk. A logico-symbolic probability theory for the management of uncertainty. *The Journal for the integrated study of artificial intelligence cognitive science and applied epistemology*, 11:417 – 484, 1994.

14. L. A. Zadeh. Fuzzy sets. *Information and control*, 8:338–353, 1965.

15. L. A. Zadeh. A theory of approximate reasoning. *Int. J. Hayes, D. Michie and L. I. Mikulich (eds); Machine Intelligence*, 9:149–194, 1979.

Appendix A: Representation within a fuzzy context

Within a fuzzy context, a fuzzy set A is associated with each vague term. A is defined by a membership function μ defined on numerical scale and which indicates the gradual membership to A. This fuzzy set is characterized even by some traditional subsets like the core, the support and the fuzzy part. These three subsets are defined respectively by Core(A)=$\{x \,|\mu_A(x) = 1\}$, Supp(A)=$\{x \,|\mu_A(x) > 0\}$, F(A) = $\{x \,|0 < \mu_A(x) < 1\}$ [2]. They represent some semantic characteristics of the term to which A is associated. The fuzzy part cardinality indicates the fuzziness of the term and the support cardinality indicates its precision. In the fuzzy sets usually used, whose have membership L-R functions, the fuzzy part is the union of two disjoined subsets: the left fuzzy part and the right fuzzy part. These fuzzy sets are defined even by numerical parameters like $A = \{X, < \alpha, a, b, \beta >, L, R\}$ (cf. Figure 9). These parameters indicate the sizes of the core, the support and the fuzzy parts, and their positions in the universe X. L(x) and R(x) are two functions that define $\mu_A(x)$ respectively on left and right fuzzy parts.

In this context, the nuances, like "very" and denoted as m_α, are fuzzy modifiers which allow to define nuanced terms denoted as "$m_\alpha A$", whose membership L-R functions simply result from A by using a translation, dilation or contraction. Some modifiers produce both (a translation and contraction) or (a translation and dilation). These modifiers are defined by Desmontils and Pacholczyk [4] in such a way that a finite and ordered set of modified fuzzy sets cover X (cf. Figure 10).
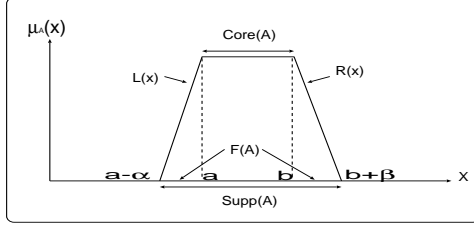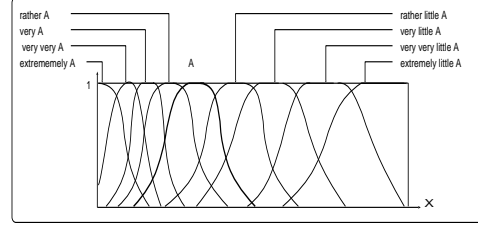
Fig. 9: Representation on fuzzy set

Fig. 10: Translation modifiers

Appendix B: M-valued logic and Symbolic Quotient of cardinalities

We begin this appendix by giving some basic elements of the many-valued predicate logic used here and which can be found in [12]. After, we describe briefly a symbolic approach for handling symbolic cardinalities of subsets. More details about this approach can be found in [7, 8]. Let $\mathcal{L}$ be a language of M-valued predicates and $\mathcal{F}$ the set of the formulas on $\mathcal{L}$. We call an interpretation structure $\mathcal{A}$ of $\mathcal{L}$, the pair $\mathcal{A} = <\mathcal{D}, \{R_n | n \in N\}>$, where $\mathcal{D}$ designates the domain of $\mathcal{A}$ and R_n designate the multisets associated with the predicate P_n of the language. We call a valuation of variables of $\mathcal{L}$, a sequence denoted as $v = <v_0, v_1, ..., v_{n-1}, v_n, v_{n+1}, ...>$. The valuation $v(n/a)$ is defined by $v(n/a) = <v_0, v_1, ..., v_{n-1}, a, v_{n+1}, ...>$.

Definition 15. *For any formula ϕ of $\mathcal{F}$, the relation of partial satisfaction 'v satisfies ϕ to a degree τ_α in $\mathcal{A}$', denoted as $\mathcal{A} \models_\alpha^v \phi$, is defined as follows:*

1. $\mathcal{A} \models_\alpha^v P_n(z_{i_1}, ..., z_{i_{\eta(n)}}) \Leftrightarrow (z_{i_1}, ..., z_{i_{\eta(n)}}) \in_\alpha R_n$.
2. $\mathcal{A} \models_\alpha^v \neg\phi \Leftrightarrow \mathcal{A} \models_\beta^v \phi | \tau_\beta = \sim \tau_\alpha$
3. $\mathcal{A} \models_\alpha^v \phi \bigcup \psi \Leftrightarrow \mathcal{A} \models_\gamma^v \phi \ et \ \mathcal{A} \models_\beta^v \psi | \tau_\gamma \vee \tau_\beta = \tau_\alpha$
4. $\mathcal{A} \models_\alpha^v \phi \bigcap \psi \Leftrightarrow \mathcal{A} \models_\gamma^v \phi \ et \ \mathcal{A} \models_\beta^v \psi | \tau_\gamma \wedge \tau_\beta = \tau_\alpha$
5. $\mathcal{A} \models_\alpha^v \phi \supset \psi \Leftrightarrow \mathcal{A} \models_\gamma^v \phi \ et \ \mathcal{A} \models_\beta^v \psi | \tau_\gamma \rightarrow \tau_\beta = \tau_\alpha$

Definition 16. *A formula ϕ is said to be $\tau_\alpha - true-in-\mathcal{A}$, if and only if, there exists a valuation v such that v $\tau_\alpha - satisfies$ ϕ in $\mathcal{A}$.*

Definition of the predicate Rcard

Given that $\mathcal{A}^*$ is an arbitrary interpretation of $\mathcal{L}$ of the universe of discourse Ω, let $\mathcal{C}$ the set of open formulas ϕ of $\mathcal{F}$ such as for any valuation v of Ω, ϕ is *completely* satisfied in $\mathcal{A}^*$ or is *not-at-all* satisfied in $\mathcal{A}^*$. Thus, in the interpretation $\mathcal{A}^*$, each formula of $\mathcal{C}$ can refer to a subset of the individuals of Ω who satisfy this formula.

To define the symbolic quotients of cardinalities assigned to sets referred by formulas of $\mathcal{C}$, we introduce a new many-valued predicate called **Rcard** added to the language $\mathcal{L}$ and defined on the formulas of $\mathcal{C}$. This predicate will modelize the notion of the quotient of cardinalities between sets referred by formulas of $\mathcal{C}$. The degrees of truth of this predicate represent Symbolic Quotients of Cardinalities of two subsets. We extend the structure of interpretation of the language $\mathcal{A}^*$ to $\mathcal{A}$ of the domain $\mathcal{D} = \Omega \cup \mathcal{C}$, and we suppose that the noted variables φ and Ψ indicate the arguments of **Rcard** and that any valuation v comprises, moreover, $(v_0, v_1) \in \mathcal{C} \times \mathcal{C}$ which are associated with the arguments φ and Ψ.

Definition 17. *The Rcard predicate is defined formally in the following way:*

- *For any interpretation $\mathcal{A}$, Rcard is a binary predicate of the language, defined on $\mathcal{C} \times \mathcal{C}$ and for which $\forall\phi \in \mathcal{C}$, $\forall\psi \in \mathcal{C}$ $Rcard(\phi, \psi) \in \mathcal{F}$.*
- *Any interpretation $\mathcal{A}$ associates with the Rcard predicate a multiset of $\mathcal{C} \times \mathcal{C}$, denoted as $\mathcal{RPS}$, such as for any valuation v and if ϕ and ψ are arbitrary elements of $\mathcal{C}$, then we have: $\mathcal{A} \models_\alpha^{v(0/\phi)(1/\psi)} Rcard(\varphi, \Psi) \Leftrightarrow <(\phi, \psi)>\in_\alpha \mathcal{RPS} \Leftrightarrow Rcard(\phi, \psi)$ is τ_α-true-in-$\mathcal{A}$. If no confusion is possible $\mathcal{A} \models_\alpha Rcard(\phi, \psi)$ stands for $\mathcal{A} \models_\alpha^{v(0/\phi)(1/\psi)} Rcard(\varphi, \Psi)$.*

- $\forall \mu, \ \mathcal{A} \models_\mu Rcard(\phi, \psi) \Longrightarrow \{\exists \alpha, \exists \lambda \mid \alpha \leq \lambda, \ \mathcal{A} \models_\alpha Rcard(\phi, \top) \ et \ \mathcal{A} \models_\lambda Rcard(\psi, \top)\}$.

Let us indicate by $\mathcal{Q}_M$ the set of M symbolic degrees of Quotients of cardinalities: $\mathcal{Q}_M = \{Q_\alpha, \alpha \in [1, M]\}$. The term "$\mathcal{A} \models_\alpha Rcard(\phi, \psi)$", will mean that "the cardinality of the individuals set who satisfy ϕ in $\mathcal{A}$ is Q_α among that of the individuals set who satisfy ψ in $\mathcal{A}$". Thus, with each symbolic degree of truth τ_α-true of $Rcard(\phi, \psi)$, we associate a symbolic degree of quotient of cardinalities i.e. Q_α of $\mathcal{Q}_M$. To handle in an equivalent way of the sets and formulas referring to these sets, we use a relation, denoted as $\unlhd$, which expresses the quotient of cardinalities of two subsets. Let us suppose that ϕ and ψ are formulas which refer respectively to the subsets A and B of Ω in interpretation $\mathcal{A}$, such as the cardinality of B is larger than that of A. The equivalence between the two notations, enables us to use "$A \unlhd_\mu B$" in the place of "$\mathcal{A} \models_\mu Rcard(\phi, \psi)$". This equivalence can be established as follows: $\mathcal{A} \models_\mu Rcard(\phi, \psi) \Leftrightarrow A \unlhd_\mu B \Leftrightarrow$ "the cardinality of A is Q_μ among that of B".

Axiomatic governing Rcard

We establish for the relation $\unlhd$ some properties equivalent to the ones satisfied by the function of traditional cardinality. In this appendix, we limit ourselves to define the axiomatic governing the predicate Rcard. Other properties can be found in [7, 8].

Axiom 2 (property of the union). *Let A and B two subsets of E. If $\{A \cap B = \emptyset, \ A \unlhd_\alpha E, B \unlhd_\beta E \ and \ \alpha + \beta \leq M + 1\}$ then $(A \cup B) \unlhd_\gamma E$ with $Q_\gamma = S(Q_\alpha, Q_\beta)$.*

Where **S** is an application representing the *"symbolic sum"*. This one is an application of $\mathcal{Q}_M \times \mathcal{Q}_M$ in $\mathcal{Q}_M$ which satisfies the properties of the traditional sum. We choose an application inspired from Lukasiewicz's T-conorm [12]. This application is defined as follows:

$$\text{If } \alpha + \beta \leq M + 1 \text{ then } S(Q_\alpha, Q_\beta) = Q_{(\alpha+\beta-1)}$$

This axiom represents a symbolic generalization of the following expression:

$$\frac{|A \cup B|}{|E|} = \frac{|A| + |B| - |A \cap B|}{|E|} = \frac{|A|}{|E|} + \frac{|B|}{|E|} \text{ (because } |A \cap B| = 0).$$

Axiom 3 (weak Transitivity). *If $\{B \unlhd_\beta A \ and \ A \unlhd_\alpha E\}$ then $B \unlhd_\gamma E$ with $Q_\gamma = I(Q_\alpha, Q_\beta)$.*

Where **I** is an application representing the *"symbolic multiplication"*. The latter is an application of $\mathcal{Q}_M \times \mathcal{Q}_M$ in $\mathcal{Q}_M$ which satisfies the properties of the traditional multiplication. This application satisfies some properties proposed by Pacholczyk [13]. We can choose the multiplication presented in Table 1. This axiom represents a symbolic generalization of the following relation:

$$\frac{|B|}{|A|} \times \frac{|A|}{|E|} = \frac{|B|}{|E|}$$

I	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9
Q_1	Q_1	Q_1	Q_1	Q_1	Q_1	Q_1	Q_1	Q_1	Q_1
Q_2	Q_1	Q_2	Q_2	Q_2	Q_2	Q_2	Q_2	Q_2	Q_2
Q_3	Q_1	Q_2	Q_2	Q_2	Q_2	Q_2	Q_2	Q_3	Q_3
Q_4	Q_1	Q_2	Q_2	Q_2	Q_2	Q_2	Q_3	Q_4	Q_4
Q_5	Q_1	Q_2	Q_2	Q_2	Q_2	Q_3	Q_4	Q_5	Q_5
Q_6	Q_1	Q_2	Q_2	Q_2	Q_3	Q_4	Q_5	Q_6	Q_6
Q_7	Q_1	Q_2	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_7
Q_8	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_8
Q_9	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9

Table 1: multiplication table