

Chapter 2

Meaning as a Quantity

Considering the meaning of words is necessary for understanding both our own reasoning and that of others; with meaningless words people's plain language and ordinary reasoning would be for nothing, and communication impossible. Hence, a previous symbolic analysis of the meaning of words is important for better comprehending what ordinary reasoning is. In addition, meaning is a twofold concept: it has two sides, the qualitative and the quantitative, reflecting the situational use of words, namely that its use is context dependent and purpose driven. Often, and from a scientific point of view, only considering meaning from its qualitative side is insufficient; it is the quantitative that facilitates the degrees up to which words actually mean something in a given context. Meaning should be somehow measured.

2.1. Given a universe X in which a predicative word P is acting through the elemental statements “ x is P ”, let's symbolically denote by $<_P$ the previously captured relationship “less P than”,

$$x <_P y \Leftrightarrow x \text{ is less } P \text{ than } y,$$

translating the either empirically, or theoretically, recognized fact that x verifies less than y the property p named P . Then, the graph $(X, <_P)$ represents how P semantically organizes the universe of discourse X , and denotes a primary *qualitative meaning* of P in X . The inverse relationship “more P than” x is more P than y , coincides with $y <_P x$, and when it simultaneously holds $x <_P y$ and $y <_P x$, is that x is “equally P than” y , symbolically written $x =_P y$.

Note that, at the end, people's (intelligent) talking and telling tries to introduce some organization, or ordering, between the concepts/words under consideration, and for trying to answer the questions leading to telling something; in this respect, it seems natural to consider that the symbolic relation $<_P$ producing the qualitative

meaning is the way in which P semantically acts on X , introducing in it some organization. Usually, words are but names of concepts mastered though the meaning of the words.

It should be pointed out that as the relation $<_P$ is, when P is a word managed in plain language, empirically recognized, it implies some subjectivism that is, nevertheless, shared with others; $<_P$ is here presented as just a kind of primitive idea, as if it were “point” and “line” in the old Euclidean *Elements*.

It should be pointed out that the relation $<_P$ is not, in general, a linear or total relation; that is, there can exist pairs of elements x and y such that it is neither $x <_P y$, nor $y <_P x$; in this case, x and y are not meaning-comparable and are denoted by $x \text{ NC}_P y$. For instance, in the toy example of the word $P = \text{big}$ in $X = [0, 10]$, it is $x <_{\text{big}} y$ if and only if $x \leq y$, in the linear order \leq of the real line; that is, $<_P = \leq$ is a total relation under which all the elements of the universe $[0, 10]$ are big-comparable; hence x is “more” big than y whenever is $y \leq x$, and x is “equally big than” y whenever both numbers coincide, $x = y$. The qualitative meaning of “big” in $[0, 10]$ is given by the graph $([0, 10], \leq)$, in which there is the unique maximal 10 and the unique minimal 0; obviously, in the real interval $[0, 10]$, 10 is always a prototype of big, and 0 is always an antiprototype. In the same interval, calling “medium” the property of being around 5, the situation is different provided, for instance, the prototypes were those x in $[4.9, 5.1]$, the antiprototypes were those x in $[0, 3] \cup [7, 10]$, and

$$x <_{\text{medium}} y \Leftrightarrow x \leq y \leq 4.9, \text{ or } 5.1 \leq y \leq x,$$

in which case there are elements, such as 4 and 6, that would not be comparable, but isolated. Note that provided the prototypes were the elements in the open interval $(3, 7)$, and the antiprototypes those in $[0, 3] \cup [7, 10]$; then the use of “medium” would be precise, and given by the necessary and sufficient condition (definition),

$$\text{‘}x \text{ is medium’ if and only if } 3 < x < 7.$$

Of the symbolic relation $<_P$, it can be easily accepted that it is always reflexive, such that $x <_P x$ for all x in X , but not, for instance, that it is a partial order; symmetry, antisymmetry, transitivity, and so on cannot always be supposed as properties $<_P$ holds.

Once a graph $(X, <_P)$ is recognized as a qualitative meaning of P in X , the mappings $m_P: X \rightarrow [0, 1]$, verifying the axioms:

- (1) $x <_P y \Rightarrow m_P(x) \leq m_P(y)$
- (2) x is maximal in the graph $\Rightarrow m_P(x) = 1$
- (3) y is minimal in the graph $\Rightarrow m_P(y) = 0$

can be defined, and called measures of the meaning of P . Note that the interval of values $[0, 1]$ can be changed by any closed interval of the real line, and taking its extremes instead of, respectively, 0 and 1. As in the case of probabilities, these three axioms do not allow us to specify a unique measure, but more information on the measure's contextual characteristics is necessary for it.

For instance, in the former case of "big" with qualitative meaning given by $([0, 10], \leq)$, provided it were known that the measure should be linear, $m_{\text{big}}(x) = ax + b$, from (2) and (3), it follows the single measure $m_{\text{big}}(x) = x/10$, also verifying (1) because it is a nondecreasing function; but provided it were contextually known that the measure should be quadratic, $m_{\text{big}}(x) = ax^2 + bx + c$, because then it should be $c = 0$, $100a + 10b = 1$, and $2ax + b \geq 0$ for all x in $[0, 10]$, many quadratic measures would be possible, for instance, $x^2/100$, $x^2 + 0.99x$, $-2x^2 + 20.1x$, and the like.

In any case, a full use of P in X associated with the relationship "less P than," is given by the quantities $(X, <_P, m_P)$; each time one of these quantities is specified, a full meaning is scientifically designated a quantity, or specified, for P in X . It should be noted that all measures preserve the relationship "equally P than":

$$x =_P y \Leftrightarrow x <_P y \text{ and } y <_P x \Rightarrow m_P(x) \leq m_P(y) \text{ and } m_P(y) \leq m_P(x), \text{ or } m_P(x) = m_P(y).$$

In this form, each meaning P can have in X can be seen as a quantity.

What about precise words? Because a precise word P specifies, by definition, a subset \mathbf{P} in the universe of discourse, all elements in \mathbf{P} are prototypes and it is $x =_P y$, for each pair x, y of them, and those in \mathbf{P}^c are antiprototypes with each pair also being equally not P ; hence, the first are maximal and the second minimal. The universe is partitioned in the maximal and the minimal, each class having equal measure for all its elements. Consequently, there is just a single measure specified by

$$m_P(x) = 1 \text{ if } x \text{ is in } \mathbf{P}, \text{ and } m_P(x) = 0 \text{ if } x \text{ is in } \mathbf{P}^c,$$

which is just the characteristic function of \mathbf{P} . The reciprocal is obvious, and then $m_P^{-1}(1) = \mathbf{P}$; thus, precise words P are specified by the single graph $(X, =_P, m_P)$, with the measure only taking its values in the subset $\{0, 1\}$ of the interval $[0, 1]$.

Note that it can be cases of words without prototypes or antiprototypes, also partitioning the universe in several subsets of equally P elements, but with measures that, constant in each of these parts, have values not 0 or 1, but only values in the open interval $(0, 1)$; of these words it could be said that their use, or meaning, is pseudo-precise in X . Changing the measure m by $m/\text{Sup } m$, provided $\text{Sup } m = \text{Max } m < 1$, at least an element with measure one would appear.

2.2. Once a measure m_P is specified, it defines in X a new, and linear, relation \leq_m ,

$$x \leq_m y \Leftrightarrow m_P(x) \leq m_P(y),$$

that is larger than the relation $<_P$, because

$$x <_P y \Rightarrow m_P(x) \leq m_P(y) \Leftrightarrow x \leq_m y.$$

That is, $<_P \subseteq \leq_m$, and, for coinciding, $<_P$ should be linear; nevertheless, in general, both relations are not coincidental, and the second being larger than the first gives a “larger and linear new meaning” (X, \leq_m, m_P) that, reached after m_P is known, can be called a *working meaning* of P in X ; noncomparable elements under $<_P$ are always comparable under \leq_m .

It can be said that measuring enlarges meaning by linearizing the qualitative meaning $<_P$ up to \leq_m . Once a working meaning substitutes the qualitative meaning, new working prototypes can appear, and working antiprototypes, those x in X , respectively verifying $m_P(x) = 1$ and $m_P(x) = 0$, that can be more than the original prototypes and antiprototypes. This deserves some comments.

The first comment refers to how a fuzzy set and its several membership functions can be understood. Up to now it is not well known what a fuzzy set in X with linguistic label P actually is, and it is usually confused with just one of their membership functions; but because, unless the linguistic label’s use is precise in the universe, there is not a single membership function characterizing the fuzzy set, such identification is actually a one-to-many correspondence. It is usual to consider that fuzzy sets are represented by the functions in $[0, 1]^X$, something only useful for purely mathematical purposes and once a membership function is specified or, better, designed. It is not the same with sets that can be quietly identified with the functions in $\{0, 1\}^X$ thanks to the unicity of their measure, or characteristic function. The concept of a fuzzy set is not only a matter of degree, as its originator Lotfi A. Zadeh likes to say, but its membership functions are a matter of careful design. To capture what is actually a fuzzy set is a question that should be initially posed in the setting of plain language.

In the first place, such a question refers to the empirical fact that predicative words “collectivize” in the universe of discourse, that they generate “linguistic collectives” well anchored in plain language. For instance, in the universe of London’s inhabitants, the word “young” generates the linguistic collective of “young Londoners”; in the universe of the real numbers the word “big” creates the linguistic collective of “big numbers”; in a universe of buildings the word “high” creates the linguistic collective of “high buildings”, and so on. Obviously, linguistic collectives are well understood by the speakers, but are kinds of gaseous or cloudy entities for which no criteria of individuation are known, and, inasmuch as, following W.V.O. Quine, “There is no entity without identity”, linguistic collectives should be approached through ways of which, right now, the one only at hand comes from the quantities specifying the meaning of the word at each universe of discourse.

It could be said that the linguistic collective P generates in X is in the qualitative state $(X, <_P)$, and that to each qualitative state it corresponds to being in several quantitative states each given by a measure m_P . Each full meaning given by a quantity $(X, <_P, m_P)$ shows both a qualitative and a quantitative state of the linguistic collective; it is a state of the collective. States reflect the available information on the qualitative and the quantitative use of P in X .

It is a view reflecting the several situations in which the collective can be seen, and under it, a fuzzy set in X is nothing else than the linguistic collective generated by its linguistic label. It just consists in renaming the collective of P in X as the fuzzy set labeled P . In the case where the linguistic label P is precise in X , the collective “vitrifies” in the set specified by P in X ; the collective has just the qualitative state $(X, =_P)$, and just the quantitative state given by its characteristic function m_P giving the vitrified, or crisp, quantitative state $m_P^{-1}(1)$. Precisely used linguistic labels are in a single state.

When the linguistic label is imprecisely used in X , the collective or fuzzy set will have as many qualitative states as qualitative meanings $(X, <_P)$ can be recognized, and each measure for them is nothing else than a membership function for the fuzzy set labeled P . Note, for instance, that there is no difference between the former measures m_{big} and the membership functions that can be attributed to the fuzzy set in $[0, 10]$ labeled “big”.

Conceptualizing a fuzzy set as just an abstract entity, or concept, existing in language, helps to refine Zadeh’s intuitive view that each membership function is an extensional meaning of the corresponding linguistic label.

In the imprecise case, the designer of a membership function of a fuzzy set in X with linguistic label P should proceed by taking into account all the information available to her or him on the use of P in X that is, usually, incomplete, sometimes not containing all the relations $<_P$, and to which often the designer still adds some reasonable hypotheses on the shape of the membership function that can be suitable for the current problem. For instance, and in the toy example of the fuzzy set in $[0, 10]$ with linguistic label “big”, the designer could consider that, with her or his current scarce information, the best that can be done is to take the simple and above-mentioned linear measure $x/10$, or if the designer can suppose it should be quadratic, just take its square $(x/10)^2 = x^2/100$. In short, the designer is often limited to consider some (possibly scarce) information of how P behaves in X , and some characteristics of the current problem for which the design is done. For instance, by estimating if the measure of “5 is big” should be 0.5, clearly less than 0.5, or clearly bigger than 0.5, and so on.

Hence in the praxis of fuzzy logic it cannot always be supposed that a membership function is actually a measure but, in the best case, that the membership function is a universal approximation of some measure; that is, a designed membership function μ_P could be seen as a good enough one provided some measure m_P exists such that, for instance, it would verify

$$|m_P(x) - \mu_P(x)| \leq \varepsilon, \text{ for all } \varepsilon > 0, \text{ and all } x,$$

or, if it is possible, that minimizes the function

$$\text{Sup } |m_P(x) - \mu_P(x)|.$$

Anyway, and in each case, this requires previously counting with a measure m_P that makes the membership function μ_P unnecessary. Provided it were proven that such a criterion of approximation actually exists, perhaps a nice existential theorem for characterizing good membership functions could be obtained and become useful for the processes of designing them.

In short, and in general, the designed membership functions only can be seen as potential approximations to measures, with which only a working meaning \leq_μ is often available and, being a linear relation (a partial order indeed), cannot always coincide with the qualitative meaning. In praxis, the modeling by fuzzy sets is just an uncertain approximation to the meaning of their linguistic labels, and true measures are but “ideal” membership functions such as the uniform probability $1/6$ reflects an “ideal die” in probability theory; for this reason, membership functions should be carefully designed on the basis of the best information available on the behavior of P in X , and with the best possible reasonable hypotheses on their shape coherently with the requisites of the current problem.

It should be pointed out that, in plain language and ordinary reasoning, it is sometimes difficult and even unnecessary, to attribute numbers for measuring the meaning; for instance, there are no rare expressions such as “It is highly possible that John is rich,” “The degree up to which Jane is wise is up to the middle,” or “Susan is extremely intelligent.” Hence, in praxis it could be suitable to substitute the interval $[0, 1]$ in which the measure’s values range by sets such as the subintervals in $[0, 1]$ or the fuzzy numbers in $[0, 1]$ or even a set of linguistic labels. With them, the meaning of the former examples can be measured by, respectively, a fuzzy number μ_{high} , the interval $(0.7, 1]$, and the word “extremely”. Because fuzzy numbers and intervals are not totally ordered, these kinds of values can still present the advantage, when $<_P$ is not linear, of having more possibilities for the coincidence between the qualitative and the working meanings. Anyway, the set of such values should be endowed with some algebraic ordering, necessary for defining a measure, and contextually coming from their uses.

In science complex numbers are sometimes taken to measure some variables and, in the same vein, instead of the unit real interval $[0, 1]$, the complex unit interval $\{a + ib; a, b \text{ in } [0, 1]\}$ can be taken by submitting the measure to verify the three former axioms but with its values in the complex unit interval, endowed with the natural (partial) order of complex numbers:

$$a + bi \leq^* c + di \Leftrightarrow a \leq c \ \& \ b \leq d,$$

with maximum and minimum, respectively, $1 + i$, and $0 + 0i = 0$, giving a non-linear working relation $\leq_m = \leq^*$ for all measures m ; now the working meaning is

not a total or linear relation, but a partial ordering. Because the complex unit interval can be seen as isomorphic with the set of closed subintervals of $[0, 1]$, taking one or another as the range for the measures does not matter. Once accepted as a natural way of ordering such complex numbers, intervals, fuzzy numbers, words, and the like, as well as which are their respective maximum and minimum, it is necessary to take the suitability for the considered problem of the calculus with them into account, something that should be done at each concrete problem for having, provided it were the case, a computation with the corresponding words, or statements.

2.3. Let's apply what has been said to the (debatable) concept of truth, deriving from the word "true" applied to statements; it is usual to identify expressions such as "What you say is true" and "You are saying the truth," but here "true" is understood as a predicative word, and not as naming the concept "truth" which can be seen as its mother-predicate. Concepts are but abstractions generated after using a word as its mother-predicate to either physical or virtual objects, and usually once such a word migrated between several universes and suffered some more or less slight modifications in its respective meanings. It is not "tall" that comes from the concept of "tallness", but this (abstract) concept was generated after applying "tall" to several collections of objects including trees, mountains, people, and so on, and passing from one to another by analogy; each time, "tall" refers to some particular objects, but "tallness" refers to all of them. In the same vein, "true" refers to each statement, but "truth" refers to all of them.

Truth is a concept that, understood as an absolute and universal one, has been conducive to aberrations and, indeed in both its past and present human history, carries terrifying consequences coming from such an understanding of "Truth" with a capital letter.

It should be pointed out that the qualitative meaning of a word applied to the limiting case of a universe of discourse just consisting in a singleton $\{x\}$, is reduced by the reflexive property to the minimum relation $<_P = \{(x, x)\}$, of which nothing can follow, and, first of all, it suggests some comments referring to when it can be said that the meaning of a word is "metaphysical", that a word is meaningless in some universe of discourse and hence cannot generate a concept; that, out of the singleton, it does not collectivize.

Once a qualitative meaning $(X, <_P)$ is captured, it can be said that P is pseudo-measurable in X , because it is what allows defining measures for P and there exists at least the one defined by assigning 1 to the prototypes, 0 to the antiprototypes (provided both exist), and a fixed and common value to the other elements in X , for instance, 0.5, even if these measures could have nothing to do with the context in which the word is used and consequently do not reflect anything but a lot of ignorance on P in X . When at least a measure related to the context can be specified, it can be said that P is measurable in X , or, perhaps better, is effectively measurable; for instance, "big" is effectively measurable in $[0, 10]$. When no relation $<_P$ can be captured, it can be said that the word is meaningless in X , that its use is currently metaphysical because it is not even pseudo-measurable, and

measurability is a main characteristic science requires for the predicates it manages. It is worth to remember Lord Kelvin's shortened statement, "If you cannot measure it, it is not Science".

That P is meaningless in X does not imply that it should also be meaningless in all possible universes of discourse; for instance, "big" is measurable with real numbers, but meaningless if applied to dreams, because it seems actually impossible to recognize previously when a dream is "less big than" another one. Nevertheless, what cannot be inferred is that a nonmeasurable "big" can never be useful for reaching some new idea; stating "a dream is big" could serve as a useful metaphor, or analogy, able to induce the study of something related to dreams; P is meaningless in X does not imply that P cannot excite someone towards a further searching, but psychological matters are beyond what we are trying to analyze here, even if it is manifest that metaphorical thinking is important for conducting creative reasoning; provided, of course, it were to be produced jointly with having some knowledge of the corresponding subject as pointed out before with the Kekulé example. It should be noted that, although for scientific purposes measurability is essential, neither all that is relevant for ordinary reasoning is measurable, nor what is measurable always important for ordinary reasoning.

What about the meaning of the word $T = \text{true}$ in a set $X [P]$ whose elements are the elemental statements x is P , for x in X , and after knowing that P is measurable in the universe X ? The first to be captured is the relationship "less true than":

$$x \text{ is } P \text{ is less true than } y \text{ is } P \Leftrightarrow x \text{ is } P <_T y \text{ is } P,$$

and the second is recognizing which are the maximal and minimal statements in the graph $(X [P], <_T)$, provided they were to exist, that is, specifying a qualitative primary meaning of T in $X [P]$. Note that, instead of X and P , it could be considered, for instance, the union universe $X \cup Y$, and the two words $\{P, Q\}$, to capture

$$x \text{ is } P \text{ is less true than } y \text{ is } Q,$$

for x in X and y in Y , with P acting in X and Q in Y with respective qualitative meanings $(X, <_P)$, and $(Y, <_Q)$; but, for simplicity, and even if several words could be taken into account, only one universe and a single word are considered right now.

True = T names a property of the elements in $X [P]$ referring to the actual verification of the property named P for the elements in X , the reality of the statement " x is P ". That is, the character of "true", a statement " x is P " can show, is directly related to the verification by x of the property named by P (as more P is x , more true is x is P), and that, hence, some relation between the qualitative meanings $<_P \subseteq X \times X$, and $<_T \subseteq X [P] \times X [P]$, should exist. It is supposed for it that "If $x <_P y$, then x is $P <_T y$ is P ". It seems, consequently, that for specifying a measure of T it should be linked with one of P , and the question is how it can be done.

Inasmuch as it is $m_P: X \rightarrow [0, 1]$ and $m_T: X [P] \rightarrow [0, 1]$, let it be $t: [0, 1] \rightarrow [0, 1]$ a nondecreasing mapping such that $t(0) = 0$, and $t(1) = 1$, namely an order morphism of the ordered unit interval $([0, 1], \leq)$, under which conditions can it be $m_T(x \text{ is } P) = (t \circ m_P)(x)$ a measure of T ?

Because $x <_P y$ implies $x \text{ is } P <_T y \text{ is } P$, it is

$$m_P(x) \leq m_P(y), \text{ and } t(m_P(x)) \leq t(m_P(y)), \text{ or } m_T(x \text{ is } P) \leq m_T(y \text{ is } P),$$

that is, the first axiom of a measure for T is verified. For what concerns the other two axioms relative to maximal and minimal elements, it is immediate that if x is maximal (resp., minimal) for $<_P$ it follows $m_T(x \text{ is } P) = t(1) = 1$ (resp., $m_T(y \text{ is } P) = t(0) = 0$), but the question seems to be opened when “ x is P ” is maximal, or “ y is P ” is minimal for $<_T$ without necessarily implying that x or y are, respectively, maximal and minimal for $<_P$. Nevertheless, by supposing that T is “coherent” with P , that is,

x is maximal (resp., minimal) for $<_P$ if and only if $x \text{ is } P$ is maximal (minimal) for $<_T$, the problem is solved. Because coherence is not a bizarre condition, under it, $t \circ m_P$ is a measure of T .

It is classically and usually said that $m_T(x \text{ is } P)$ is a “degree of truth” of $x \text{ is } P$, thus when it is $m_T = t \circ m_P$, the degrees of truth are obtained through a “truth-function” t . Then it should be noticed that it is $m_T(x \text{ is } P) = m_P(x)$ for all x in X , provided $t = \text{id}_{[0, 1]}$, a case in which the degree of truth of “ $x \text{ is } P$ ” just coincides with that in which $x \text{ is } P$, as it is classically understood. Were, for instance, $t(x) = x^2$, it would be $m_T(x \text{ is } P) = m_P(x)^2$, and then, if $x \text{ is } P$ with degree 0.7, “ $x \text{ is } P$ is true” is with degree 0.49; if it were $t(x) = x^{1/2}$, then $x \text{ is } P$ would be true with degree $\sqrt{0.7} = 0.836$. Of course, the truth-function t should be chosen in each case according to the information available on the characteristics of the current situation, and by searching if, under them, the degrees of T should be lower or bigger than those of P , if t is, respectively, contractive ($t(x) \leq x$), or expansive ($x \leq t(x)$) for the x in X .

Summing up, provided T were coherent with P , truth-functions t would allow us to obtain measures of T from those of P , and whenever t is one to one and onto (bijective, an order automorphism of the unit interval) there are no more statements with measure one or zero for T than, respectively, the prototypes and the antiprototypes of P ; nevertheless, a nonbijective truth-function, such as

$$\begin{aligned} t(x) &= 0 \text{ for } x \text{ in } [0, 0.4], \\ t(x) &= 1 \text{ for } x \text{ in } [0, 6, 1], \text{ and} \\ t(x) &= 5x - 2 \text{ for } x \text{ in } [0.4, 0.6], \end{aligned}$$

shows more statements with one or zero degrees of true than the prototypes and the antiprototypes of P .

2.4. Plain languages are not the creation of a single person, but are slowly generated by linguistic interactions between groups of people; plain languages are socially constructed over the course of time, and words acquire meaning along such interactions. Without sharing common meanings, people cannot actually understand each other, and communicating is very difficult if not impossible; linguistic meaning is a social construction that, to some extent, deserves a symbolical analysis to see how people can arrive at sharing a common meaning for words.

For such a goal, suppose that a person p_1 manages a word P in X under the qualitative meaning given by the graph $(X, <_P^1)$, and that this person utters to another p_2 elemental statements “ x is P ”. Person p_2 will understand what p_1 is saying just provided she either captures the relation $<_P^1$, or, at least a nonempty part of it. That is, she manages P with a qualitative meaning $(X, <_P^2)$ such that the intersection of the respective relations $<_P^1$ and $<_P^2$ is not empty; on the contrary, if such an intersection is empty, p_2 cannot understand what p_1 tells her. To understand what p_1 says, p_2 can accept as the meaning of P , either all $<_P^1$, or a part of it, that will be denoted by $<_P^2$ and become a common meaning of P for both p_1 and p_2 . Provided a third person p_3 enters later on in the conversation, the situation would be repeated, and a common meaning by the three people is reached by either the intersection of the three corresponding meanings, or by accepting a part of that meaning previously accepted by the first two, and so on.

Of course, when time passes, P can migrate from X to another universe Y , and so on; with it, and at the end, one or several meanings can result associated with the word P . Of course, it can happen that p_2 has no meaning for P , or that P is an unknown word to her, and, then, either the communication between p_1 and p_2 is impossible, or p_1 explains to p_2 the meaning of P by, for instance, practical exemplification, such as when exemplifying “the door is closed”, by opening the door and saying “open”, and closing the door and saying “closed” through a practical, or visual, description of the meaning of P .

Concerning how a quantity reflecting a full meaning of P in X can finally appear, a way for it could consist in aggregating the several measures assigned by each successive person in such a way that the aggregation preserves the verification of the three axioms a measure should satisfy. For instance, given the measures each of n people $(m_P^k, 1 \leq k \leq n)$ assigns to its respective qualitative meaning $(<_P^k)$, the function

$$m_P(x) = \min (m_P^1(x), \dots, m_P^n(x)),$$

is a measure for the not empty intersection \langle_p of all relations \langle_p^k because:

- (1) $x \langle_p y \Leftrightarrow x \langle_{p^1} y \ \& \ \dots \ \& \ x \langle_{p^n} y \Rightarrow m_p^1(x) \leq m_p^1(y) \ \& \ \dots \ \& \ m_p^n(x) \leq m_p^n(y) \Rightarrow \min(m_p^1(x), \dots, m_p^n(x)) \leq \min(m_p^1(y), \dots, m_p^n(y)) \Leftrightarrow m_p(x) \leq m_p(y)$.
- (2) If x is a maximal for \langle_p , then it should be a maximal for all the \langle_p^k . Hence, $m_p(x) = \min(1, \dots, 1) = 1$.
- (3) If x is minimal for \langle_p , then it should be minimal for at least one \langle_p^k , and because between the brackets of the min at least one zero will appear, it is $m_p(x) = 0$.

Of course, min is not the only possible function allowing that proof; any n -place function being nondecreasing at each place, taking the value one for the argument $(1, \dots, 1)$, and the value zero for those arguments containing at least one zero, permits repeating the proof. In sum, there are many ways for reaching a common full meaning, and some of them can be symbolically represented and linguistically interpreted.

It should be pointed out that what has been developed does not pretend to be the only way in which all common meaning is reached; it simply tries to show that there are ways for it that are describable through a symbolic formalism. Anyway, the full theoretical problem still remains an open one in which the role analogy plays in it should be considered, and for whose solution the concurrence of some evidence not only obtainable by means of clever observation seems necessary, but especially by well-designed controlled processes of experimentation on how meaning evolves in plain language. Possibly such types of problems require a new experimental methodology for studying language, supported by what computer technology can facilitate today for reaching it.

The importance words' migration can have was mentioned before, therefore let's add something in this respect; at the end, everything that has been reported in a symbolic form has at its back Ludwig Wittgenstein's ideas on meaning as use, family resemblances, and language games, presented in his (posthumous) second book, *Philosophical Investigations*.

In this respect, an example could be in order: the word "big" applied to numbers can be seen as a migration of "tall" applied to people after the centimeter to state numerical height is introduced. Indeed, if John's height is 190 cm, because 190 is a big number between those in $[0, 200]$, interpreting "John is tall" as the linguistic evaluation of John's height as "big", is seen the migration between people and numbers of "tall" into "big". In an analogous way, "small" can be imagined as a migration of "short" from people to numbers, and further said that these pairs of words show a kind of linguistic family resemblance going from linguistically playing with people to playing with numbers. Perhaps this comment could help to find a (mathematical) way to analyze a linguistic phenomenon deeply under which people often learn how to understand and manage words, how to play with them in plain language.

Some naïve comments on syntax and semantics aspects are still in order, and once clearly said that if both aspects are basic for managing a plain language well,

the meaning of statements not syntactically well constructed is often well understood, as happens, for instance, with children's speech. But, without capturing the meaning of statements, its semantics, there is no way of understanding them; for instance, sometimes children signalize something of what they refer to and thus what they are trying to say is recognized. For comprehending a language, semantics is essential, but syntax is often, although important, an accessory; for instance, in a talk on pigeons it does not matter very much if someone ignores how "pigeon" is written, but what is essential is not confusing a pigeon with an eagle, or a raven, or a vulture, and so on, and less again with something not being a flying bird, that is, recognizing the meaning of "pigeon" in the universe of "flying birds".

In logic, complex statements are supposed to be constituted of elemental statements " x is P ", " y is Q ", and the like, joined by connectives such as "and", "or", "not", and, sometimes in the conditional form "if/then", or affected by the only quantifiers "exists" and "all". But in the case of plain language, and in addition, those statements are often affected by linguistic modifiers such as "they are 'very'" and "more or less", and linguistic quantifiers, such as "many", or "several", or "few", as it is done with fuzzy sets. Then, and both in classical logic and in fuzzy logic, the meaning of a complex statement is supposed to be captured from the meaning of its components, after knowing which are the meanings of conjunction, disjunction, negation, conditionals, and so on, and presuming they are previously and contextually specified by either characteristic or membership functions. Nevertheless, in ordinary situations it is often the case of first capturing the meaning of a full complex statement, and second those of its components thanks to the contextual meaning the full statement facilitates for them.

Anyway, and for a theoretical study, the, let's say classical, way of constructively studying the meaning of complex statements through its parts cannot be avoided, hence, instead of a (currently unknown) synthetic and systematic procedure for directly analyzing the meaning of complex statements, it seems suitable, for a first symbolic study, to attempt one of analytic type in which the meaning of the components already takes into account the context in which they are inscribed. The next chapters are devoted to such a task, and, as always in this book, without presupposing laws that can reduce the analysis to be enclosed in a restricted, and up to some extent artificial, mathematical framework.

2.5. Finally, it should be pointed out why here it is preferable to say "ordinary, or plain language", instead of the usual expression "natural language", a preference just coming from the adjective "natural" used to show its opposite character to the axiomatic and "artificial languages" typical of formal logic and basically used to expose proofs clearly being sure that they contain no jumps. Because natural/plain language, during the process of representing, has to adapt itself to the pressures of its own capabilities and the corresponding representing goals, it never finishes being as "natural" as it had, presumably, begun; and this is because at the beginning it was still not committed to a specific work. There is no way of faithfully representing language and reasoning except with plain language; there can be thinking without language, but not human reasoning without language.

In addition, each plain language is permeated by traits coming from the cultural environment of their speakers; even English, today the almost universally common language in science, at least, is currently being influenced by the ways of speaking and writing it by people whose native language is not English, or whose native country is not an English-speaking one. In sum, plain language is not properly “natural” in the same sense that the brain or the Amazon forest is natural, but a result of many cultural, historical, geographical, and intellectual influences. Ordinary people not having attended school and not writing or reading their own plain language well, still speak it well enough and communicate easily with people educated in universities; the first perhaps don’t manage it perfectly from the point of view of syntax, but well enough from that of semantics for conducting the second to capture the meaning of what they express.

It should be remarked that the same ownership title of Spanish belongs to a farmer in Mexico, or a Chilean living in Paris, the President of Argentina, the King of Spain, or a writer who won the Nobel Prize by his work originally written and published in Spanish. Plain languages are a shared common property of, at least, all its native speakers, and are among the most complex dynamical systems today science is faced with; new perspectives are needed for its scientific domestication.

In the end, mechanizing a plain language towards undoing the Gordian knot of artificial intelligence cannot be done without knowing its use, and once represented in a form preserving its flexibility. A form not previously constrained by logical laws not necessarily holding in plain language, as is the case with the commutative law of conjunction ($p \& q = q \& p$), almost always accepted in logic, but that, because time often intervenes in plain language, is a law that cannot be universally supposed, and as shown, for instance, by “She enters the room and starts crying”, and “She starts crying and enters the room”, two statements depicting different situations, and whose identification can consequently lead, later on, to committing some mistakes.

In the same vein is the identification of conditional statements “If p , then q ” ($p \rightarrow q$), with the affirmative statements “not p , or q ” ($p' + q$), regardless of, for instance, whether the antecedent’s negation (not p) can be suitably described and rightly represented after capturing its meaning in the corresponding setting, or if the underlying formal framework allows us to follow q from p and $p \rightarrow q$ (modus ponens, MP), as happens in the frame of Boolean algebras, but neither with $p \cdot q = p' + q$, in those of De Morgan algebras and ortholattices, nor generally in those of the standard algebras of fuzzy sets, where the inequality $p \cdot (p' + q) \leq q$, formally representing the rule of modus ponens, does not hold for all the pairs p, q , and all representations of “and” (\cdot), “or” ($+$), “not” ($'$), and “it follows” (\leq).

For instance, if such inequality is considered in a setting endowed with a De Morgan algebra framework, by taking any element p , and $q = 0$, it follows that $p \cdot (p' + 0) \leq 0 \Leftrightarrow p \cdot p' = 0$; that is, p is one of the Boolean elements in the De Morgan algebra and, hence, the inequality does not hold for any pair p, q . Of course, in a setting endowed with a Boolean algebra’s structure, and because both the distributive and the noncontradiction laws hold in it, it follows that $p \cdot p' + p \cdot q = 0 + p \cdot q = p \cdot q \leq q$, and the MP-inequality $p \cdot (p' + q) = p \cdot q \leq q$, holds for all pairs p, q .

In the case where the setting is of fuzzy sets μ, σ , and so on, endowed with the framework given by a standard algebra, the MP-inequality should be expressed by the functional inequality

$$\mu \cdot (\mu' + \sigma) \leq \sigma, \text{ or } T \circ (S \circ (N \circ \mu \times \sigma)) \leq \sigma,$$

with T a continuous t -norm representing the linguistic “and”, S a continuous t -conorm representing the linguistic “or”, and N a strong negation function for the linguistic “not”, whose study can be reduced to solving the (numerical) functional inequality

$$T(a, S(N(a), b)) \leq b, \text{ for all } a, b \text{ in } [0, 1],$$

in which $b = 0$ ($N(0) = 1$) implies $T(a, N(a)) = 0$, showing the MP-inequality cannot hold for all pairs of numbers a, b , except for some triplets (T, S, N) such as the one constituted by Lukasiewicz’s t -norm, $W(a, b) = \max(0, a + b - 1)$, its dual t -conorm, $W^*(a, b) = \min(1, a + b)$, and the strong negation function, $N = 1 - \text{id}$:

$$\begin{aligned} W(a, W^*(1 - a, b)) &= W(a, \min(1, 1 - a + b)) \\ &= \max(0, a + \min(1, 1 - a + b) - 1) = \min(a, b) \leq b, \end{aligned}$$

for all pairs of numbers a, b in $[0, 1]$. Notice that with $T = \min$, $S = \max$, and $N = 1 - \text{id}$, the inequality $\min(a, \max(1 - a, b)) \leq b$ does not hold for the pairs $a = 0.5$ and $b = 0$; and, analogously, with the triplet given by $T = \text{prod}$, $S = \text{prod}^* = \text{sum} - \text{prod}$, $N = 1 - \text{id}$, the inequality is $\text{prod}(a, \text{prod}^*(1 - a, b)) = a \cdot (1 - a + b - (1 - a) \cdot b) = a \cdot (1 - a + a \cdot b) \leq b$, and it also does not hold for $a = 0.5$ and $b = 0$.

2.6. To end this chapter, and even if later on the subject is reconsidered, let’s advance something of a preliminary character on the use in ordinary language of the words “uncertain”, and “probable”, that is, on describing and differentiating their respective qualitative and quantitative meanings; it is a subject that, linked with what has been presented on the meaning of words, is in touch with the debate between the two main interpretations of the probability mathematical concept, the objective and the subjective. The first comes from an observed convergence of the outcomes’ frequencies in random experiments whose possible outcomes can be well described, and the second from the experienced opinion of a “rational person” assigning a priori probabilities to events, that, once transformed into a posteriori probabilities by means of the Bayes formula, helps whoever, based on them, wants to take the risk of betting some money on their appearance. In both cases, nevertheless, a common belief is shared on the actual possibility of perfectly classifying the universe by a union of disjoint classes, something necessary for posing the additive law of probability in which both interpretations coincide, and that, jointly with assigning a probability equal to one to the sure event, allows us to obtain the law for the probability of the negation of outcomes.

Notwithstanding, the understanding of events as subsets, also underlying both interpretations, corresponds to naming the events by precise words, something that is not always the case in ordinary reasoning expressed in plain language, as it is not always to assign numerical probabilities as can be exemplified by typical utterances such as, “It is highly probable than John is rich,” or “It is improbable that Laura will join us,” and so on. The meaning of “probable”, the mother-predicate of the probability’s concept, is still to be clarified in plain language.

Because the meaning of “uncertain”, mother-predicate of the uncertainty’s concept, is also not clarified, and because uncertainty, as with imprecision in language, permeates almost all branches of science, there is confusion between probability and uncertainty deserving to be clarified. All these questions are considered further.

References

1. E. Trillas, C. Moraga, S. Termini, A naïve way of looking at fuzzy sets. *Fuzzy Sets Syst.* **292/C**, 380–395 (2016)
2. E. Trillas, On a Model for the Meaning of Predicates, in *Views of Fuzzy Sets and Systems from Different Perspectives*, ed. by R. Seising (Springer, Berlin, 2009), pp. 175–205
3. E. Trillas, R. Seising, On meaning and measuring: A philosophical and historical view. *Agora* **34(2)**, 1–23 (2015)
4. W.V.O. Quine, *Philosophy of Logic* (Prentice-Hall, New York, 1970)
5. J. Bairwise, J. Perry, *Situations and Attitudes* (The MIT Press, 1983)
6. L. Wittgenstein, *Philosophical Investigations* (Basil Plackwell, Oxford)
7. P. Halmos, *Naïve Set Theory* (Van Nostrand, New York, 1960)
8. E. Trillas, On the use of words and fuzzy sets. *Inf. Sci.* **176(11)**, 1463–1487 (2006)
9. E. Trillas, L. Eciolaza, *Fuzzy Logic* (Springer, Berlin, 2015)
10. B. Russell, *The Problems of Philosophy* (Williams & Norgate, London, 1912)
11. E. Trillas, S. Guadarrama, Fuzzy representations need a careful design. *Int. J. Gen Syst.* **39(3)**, 329–346 (2010)
12. E. Trillas, C. Moraga, Reasons for a careful design of fuzzy sets, in *Proceedings of the EUSFLAT* (2013), pp. 140–145
13. E. Trillas, I. García-Honrado, Unended Reflections on Family Resemblances and Predicates Linguistic Migration, in *Proceedings of the EUSFLAT* (2011), pp. 598– 604
14. E. Trillas, A. Sobrino, C. Moraga, On family resemblances with fuzzy sets, in *Proceeding of the IFSA-EUSFLAT* (2009), pp. 306–311
15. Lord Kelvin (W. Thompson), Lecture on electrical units of measurement, in *Popular Lectures and Addresses*, vol. I (1889), p. 73
16. A.N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, New York, 1956)
17. B. De Finetti, Sul significato soggettivo della probabilità. *Fundam. Math.* **XVII**, 298–329 (1931)