Vagueness and Fuzzy logic

1 Fuzzy sets and fuzzy logic

Standard set theory: a set corresponds with a characteristic function from D to $\{0, 1\}$.
Fuzzy set theory: a fuzzy set corresponds with a function from D to $[0, 1]$.
Thus, if $D = \{x, y, z\}$, m_A is a fuzzy set s.t. $m_A(x) = 0.2, m_A(y) = 0.9, m_A(z) = 0.3$, for example.

In fuzzy logic, also sentences denote values in $[0, 1]$.

Why should we be interested in fuzzy logic for natural language?

1. Vague predicates don't have sharp cutoff points. $\rightarrow$ no finite number of truth values seems to be enough.
2. It perhaps can be argued for in terms of comparatives.
3. Truth values behave compositionally.
4. It treats the sorites paradox very nicely.

Ad (1). Obvious.

Ad (2). One argument (Forbes 1983, pp. 241-242):
Consider a pair of men, a and b, such that

(1) a is taller than b.

We can infer that

(2) a. a is tall to a greater degree than b, so
b. a satisfies the predicate 'is tall' to a greater degree than b, so
c. 'a is tall' has a higher degree of truth than 'b is tall'.

Ad (3). Given a primitive truth value (or degree) of atomic formulas, we can determine the value of complex formulae as follows (according to Lukasiewcz):

1. $m_0(\neg p) = 1 - m_0(p)$
2. $m_0(p \wedge q) = min\{m_0(p), m_o(q)\}$
3. $m_0(p \vee q) = max\{m_0(p), m_0(q)\}$
4. $m_0(p \rightarrow q) = 1$, if $m_0(p) \leq m_0(q)$, and $1 - (m_0(p) - m_0(q))$ otherwise.

m_0 is *not interactive*, meaning that $m_0(p \wedge q) \in \{m_0(p), m_0(q)\}$. Alternative treatments of conjunction and disjunction do not have this property and are *interactive*:

$m_1(p \wedge q) = m_1(p) \times m_1(q)$ (problem: $m_1(p \wedge p) < m_1(p)$, if $0 < m_1(p) < 1$)
$m_1(p \vee q) = m_1(p) + m_1(q) - m_1(p) \times m_1(q)$ (problem: $m_1(p \vee p) > m_1(p)$, if $0 < m_1(p) < 1$)

$m_2(p \wedge q) = |m_2(p) + m_2(q) - 1|$
$m_2(p \vee q) = min\{1, m_2(p) + m_2(q)\}$

$m_3(p \wedge q) = min\{m_3(p), m_3(q)\}$, if $m_3(p) = 1$ or $m_3(q) = 1$, 0 otherwise
$m_3(p \vee q) = max\{m_3(p), m_3(q)\}$, if $m_3(p) = 0$ or $m_3(q) = 0$, 1 otherwise

A generalization A t-norm $\mathbf{t}$ is a binary function from $[0,1]^2$ to [0,1] that is commutative, associative and monotone increasing with 1 as neutral element and 0 as zero element. That means that for arbitrary $x, y, z, u, v \in [0,1]$ the following holds:

1. $x\mathbf{t}y = y\mathbf{t}x$
2. $x\mathbf{t}(y\mathbf{t}z) = (x\mathbf{t}y)\mathbf{t}z$
3. if $x \leq u$ and $y \leq v$, then $x\mathbf{t}y \leq u\mathbf{t}v$
4. $x\mathbf{t}1 = x$ and $x\mathbf{t}0 = 0$

For any t-norm t one can define the *intersection* $\cap_{\mathbf{t}}$ for vague sets as follows:

$$m_{A \cap_{\mathbf{t}} B}(x) = m_A(x)\mathbf{t}m_B(x), \text{ for all } x \in D$$

Now one can derive $m_0, ..., m_3$ as special cases, if one defines:

$x\mathbf{t}_0 y = min\{x, y\}$ for $x, y \in [0,1]$
$x\mathbf{t}_1 y = x \times y$ for $x, y \in [0,1]$
$x\mathbf{t}_2 y = |x + y - 1|$ for $x, y \in [0,1]$

$x\mathbf{t}_3 y = min\{x, y\}$, if $x = 1$ or $y = 1$ for $x, y \in [0, 1]$, 0 otherwise

They all are extensions of classical intersection (correspond in case of $\{0, 1\}$)

The dual of the t-norn is the s-norm, which can be defined as follows:

$$x\mathbf{s}_{\mathbf{t}} = 1 - (1 - x)\mathbf{t}1 - y)$$

For any s-norm $\mathbf{s}$ one can derive the *union* $\cup_{\mathbf{s}}$ for vague sets as follows:

$$m_{A \cup_{\mathbf{s}} B}(x) = m_A(x)\mathbf{s}m_B(x), \text{ for all } x \in D$$

Now one can derive $m_0, ..., m_3$ for disjunction as special cases:

$x\mathbf{s}_0 y = max\{x, y\}$
$x\mathbf{s}_1 y = x + y - (x \times y)$
$x\mathbf{s}_2 y = min\{1, x + y\}$
$x\mathbf{s}_3 y = max\{x, y\}$, if $x = 0$ or $y = 0$, 1 otherwise

All of these are extensions of classical union (correspond in case of $\{0, 1\}$)

Ad (4) The Sorites
H1. $Bald(x_0)$ value 1.
IH_1 $Bald(x_0) \rightarrow Bald(x_1)$ value almost 1
...
IH_{100} $Bald(x_{99}) \rightarrow Bald(x_{100})$ value almost 1.
C $Bald(x_{100})$ value 0.

We naturally assume H1 and C. It is also not unnatural to assume IH1 – IH100. But if we also assume that C only follows from premisses $A_1, ..., A_n$ just in case the value of C is at least as high as the lowest value of the premisses, we have to conclude that the conclusion doesn't follow. This treatment of the sorites is natural, because it explains why we tend to accept each of the induction hypotheses.

2 Reasons to be sceptical to use Fuzzy Logic for NLS

There are reasons to be sceptical about fuzzy logic as a master theory for dealing with vagueness in natural language.It is true that soon after Zadeh's original paper in 1965, fuzzy logic was appliedto account for vagueness, approximators, and comparatives in natural language (Goguen, 1969; Machina 1972, Lakoff, 1973). According to the vast majority of linguists and philosophers, however, there are considerable problems with this proposal, as shown in particular inFine (1975) and Kamp (1975), but also in Klein (1980), Williamson (1994), and Kamp & Partee (1995).

The simplest way to illustrate the limitations of fuzzy logic is perhaps the following: Suppose Bert is tall to degree 0.5. The standard treatment of negation in Fuzzy Logic is that it reverses truth values so that 'Bert is not tall' is also true to degree 0.5. Fine (1975) and Kamp (1975) already objected that fuzzy logicians would not treat 'Bert is tall and Bert is not tall' as a contradiction, i.e., of having value 0. Instead this sentence is counterintuitively treated as having the value 0.5, the same value as 'Bert is tall'.

Fuzzy logic makes similar wrong predictions for many other complex sentences as well. Suppose Fred is tall to degree 0.4. Now consider, for instance, (a) 'Fred is tall and Bert is not tall' and (b) 'Fred is tall and Bert is tall'. These sentences should have the same value (i.e., 0.4) according to fuzzy logic. But it seems that (a) must be false: if Fred is shorter than Bert, it cannot be that Fred is tall and Bert is not. But a fuzzy logician wouldn't treat (b) as false, but true to some positive degree. As yet another example. Consider also the conditionals (c) 'If Fred is tall, Bert is tall' and '(d) 'If Fred is tall, Bert is not tall'. If the value of these sentences depends only on the values of their parts, they should have the same value. But this prediction is wrong: while (c) is plausibly true, (d) is certainly wrong.

Fuzzy logicians sometimes claim that they can straightforwardly account for comparatives. Fuzzy logicians analyze a comparative like 'John is taller than Bill' by saying that the sentence 'John is tall' must be truer than 'Bill is tall'. Now suppose that John is 6 feet 9 inches and BIll is 6 feet 8 inches. Intuitively, although John is taller than Bill, they are both unquestionably tall and thus 'John is tall' and 'Bill is tall' are equally true. But this is in contradiction with the fuzzy logic account of comparatives. As another counterexample to the Fuzzy logic account of comparatives, look at an all-or-nothing predicate like 'acute'. Even though 'acute' is an all or nothing predicate, one still has to account for the fact that object x can be more acute than object y. This cannot be done – as Fuzzy logic deals with comparatives – by saying that the statement 'Object x is acute' is truer than the statement 'object y is acute'. Even if it could be done, how would a fuzzy logician account for a sentence with an approximator like 'This object is almost acute'? A final example involving comparatives, look at examples like 'John is taller than he is short' or 'John is cleverer than Mary is rich'. Intuitively, these comparatives are odd. But if the relation 'truer than' is all there is to comparatives, no explanation for this oddity can be found.

Even if fuzzy logic works for *one*-dimensional predicates like *tall* or *wide*, it doesn't work for *more* dimensional predicates like 'big' and 'clever'. It is not the case that we can say that for any two individuals x and y either (i) x is cleverer than y, or y is cleverer than x, or (iii) x and y are equally clever. The 'cleverer than' relation seems to give rise to a partial order, but fuzzy logic requires it to give rise to a weak order.

Just as probability, Keefe (2000) points out that also degrees of truth allows only for basically one type of transformation: identity. This is problematic for two reasons. First, it is inconsistent with claims of fuzzy logicians (Machina, 1976) that "the assignment of exact values doesn't matter much. [...] what is of importance instead is the ordering relation between the values of the various propositions." (suggestions that degrees just represent an ordinal scale). Second, and more seriously, it gives rise to the question where the precise numbers come from. How do we know what the exact truth value of a sentence is?

Fuzzy logicians argue that because of *higher order vagueness*, no finite number of truth values will do. Thus we need fuzzy logic, which has an infinite number of truth values and thus no clear cut-off points. But also fuzzy logicians do not avoid all sharp boundaries. There still should be a sharp boundary between individuals which have property P to degree 1 and those who don't. But what is the last man who is tall to degree 1?

A related problem involves the analysis of modifiers proposed by Zadeh (1975). He proposed to analyze modifiers like *very* and *more or less* as follows:

$m(Very(tall(John))) = (m(tall(John)))^2$

$m(more-or-less(tall(John))) = (m(tall(John)))^{\frac{1}{2}}$

But this treatment gives rise to several problems. First, why can you say 'very old', 'very young', but not 'very middle aged'? Second, if John is 70, he is old. This presumably means that he is old to degree 1. But then it follows that John is also Very(very(very(...(very old)...))), which seems wrong. Perhaps this just means that John is not old to degree 1, but then who is, if anyone? (well ... perhaps it just means that only *absolute* adjectives (like *full, closed,...*) have values in $\{0, 1\}$, but not *relative* adjectives (like *tall, long,...*)).

A final(?) problem for fuzzy logicians is that we would like to distinguish between different sources of vagueness. For relative adjectives we can say both 'John is very tall', but also 'John is definitely tall'. Presumably, fuzzy logicians would account for both in terms of high degrees. A problem for them is that for other adjectives there is a difference: we can say 'Mary is definitely pregnant', but we cannot (seriously) say 'Mary is very pregnant'. How can fuzzy logicians account for this difference? Moreover, how can they account for the fact that one can say 'John is definitely very tall', but not 'John is very definitely tall'?