ELSEVIER

Information Sciences 138 (2001) 237–255

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

# A context-dependent method for ordering fuzzy numbers using probabilities

Ronald R. Yager [a,*], Marcin Detyniecki [b], Bernadette Bouchon-Meunier [b]

[a] *Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, USA*
[b] *LIP6 – CNRS – University of Paris VI, 4 place Jussieu, 75230 Paris Cedex 05, France*

Received 1 April 2000; received in revised form 4 January 2001; accepted 10 March 2001

**Abstract**

In this work we suggest an approach to comparing fuzzy numbers motivated by a probabilistic view of the underlying uncertainty. An important aspect of the method suggested is its context dependency, the numbers being compared affect the process. We discuss the effect of decision attitude and show that this approach is particularly useful for aiding decision makers having a temperate decision attitude, optimism in the face of adversity and conservatism in the face plenty. 

## 1. Introduction

In applications in which we use fuzzy sets we are often faced with the problem of comparing fuzzy numbers and deciding which one is bigger. For example in [1] Yager discusses the use of fuzzy subsets for the representation of perceived time in multi-media systems. Here we may be faced with the problem of ordering two events, with respect to which started first, where the start times of each of the events are expressed as fuzzy subsets. In decision making environments, in which we must select a course of action, the payoffs associated

* Corresponding author. Tel.: +1-212-249-2047; fax: +1-212-249-1689.
*E-mail addresses:* yager@panix.com (R.R. Yager), marcin.detyniecki@lip6.fr (M. Detyniecki), bernadette.bouchon-meunier@lip6.fr (B. Bouchon-Meunier).

0020-0255/01/$ - see front matter 
PII: S 0 0 2 0 - 0 2 5 5 ( 0 1 ) 0 0 1 3 2 - 3

with the different courses of action may be expressed in terms of fuzzy quantities. To accomplish this task requires that we compare fuzzy numbers. While it is clear how to compare scalar numbers, the comparison of fuzzy numbers is not always obvious as there exists a fundamental underlying uncertainty. From very early in the development of fuzzy set theory [2–5] the problem of comparing fuzzy subsets was apparent and was seen to be an important and complicated issue. A number of different techniques have been suggested in the literature, see [6,7] for a review of these methods. What seems to be apparent is that there exists no unique superior way to make comparisons but a number of different methods, each having some preferred qualities. At a more fundamental level there is a core subjectivity in this process, although some basic properties must be satisfied. Situations of this type, where one is trying to represent subjectivity and individual preferences, usually benefit from the introduction of various different methods and models, especially those in which a semantics exists connecting the formal model with human cognitive aspects. This semantic connection often more easily allows people to include subjective preferences in their selection of a method.

In this work we suggest an approach to comparing fuzzy numbers motivated by a probabilistic view of the underlying uncertainty. An important aspect of the method suggested is its context dependency. This context dependency is a manifestation of the fact that each of the fuzzy sets being ordered participates in the process by providing a measuring tool, a scale, on which all the fuzzy numbers being ordered are measured.

## 2. Comparing fuzzy numbers using a probability distribution

When expressing the value of a variable $V$, such as the beginning time of an event or the payoff associated with a decision, as a fuzzy subset we are inducing a possibility distribution [8] over the domain of the variable. In particular, the statement $V$ is $F$, where $F$ is a fuzzy set, can be seen as inducing a possibility distribution $\Pi$ on the domain of $V$ such that for each $x$, $F(x) = \Pi(x)$. From the initial introduction of the idea of possibility, starting with Zadeh's "possibility/probability consistency principle", [8] considerable interest has focused on the relationship between possibility and probability. A number of researchers have looked at the problem of associating probability distributions with possibility distributions and have suggested methods of conversion [9]. All of these share the fundamental properties that $F(x) = 0$ implies $P(x) = 0$, $F(x) = F(y)$ implies $P(x) = P(y)$ and $F(x) > F(y)$ implies $P(x) \geqslant P(y)$. In [10] Yager suggested a procedure for instantiating a possibilistic variable by converting its possibility distribution into a probability distribution. Once having this probability distribution the instantiation of the possibilistic variable can be implemented by the performance of a random experiment using the associated probability distribution. The conversion of the possibility distribution into a probability distribution suggested in [10] was via a simple normalization, i.e.

$$P(x) = \frac{F(x)}{\sum F(x)}.$$

As already noted a problem of longstanding interest in the field of fuzzy sets is the ordering of fuzzy numbers, fuzzy subsets over the real line. Motivated by this idea of converting possibility distributions into probability distributions we shall suggest a procedure for comparing the ordering of two fuzzy numbers based upon a calculation of the probability of the ordering of the associated probabilistic variables.

Let $F$ and $G$ be two fuzzy numbers. Let $f$ and $g$ be their associated probability distributions and let $X$ and $Y$ be two random variables having the probability distributions $f$ and $g$, respectively. Assume that we pick an element $x$ randomly for $X$ taking into account the distribution $f$ and an element $y$ for $Y$ taking into account the distribution $g$. Consider the probability that $x$ is greater than or equal to $y$. We will note this probability $P(Y \leqslant X)$. We observe that this probability should give us an idea of which of the fuzzy numbers $F$ or $G$ is greater. For example, if the fuzzy set $F$ is completely to the right of the fuzzy set $G$, $F > G$, it will be the case that the random variable $X$ is always larger than the variable $Y$ we would get $P(Y \leqslant X) = 1$. The interesting cases are when the two probability distributions are not disjoint.

In order to obtain $P(Y \leqslant X)$, we need some results from probability theory. We note that to calculate $P(Y \leqslant X)$, we can compute the equivalent probability $P(Z \geqslant 0)$ where $Z$ is the random variable $Z = X - Y$. The probability distribution of $Z$ is the convolution [11] of the probability distribution of $X$, $f(x)$, and the probability distribution of $Y$, $g(y)$. Let us call the probability distribution of $Z$, $h$, we then have

$$h(z) = \int_{-\infty}^{\infty} f(z+y)g(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} f(z-y)g(-y)\,\mathrm{d}y. \tag{1}$$

And so the probability $P(Y \leqslant X) = P(Z \geqslant 0)$ will be computed by

$$P(Z \geqslant 0) = \int_{0}^{\infty} h(z)\,\mathrm{d}z. \tag{2}$$

It is interesting to note here that there are a number of important mathematical properties associated with convolution:

Convolution is *commutative*

$$h = f \otimes g = g \otimes f,$$

Convolution is *associative*

$$h = f \otimes (g \otimes m) = (f \otimes g) \otimes m,$$

Convolution is *distributive*

$$h = f \otimes (g + m) = (f \otimes g) + (f \otimes m).$$

We also can write the formulas (1) and (2), for a discrete space, where $f$ and $g$ are the discrete probability distribution and therefore are normalized.

$$h[m] = \sum_{i=-\infty}^{+\infty} f[m+i]g[i]. \tag{3}$$

And so the probability $P(Y \leqslant X)$ will be computed by

$$P(X \leqslant Y) = \sum_{j=0}^{+\infty} h[j]. \tag{4}$$

In the preceding we have obtained a general formulation to compare the ordering of any two distributions, discrete or continuous.

## 3. Comparing uniform distributions

In order to understand the behavior of the complex mathematical expressions resulting from the convolution let us consider the case of the uniform probability distribution. Assume that we have two random variables, each one of them has a uniform probability distribution: $[x_1, x_2]$ for $X$ and $[y_1, y_2]$ for $Y$. See Fig. 1.

In this case Eq. (1) becomes

$$h(z) = \frac{1}{y_2 - y_1} \int_{-\infty}^{\infty} f(z+y)\,\mathrm{d}y. \tag{5}$$

Significantly different cases occur depending on the relative position of $x_1$ and $x_2$ with respect to $y_1$ and $y_2$. We studied all the particular cases and we summarize the result in 6 particular cases, divided into 3 cases of 2 members each: the disjoint cases, the overlapping cases and the included cases.

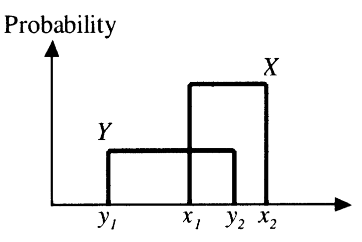


Fig. 1. Uniform probability distributions for $X$ and $Y$.

### 3.1. The disjoint cases

The two intervals $[x_1, x_2]$ and $[y_1, y_2]$ are disjoint, $x_2 \leqslant y_1$ or $x_1 \geqslant y_2$. In the first case, $x_2 \leqslant y_1$ (see Fig. 2), we have the whole probability interval for $X$ before the probability interval for $Y$. We are sure that any value $x$ for $X$ is smaller than any $y$ for $Y$. Mathematically the probability that any $x$ is bigger than any $y$ is zero: $P(Y \leqslant X) = 0$. We can make the same reasoning for the second case in order to obtain that we are sure that any $x$ will be bigger than any $y$. The probability that $X$ is bigger than $Y$ equals then one: $P(Y \leqslant X) = 1$. We have so for the disjoint cases:

if $x_2 \leqslant y_1$, then $P(Y \leqslant X) = 0$;
if $x_1 \geqslant y_2$, then $P(Y \leqslant X) = 1$.

We note here that in the disjoint cases the distance between the intervals is not taken into account. In fact as long as the intervals are disjoint the probability is constant equal to 1 or 0.

### 3.2. The overlapping cases

As in the preceding we have two cases. The first one, Fig. 3, is where the end of the $X$ interval overlaps with the beginning of the $Y$ interval, $x_1 \leqslant y_1 \leqslant x_2 \leqslant y_2$.

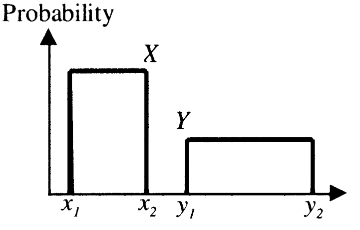


Fig. 2. Disjoint case $x_2 \leqslant y_1$.

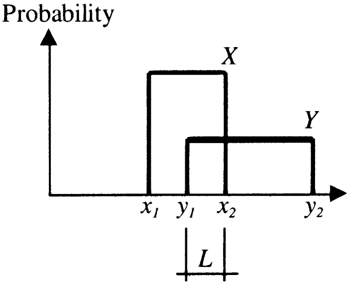


Fig. 3. Overlapping case $x_1 \leqslant y_1 \leqslant x_2 \leqslant y_2$.

In this case, $x_1 \leqslant y_1 \leqslant x_2 \leqslant y_2$, using formulas (2) and (5) we obtain that

$$P(Y \leqslant X) = \frac{1}{2}\frac{|L||L|}{|X||Y|},$$

where $|X| = x_2 - x_1$, the length of the $X$ interval, where $|Y| = y_2 - y_1$, the length of the $Y$ interval and where $|L| = x_2 - y_1$ is the interval that overlaps. An intuitive derivation of this result can also be made. We note that when $X$ falls in the interval $[x_1, y_1]$, which it does $(|X| - L)/|X|$ of the time, $Y$ is always bigger than $X$. When $X$ falls in the interval $[y_1, x_2]$ if $Y$ is greater than $x_2$, which it is $(|Y| - L)/|Y|$ of the time, $Y$ is always bigger than $X$. When both $X$ and $Y$ fall in the interval $[y_1, x_2]$, which occurs $(L/|Y|)(L/|X|)$ of the time then can $X$ be bigger than $Y$, however there is a fifty–fifty chance of this happening in this interval hence we get

$$P(Y \leqslant X) = \frac{1}{2}\frac{|L||L|}{|X||Y|}.$$

Significantly it is worth noting that since $L < |X|$ and $L < |Y|$ then $P(Y \leqslant X) < 0.5$. Also for a fixed $|X|$ and $|Y|$ as $L$ increases the value $P(Y \leqslant X)$ increases, we move closer to indistinguishability. We note that $|X|$ and $|Y|$ provide some measure of the uncertainty associated of the variables. With

$$L = x_2 - y_1 = \tfrac{1}{2}(x_2 + x_1) + \tfrac{1}{2}(x_2 - x_1) - \tfrac{1}{2}(y_2 + y_1) + \tfrac{1}{2}(y_2 - y_1),$$

we see that

$$L = \mu_x - \mu_y + \tfrac{1}{2}(|X| + |Y|),$$

where $\mu_x$ and $\mu_y$ are the respective means.

The second overlapping case (see Fig. 4) is when the beginning of the $X$ interval overlaps with the end of the $Y$ interval, $y_1 \leqslant x_1 \leqslant y_2 \leqslant x_2$. In this case we have

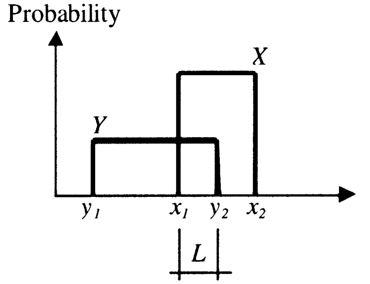


Fig. 4. Overlapping case $y_1 \leqslant x_1 \leqslant y_2 \leqslant x_2$.

$$P(Y \leqslant X) = 1 - \frac{1}{2}\frac{|L||L|}{|X||Y|},$$

where again $|X| = x_2 - x_1$, $|Y| = y_2 - y_1$ but where $|L| = y_2 - x_1$, the interval overlaps. Here since $L < |X|$ and $L < |Y|$ then $P(Y \leqslant X) > 0.5$.

Thus we see that generally in the overlapping case the ordering will be based upon positioning of the first element in the interval.

### *3.3. The inclusion cases*

Here we have the case when $X$ is inside of $Y$ and the symmetric case when $Y$ is inside of $X$. For the first case, when $X$ is inside of $Y$, we have $y_1 \leqslant x_1 \leqslant x_2 \leqslant y_2$ (see Fig. 5) Using formulas (8) and (2) we obtain that $P(Y \leqslant X) = |L_X|/|Y|$ where $|Y| = y_2 - y_1$ and where $|L_X| = ((x_2 + x_1)/2) - y_1$, the distance from the middle of $X$ to the left border of $Y$.

An interesting and useful reformulation of this can be obtained

$$P(Y \leqslant X) = \frac{1}{|Y|}\left(\frac{x_2 + x_1}{2} - y_1\right) = \frac{1}{|Y|}\left(\frac{x_2 + x_1}{2} - y_1 - \frac{1}{2}y_2 + \frac{1}{2}y_2\right),$$

$$P(Y \leqslant X) = \frac{1}{|Y|}\left(\frac{x_2 + x_1}{2} - \frac{y_2 + y_1}{2} + \frac{y_2 - y_1}{2}\right).$$

Noting that $\frac{1}{2}(x_2 + x_1)$ is the mean of $X$, $\mu_x$, and $\frac{1}{2}(y_2 + y_1)$ is the mean of $Y$, $\mu_y$, and that $|Y| = y_2 - y_1$ we see that

$$P(Y \leqslant X) = \frac{1}{2} + \frac{1}{|Y|}(\mu_x - \mu_y).$$

Since $P(X \leqslant Y) = 1 - P(Y \leqslant X)$ then we get

$$P(X \leqslant Y) = \frac{1}{2} + \frac{1}{|Y|}(\mu_y - \mu_x).$$

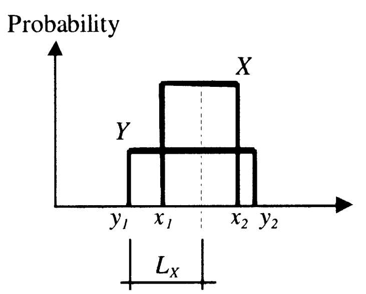


Fig. 5. Inclusion case $y_1 \leqslant x_1 \leqslant x_2 \leqslant y_2$.

In the second case, when $X$ is inside of $Y$, we have $x_1 \leqslant y_1 \leqslant y_2 \leqslant x_2$. For the case we have $P(Y \leqslant X) = |L_Y|/|X|$, where $|X| = x_2 - x_1$ and $|L_Y| = x_2 - [(y_2 + y_1)/2]$ is the distance from the middle of $Y$ to the left border of $X$. Again this can be reformulated so that

$$P(Y \leqslant X) = \frac{1}{2} + \frac{1}{|X|}(\mu_x - \mu_y)$$

and

$$P(X \leqslant Y) = \frac{1}{2} + \frac{1}{|X|}(\mu_y - \mu_x).$$

Noting that in both cases the dividing term is the spread of the broader distribution we unify this inclusion situation with one formula

$$P(Y \leqslant X) = \frac{1}{2} + \frac{1}{\mathrm{Max}(|X|, |Y|)}(\mu_x - \mu_y).$$

Here we see that $P(Y \leqslant X) > \frac{1}{2}$ if $\mu_x > \mu_y$. Thus we see this inclusion case induces an ordering based on the ordering of the respective means. In addition, we note that the term

$$\frac{1}{\mathrm{Max}(|X|, |Y|)}(\mu_x - \mu_y)$$

is inversely related to spreads $|Y|$ and $|X|$. As the uncertainty increases the clarity of the ordering diminishes.

## 4. Comparing Gaussian distributions

Another interesting probability distribution, where the comparison is relatively easy to compute is the Gaussian distribution. We recall that the general form of the Gaussian distribution is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\},$$

where $\mu$ is the mean and $\sigma$ is the standard deviation.

Let $f$ with parameters $(\mu_f, \sigma_f)$ be the Gaussian distribution associated with the random variable $X$ and let $g$ with parameters $(\mu_g, \sigma_g)$ be the Gaussian distribution associated with $Y$. It is known that a convolution of two Gaussian distribution is still a Gaussian convolution. Thus when we apply the formula (1), which is a convolution, we obtain a Gaussian distribution $h$ with the parameters:

$$\mu_h = \mu_f - \mu_g,$$
$$\sigma_h = \sqrt{\sigma_f^2 + \sigma_g^2}.$$

Using this the probability $P(Y \leqslant X)$ will be calculated by

$$P(Y \leqslant X) = \frac{1}{\sigma_h\sqrt{2\pi}} \int_0^{\infty} \exp\left\{ -\frac{(z-\mu_h)^2}{2\sigma_h^2} \right\} \mathrm{d}z.$$

This integral corresponds to the area for $z \geqslant 0$, see Fig. 6.

It is interesting to observe the role the standard deviation $\sigma$ plays in this framework. First we note that if $\mu_f = \mu_g$, then probability $P(Y \leqslant X) = \frac{1}{2}$. We can make no distinction between the two distributions. In Fig. 7, we see three different Gaussian distributions all centered on the same $\mu$ but with different $\sigma$'s. No distinction is made between any of these.

Consider now the case where $\mu_f > \mu_g$. In this situation it will always be the case that $P(X \geqslant Y) > \frac{1}{2}$, however the exact value of $P(X \geqslant Y)$ will depend upon the associated variances. For example if $\Delta\mu = \mu_f - \mu_g > 0$ and $\sigma_h = \sqrt{\sigma_f^2 + \sigma_g^2}$ then

$$P(X \geqslant Y) = \int_{-\Delta\mu/\sigma_h}^{\infty} \mathrm{e}^{-0.5z} \, \mathrm{d}z.$$

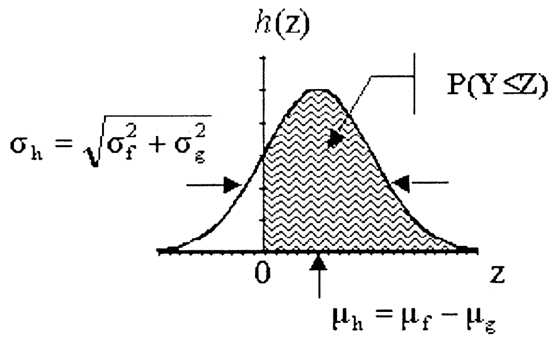


Fig. 6. $P(Y \leqslant X)$ for Gaussian distribution $h(z)$.

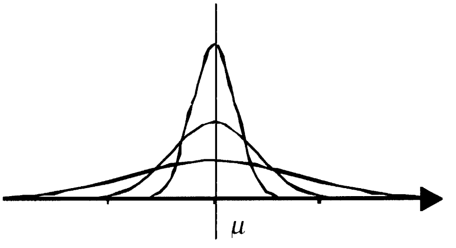


Fig. 7. Three Gaussian distributions with the same mean $\mu$.

We see that for a fixed $\Delta\mu$ the smaller the variance the larger the $P(X \geqslant Y)$. Thus if $X_1$ and $X_2$ are two fuzzy numbers with the same mean, greater than $Y$'s, but different variances, $X_1$ having a smaller variance then we have that $P(X_1 \geqslant Y) > P(X_2 \geqslant Y) > \frac{1}{2}$. The implication here is that the smaller variance, the lesser the uncertainty, will allow us the make stronger distinctions. A further fact worth noting in this regard is that if $\sigma_h$ gets very large, then $\frac{\Delta\mu}{\sigma_h} \to 0$ and hence $P(X \geqslant Y) \to \frac{1}{2}$. This of course means that if either of two variables being compared has a large variance, large uncertainty, we will get $P(X \geqslant Y) \to \frac{1}{2}$. It is to be noted that if $\Delta\mu$ gets very small, then also $\frac{\Delta\mu}{\sigma_h} \to 0$ and hence $P(X \geqslant Y) \to \frac{1}{2}$. What is also important to emphasize is that no distinction can be directly made between $X_1$ and $X_2$ if they have the same mean as here, $\Delta\mu = 0$. As we shall subsequently see indirectly a distinction exists through their relationship with other fuzzy numbers.

## 5. Comparing multiple fuzzy numbers

In the preceding we have suggested an approach to the comparison of two fuzzy numbers based upon the use of a probabilistic interpretation. Specifically, if $X$ and $Y$ are two fuzzy numbers we have calculated $\text{Prob}(X \leqslant Y)$ and $\text{Prob}(X \geqslant Y)$. Our goal in obtaining these values was to provide an ordering on the fuzzy numbers. For example, in the framework of decision making if one course of action leads to a payoff expressed as the fuzzy number $X$ and if a second leads a payoff expressed as the fuzzy number $Y$ then if $P(X \geqslant Y) \geqslant P(Y \geqslant X)$ we would select the first course of action.

In the following we shall consider the use of this tool for the ordering of multiple fuzzy numbers. To focus on the main ideas we shall restrict ourselves to the case of continuous fuzzy numbers. A useful property in this case, since $P(X = Y) = 0$, is that $P(X \geqslant Y) = 1 - P(X \geqslant Y)$. We also note that in this case of continuous fuzzy number if $X$ and $Y$ have the same distribution, then $P(X \geqslant Y) = P(Y \geqslant X) = 0.5$.

Let $X_1, X_2, \ldots, X_n$ be a collection of continuous fuzzy numbers. To induce an ordering over these fuzzy numbers we shall introduce a fuzzy relationship $S$ [12] over the space $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ in which we shall refer to $S(X_i, X_j)$ as the *strength of preference* of $X_i$ over $X_j$. To obtain the strengths of preferences we shall use the method discussed in the preceding. Specifically for any pair $X_i$ and $X_j$ we calculate $P(X_i \geqslant X_j)$ and then assign $S(X_i, X_j) = \text{Prob}(X_i \geqslant X_j)$. We note in this fuzzy relationship that $S(X_i, X_j) = 1 - S(X_j, X_i)$ and $S(X_i, X_i) = 0.5$.

To generate the ordering induced by the relationship $S$ we associate with each $X_i$ a value, $T(X_i)$, called its score, defined as

$$T(X_i) = \sum_{j=1}^{n} S(X_i, X_j).$$

We then use these scores to provide an ordering over the $X_i$ such that $X_i >_S X_j$ if $T(X_i) > T(X_j)$ and $X_i =_S X_j$ if $T(X_i) = T(X_j)$. There exist alternative ways of calculating the score function $T$ from the fuzzy relationship $S$, such as

$$T(X_i) = \sum_{j=1}^{n} (1 - S(X_j, X_i)).$$

Let us now look at this suggested procedure for ordering fuzzy numbers. Consider first the case when we are just ordering two fuzzy numbers. In this case using this method we get the property that if $P(X_1 \geqslant X_2) > P(X_2 \geqslant X_1)$, then $X_1 >_S X_2$. We see this as follows. Since $P(X_1 \geqslant X_2) + P(X_2 \geqslant X_1) = 1$ if $P(X_1 \geqslant X_2) > P(X_2 \geqslant X_1)$, then $P(X_1 \geqslant X_2) = S(X_1, X_2) = \alpha > \frac{1}{2}$. Since $S(X_i, X_i) = \frac{1}{2}$ then $S$ is represented by the following matrix:

$$S = \begin{matrix} & X_1 & X_2 \\ X_1 \\ X_2 \end{matrix} \begin{pmatrix} 0.5 & \alpha \\ 1 - \alpha & 0.5 \end{pmatrix}.$$

Using this we get $T(X_1) = \alpha + \frac{1}{2}$ and $T(X_2) = (1 - \alpha) + \frac{1}{2}$ since $\alpha > \frac{1}{2}$ then $T(X_1) >_S T(X_2)$

A general property that we can associate with this ordering procedure is the following:

If $X_i$ strongly dominates $X_j$, $S(X_i, X_k) \geqslant S(X_j, X_k)$ for all $k$ and $S(X_i, X_k) > S(X_j, X_k)$ for at least one $K$, then $T(X_1) > T(X_2)$ and hence $X_i >_S X_j$.

This procedure leads to the natural ordering in the case in which the fuzzy numbers are crisp scalar numbers, $X_j = x_j$. In the case of comparing crisp numbers we get $P(X_i \geqslant X_j) \in \{0, \frac{1}{2}, 1\}$ where $P(X_i \geqslant X_j) = 1$ if $x_i > x_j$, $P(X_i \geqslant X_j) = 0$ if $x_i < x_j$ and $P(X_i \geqslant X_j) = \frac{1}{2}$ if $x_i = x_j$. In this case it can be shown that the methodology suggested has $T(X_i) > T(X_j)$ if $x_i > x_j$ and $T(X_i) = T(X_j)$ if $x_i = x_j$.

It is significant to note, as we shall subsequently see, in using this approach each of the fuzzy subsets being ordered is essentially providing a scale on which we measure all the other fuzzy subsets, that is $P(X_j \geqslant X_k)$ for $k = 1, 2, \ldots, n$ can be seen as a collection of measurements on a scale provided by $X_j$ via the instrument $P(X_j \geqslant X_k)$. As such the ordering that will result using this method will be context dependent, since the procedure, the measuring instruments being used, depend on the objects being included in the process.

While the approach suggested allows us to compare any types of fuzzy numbers it is useful to consider the special case when all the numbers are Gaussian as this will more easily allow us to get some intuition about the method being used.

We can make some general observations regarding the ordering of a collection of Gaussian numbers, $X_i = (\mu_i, \sigma_i)$. First we note that if $\sigma_i = 0$ for all $i$, then $P(X_i \geqslant X_j) = 1$ if $\mu_i > \mu_j$. In this case, the ordering obtained by this method is based simply on the ordering of the $\mu_i$, thus $X_i >_S X_j$ if $\mu_i > \mu_j$ and $X_i =_S X_j$ if $\mu_i = \mu_j$. At the other extreme is the situation when all the $\sigma$'s get very large, $\sigma_i \to \infty$ for all $i$, in this case $S(X_i, X_j) \to 0.5$ for all pairs and the distinction between the objects decreases, $T(X_i) \approx T(X_j)$ for all pairs. In general small variance, strong certainty, tends to persevere the ordering imposed by the mean.

Let us look further at the properties associated with this approach. Without loss of generality we shall assume that the $X_i$ are indexed such that $\mu_i \geqslant \mu_j$ if $i < j$, $X_1$ has the largest mean. Consider now the special case in which we assume $X_1$ has the smallest variance, $X_1 = (\mu_1, \sigma_1)$ has the largest mean and the smallest variance. Assume $X_j$ is some other distribution, let us compare $S(X_1, X_k)$ and $S(X_j, X_k)$ for any $k$. Denoting $\Delta_{1k} = \mu_1 - \mu_k$ and $\Delta_{jk} = \mu_j - \mu_k$ it is clear since $\mu_1$ is the largest mean that $\Delta_{1k} > \Delta_{jk}$ for all $k$. Furthermore denoting

$$\sigma_{1k} = \sqrt{\sigma_1^2 + \sigma_k^2} \quad \text{and} \quad \sigma_{jk} = \sqrt{\sigma_j^2 + \sigma_k^2}$$

its clear that $\sigma_{1k} < \sigma_{jk}$. Since

$$S(X_1, X_k) = \int_{\Delta_{1k}/\sigma_{1k}}^{\infty} \mathrm{e}^{-0.5z}\,\mathrm{d}z \quad \text{and} \quad S(X_j, X_k) = \int_{\Delta_{jk}/\sigma_{jk}}^{\infty} \mathrm{e}^{-0.5z}\,\mathrm{d}z,$$

and $-\Delta_{1k}/\sigma_{1k} < \Delta_{jk}/\sigma_{jk}$ then $S(X_1, X_k) > S(X_j, X_i)$ for $k$ and hence $T(X_1) > T(X_j)$. Thus we see that if the distribution with the largest mean also has the smallest variance, the least uncertainty, it will be the highest ordered, the most preferred. The intuition behind this that a fuzzy number with a large mean and small variance is such that its instantiations will be a larger value than most. We should note that largest mean by itself does not guarantee first in the ordering, an associated large variance can diminish its standing. At the other extreme if $X_n$, the distribution with the smallest mean, also has the smallest variance then $T(X_j) > T(X_n)$ for all $j$, thus a distribution with the smallest mean and least uncertainty, smallest variance, will always be the lowest in the ordering.

As we have indicated that narrow variance, strong certainty, tends to lead to orderings preserving the mean order while more uncertainty, large variance, can lead to reversal of the mean ordering. In the following we first look at the role of uncertainty and the effect of its increase. Let $X$ be some Gaussian fuzzy number with parameters $(\mu, \sigma)$. Let $X_j = (\mu_j, \sigma_j)$ be any arbitrary fuzzy number for which $\mu > \mu_j$. It is clear in this case that $S(X, X_j) > 0.5$. In particular with

$$\Delta = \mu - \mu_j \quad \text{and} \quad \sigma^* = \sqrt{\sigma^2 + \sigma_j^2}$$

we have

$$S(X,X_j) = \int_{\Delta/\sigma^*}^{\infty} \mathrm{e}^{-0.5z}\,\mathrm{d}z.$$

We also see that if the uncertainty associated with $X$ increases to $\sigma' > \sigma$, then $\sigma^*$ increases to $\sigma^{**}$ and $\Delta/\sigma^*$ decreases to $\Delta/\sigma^{**}$, causing $S(X,X_j)$ to decrease, bringing it closer to $\frac{1}{2}$. Thus we see an increase in variance results in a loss of distinguishability of $X$ from the other elements.

On the other hand if $\mu < \mu_j$, $\Delta < 0$, then $S(X,X_j) < \frac{1}{2}$. In this case an increase in $\sigma$ also causes an increase in $\sigma^*$ to $\sigma^{**}$ and a decrease in $|\Delta/\sigma^*|$ to $|\Delta/\sigma^{**}|$. However since $\Delta/\sigma^* < 0$ and $\Delta/\sigma^{**} < 0$ then $-\Delta/\sigma^* > 0$ and $\Delta/\sigma^{**} > 0$. This results in an increase in $S(X,X_j)$, again resulting in less distinguishability.

As the following example illustrates the association of a large variance, great uncertainty, with a fuzzy number can lead to an ordering that reverses the mean ordering.

**Example.** Assume that $X_1, X_2$, and $X_3$ are three fuzzy numbers with parameters $(\mu_1, \sigma_1)$, $(\mu_2, \sigma_2)$ and $(\mu_3, \sigma_3)$ such that $\mu_1 > \mu_2 > \mu_3$ and where $\sigma_2 = \sigma_3 \approx$ "very small" and $\sigma_3$ is large (see Fig. 8).

In this case

$$S = \begin{bmatrix} 1/2 & a & b \\ \overline{a} & 1/2 & c \\ \overline{b} & \overline{c} & 1/2 \end{bmatrix},$$

where $a = S(X_1,X_2) > \frac{1}{2}$, $b = S(X_1,X_3) > \frac{1}{2}$ and $c = S(X_2,X_3) > \frac{1}{2}$. Let

$$\sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2} \quad \text{and} \quad \Delta_{ij} = \mu_i - \mu_j.$$

Since $\sigma_3 = \sigma_2 \approx 0$ then $\sigma_{23} \approx 0$ and $c = S(X_2,X_3) \approx 1$, $\overline{c} \approx 0$. Furthermore

$$a = \int_{-q}^{\infty} \mathrm{e}^{-0.5z}\,\mathrm{d}z \quad \text{and} \quad b = \int_{-r}^{\infty} \mathrm{e}^{-0.5z}\,\mathrm{d}z,$$

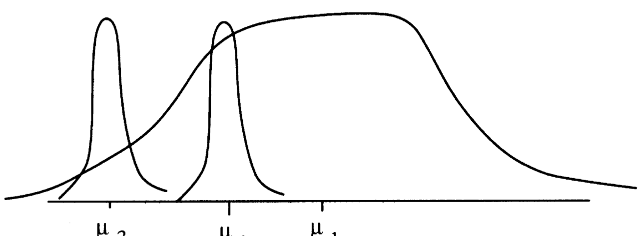


Fig. 8. Three Gaussian fuzzy numbers.

where

$$q = \frac{\Delta_{12}}{\sigma_{12}} \approx \frac{\Delta_{12}}{\sigma_1} \quad \text{and} \quad r = \frac{\Delta_{13}}{\sigma_{13}} \approx \frac{\Delta_{13}}{\sigma_1}.$$

If $\sigma_1$ is very big, then $q \to 0$ and $r \to 0$, hence $b \approx a \approx \frac{1}{2}$. This gives us

$$S = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1 \\ 1/2 & 0 & 1/2 \end{bmatrix}.$$

Thus $T(X_1) = 1.5$, $T(X_2) = 2$ and $T(X_3) = 1$ and hence $X_2 >_S X_1 >_S X_3$. Thus we see that large uncertainty can cause inversion of the ordering with respect to the mean.

## 6. Underlying decision attitude

To help provide further insight into this method we shall consider first the case of two Gaussian numbers $X_1$ and $X_2$ with parameters $(\mu_1, \sigma_1)$ and $(\mu_2, \sigma_2)$. If $\mu_1 > \mu_2$, then our ordering method results in $X_1 >_S X_2$. If $\mu_1 = \mu_2 = \mu$, then our method results in $X_1 =_S X_2$, and they are indistinguishable. As noted this indistinguishability is unaffected by any information about the variances, their relative uncertainties. Whether $\sigma_1 < \sigma_2$, or $\sigma_1 < \sigma_2$ or $\sigma_1 = \sigma_2$ we get $X_1 =_S X_2$.

In what follows, without loss of generality we shall assume that $\sigma_1 < \sigma_2$, $X_2$ has more uncertainty than $X_1$. As we have noted this does not allow us to distinguish between these two numbers when $\mu_1 = \mu_2 = \mu$. Consider now the introduction of a third fuzzy number $X_3$ with parameters $(\mu_3, \sigma_3)$ where $\mu_3 < \mu$. In this case to provide an ordering among these three Gaussian numbers we need to calculate $S(X_i, X_j)$ for all combinations. We already know that $S(X_i, X_i) = \frac{1}{2}$ for all $i$. Furthermore since these are Gaussian and $\mu_1 = \mu_2$ then $S(X_1, X_2) = S(X_2, X_1) = \frac{1}{2}$. In addition since both $\mu_1$ and $\mu_2$ are greater than $\mu_3$ we have $S(X_1, X_3) = \alpha_1 > \frac{1}{2}$ and $S(X_2, X_3) = \alpha_2 > \frac{1}{2}$. However since $\sigma_1 < \sigma_2$, as we earlier noted this implies that $\alpha_1 > \alpha_2$. The following matrix summarizes our measurement

$$S = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{c} \begin{array}{ccc} X_1 & X_2 & X_3 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & \alpha_1 \\ 1/2 & 1/2 & \alpha_2 \\ 1-\alpha_1 & 1-\alpha_2 & 1/2 \end{pmatrix} \end{array},$$

where $\alpha_1 > \alpha_2 > 0.5$. Using this matrix we get

$$T(X_1) = 1 + \alpha_1,$$
$$T(X_2) = 1 + \alpha_2,$$
$$T(X_3) = 1 + \alpha_1 + 1 - \alpha_2 + \tfrac{1}{2}.$$

Since $\alpha_1 > \alpha_2$ then $T(X_1) > T(X_2)$. Since $1 - \alpha_1$ and $1 - \alpha_2$ are both less than $\frac{1}{2}$ we get that $T(X_3) < 1.5 < T(X_2)$. Thus in this case we get the following ordering:

$$X_1 >_S X_2 >_S X_3.$$

It is significant here to note that the inclusion of the third element allowed us to make a distinction between $X_1$ and $X_2$ which we could not make between them alone. The reason for this is that the introduction of $X_3$ brought with it an additional measuring scale, the values $S(X_j, X_3)$, which allowed us to make this distinction. In particular, the distinction was based upon the different variances associated with $X_1$ and $X_2$, a factor which does not manifest itself when we only had two distributions having the same mean.

Consider now again the situation in which we have again $X_1 = (\mu_1, \sigma_1)$ and $(\mu_2, \sigma_2)$ where $\mu = \mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$. However consider the introduction of $X_3$ where $X_3 = (\mu_3, \sigma_3)$ but $\mu_3 > \mu$, $X_3$ is to the right of $X_1$ and $X_2$. Let us denote $\text{Prob}(X_3 \geqslant X_1) = \beta_1$ and $\text{Prob}(X_3 \geqslant X_1) = \beta_2$. Since $\mu_3 > \mu_1 = \mu_2$ then both $\beta_1$ and $\beta_2 > \frac{1}{2}$. However again since $\sigma_1 > \sigma_2$ then $\beta_1 > \beta_2$, thus $\frac{1}{2} < \beta_2 < \beta_1$.

In this case our $S$ matrix is

$$S = \begin{matrix} X_1 \\ X_2 \\ X_3 \end{matrix} \begin{pmatrix} 1/2 & 1/2 & 1 - \beta_1 \\ 1/2 & 1/2 & 1 - \beta_2 \\ \beta_1 & \beta_2 & 1/2 \end{pmatrix}$$

from this we get:

$$T(X_1) = 1 + 1 - \beta_1 = 2 - \beta_1,$$
$$T(X_2) = 2 - \beta_2,$$
$$T(X_3) = \beta_1 + \beta_2 + \frac{1}{2}.$$

It is clear that since $\beta_1$ and $\beta_2 > \frac{1}{2}$ then $T(X_3) > T(X_2)$ and $T(X_3) > T(X_1)$. However since $\beta_1 > \beta_2$ then $1 - \beta_1 < 1 - \beta_2$ and therefore $T(X_1) < T(X_2)$ thus here we get as our ordering

$$X_3 >_S X_2 >_S X_1.$$

The fact that $X_3$ is the biggest is as expected. However, the inversion of the ordering between $X_2$ and $X_1$ in these two situations is a notable phenomenon. Previously we had $X_1 >_S X_2 >_S X_3$ and now we get $X_3 >_S X_2 >_S X_1$. In the first case $X_1$ was preferred to $X_2$ and in this case $X_2$ is preferred to $X_1$. Let us try to understand the reason for this. First of all this inversion is a manifestation of the fact that each of the distributions which are being ordered brings with it a scale on which all other distributions are measured. Since the distribution $X_3$ used in the two cases was different the two situations are not using the same measuring instruments. Thus this situation dramatically highlights the contextual aspect of this approach to ordering. The result is strongly dependent upon the objects being ordered, not only in their sense of them being the objects to be ordered but in the additional sense that the objects being ordered participate in the process by providing the tools, measuring instruments, that are used to compare the objects being ordered. Once realizing the inherent contextual aspect of this procedure it becomes important to try to understand how this works. Let us turn to this question.

It is first worth noting the effect on the ordering process of the variances, the uncertainty, associated with tied numbers. In particular, when $X_1$ and $X_2$ were bigger than $X_3$, $\mu_1 = \mu_2 > \mu_3$, then we favored the more certain of the distributions, $X_1 >_S X_2$. On the other hand when $\mu_1 = \mu_2 < \mu_3$ we obtained $X_2 >_S X_1$, we favored the more uncertain outcome.

In order to try to understand what has occurred and to provide some further intuition for the mechanism being used here to order the fuzzy numbers let us again return to the situation in which we have two fuzzy numbers with the same mean but different variances, different uncertainties. In this situation there is some feeling that the information about the different variances, the relative certainties of the distributions should be useful to help distinguish between these two fuzzy numbers. One possibility is that we may want to give some preference to the fuzzy number with the smaller variance. This attitude may not always correctly capture the situation. Consider a case in which the two fuzzy numbers are payoffs associated with alternative actions in a decision. Our ordering then is to be used to select one of these actions, and we prefer the action with the bigger payoff. In this framework the preference for certainty, the selection of the number with the smaller variance, $X_1 > X_2$, can be seen as an implementation of a pessimistic decision attitude. Here the decision maker can be seen as trying to avoid any uncertainty because he feels that this will work in his worst interest. He has a malevolent view of nature [13].

At the other extreme is the decision maker who would prefer the fuzzy number with the larger variance, $X_2 > X_1$. This decision maker, who prefers the more uncertainty, can be seen as one implementing an optimistic attitude. Here the decision maker is seeking uncertainty as it offers the possibility for greater gain. This type of decision maker views nature as a benevolent force.

As we shall suggest in the following the decision attitude manifested in the ordering procedure we have introduced will be a more complex one than either of these two extreme cases. Let us now consider a decision maker whose attitude is one that we shall call *temperate*. This temperate decision maker looks at the situation in the following way. He feels that if he selects an alternative with a good expected payoff ("big" fuzzy number), then he prefers it to be very certain, have little variability and uncertainty, as this guarantees him a good payoff. On the other hand if he must select an alternative that is associated with a poor expected payoff ("small" fuzzy number), then he prefers large variability (uncertainty) as this may give him an opportunity to actually receive a high payoff. We see that this temperate decision maker "is conservative in the face of good" and "optimistic in the face of adversity". As we shall subsequently see it is this temperate decision attitude with respect to the resolution of uncertainty that guides our suggested approach to ordering fuzzy numbers.

Consider now this temperate decision maker in our situation, two fuzzy numbers $(\mu, \sigma_1)$ and $(\mu, \sigma_2)$ with $\sigma_1 > \sigma_2$. In order to use his rule for distinguishing payoffs based upon their variance, he must have some indication as to whether he is in a high/low situation. He must have some indication of whether $\mu$ is a good or bad payoff. Unfortunately all he has is $\mu$ and he has no additional knowledge to indicate whether $\mu$ is good or bad (high or low), he only has one point, $\mu$. In this situation this type of temperate decision maker cannot decide whether he should prefer certainty or uncertainty, hence he cannot use the distinction in variance between $X_1$ and $X_2$ to distinguish between them.

Consider now the introduction of a third fuzzy number, $X_3$, where $\mu_3 < \mu$. Our decision maker now has an additional piece of information to help locate $\mu$. This information, $u_3 < u$, provides some indication that $\mu$ is "not bad"; it is good, he could do worse. Given this information and his temperate decision attitude he now becomes slightly pessimistic, preferring certainty to uncertainty and hence $X_1 >_S X_2$.

On the other hand the introduction of a third fuzzy number $X_3$, where $\mu_3 < \mu$, has the opposite effect. Here again he has an additional piece of information which can help locate the $\mu$ in context. However, this information, $\mu_3 < \mu$, provides some indication that this $\mu$ is not a good solution, it is low, he could do better. Given this information and his temperate attitude he now becomes optimistic, preferring uncertainty to certainty and hence $X_2 >_S X_1$.

We can view this temperate attitude from a slightly different perspective. Consider that in effect when ordering the fuzzy numbers each fuzzy number as being disambiguated to some crisp number. This process being a function of $\mu_i$, $\sigma_i$ and some $\alpha_i$, measure of attitude. If a number looks around sees a lot of other numbers as being bigger or better, then its attitude is that things can get better and hence it is disambiguated in an optimistic fashion, uncertainty is preferred. On the other hand if it sees that most of its colleagues are smaller than it, it becomes worried that it could end up small and hence it prefers certainty. It becomes pessimistic. Thus we see that the decision attitude used with each fuzzy number is different, it depends on the other objects being defuzzified, it is relative and context dependent. A fuzzy number that is in a group with a lot of fuzzy numbers smaller than itself will tend to act conservative, prefer certainty while when in the midst of a collection of fuzzy numbers bigger than itself it will prefer uncertainty.

Our temperate decision maker possesses a type of decision attitude that is conservative (pessimistic) in the face of plenty, and aggressive (optimistic) in the face of adiversity. It is this type of decision attitude that is implicitly being used within the approach to ordering fuzzy numbers that we have suggested.

A dual of this temperate decision maker is one who would prefer variability, be optimistic, in the face of plenty, ‘‘get while the getting is good’’, and would be pessimistic, prefer certainty, in the face of adversity, ‘‘cut his losses’’.

In the preceding we have tried to provide some understanding of the procedure we introduced for ordering fuzzy numbers. It is important to note that implicit decision attitude, temperance, is one that leads to context dependence, the ordering can be substantially affected by the objects being ordered. As we have indicated the reason for this is that the decision maker takes into account the other numbers to be ordered when doing the ordering.

## 7. Conclusion

In this work we introduced an approach to comparing fuzzy numbers motivated by a probabilistic view of the underlying uncertainty. An important aspect of the method suggested was its context dependency, the numbers being compared affect the process. This context dependency was shown to be a manifestation of the fact that each of the fuzzy sets being ordered participates in the process by providing a measuring tool, a scale, on which all the fuzzy numbers being ordered are measured. We discussed the effect of decision attitude. We showed that this approach is particularly useful for aiding decision makers having a temperate decision attitude, optimism in the face of adversity and conservatism in the face plenty.

## References


[1] R.R. Yager, Fuzzy temporal methods for video multimedia information systems, Journal of Advanced Computational Intelligence 1 (1997) 37–45.

[2] S.M. Baas, H. Kwakernaak, Rating and ranking multiple-aspect alternatives using fuzzy sets, Automatica 13 (1977) 47–58.

[3] R.R. Yager, Ranking fuzzy subsets over the unit interval, in: Proceedings of the Seventeenth IEEE Conference on Cybernetics and Society, 1978, pp. 921–925.

[4] J.F. Baldwin, N.C.F. Guild, Comparison of fuzzy sets on the same space, Fuzzy Sets and Systems 2 (1979) 212–231.

[5] D. Dubois, H. Prade, Ranking fuzzy numbers in the setting of possibility theory, Information Sciences 30 (1983) 183–224.

[6] G. Bortolan, R. Degani, A review of some methods for ranking fuzzy subsets, Fuzzy Sets and Systems 15 (1985) 1–19.

[7] G. Facchinetti, R.G. Ricci, S. Muzzioli, Note on ranking fuzzy triangular numbers, International Journal of Intelligent Systems 13 (1998) 613–622.

[8] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.

[9] G.J. Klir, B. Yuan, Fuzzy Sets, Fuzzy Logic and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh, World Scientific, Singapore, 1996.
[10] R.R. Yager, On the instantiation of possibility distributions, Fuzzy Sets and Systems (to appear).
[11] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.
[12] L.A. Zadeh, Similarity relations and fuzzy orderings, Information Sciences 3 (1971) 177–200.
[13] R.R. Yager, A game theoretic approach to decision making under uncertainty, International Journal of Intelligent Systems in Accounting, Finance and Management, 1999, pp. 131–143.