4.3 Ranking Using Hamming Distance

This family of ranking methods involves the measurement of the Hamming distance between two fuzzy sets. The Hamming distance between fuzzy sets M and N is defined as:

$$d(M,N) = \int_{-\infty}^{+\infty} | \mu_M(x) - \mu_N(x) | \, dx \qquad (4.23)$$

for continuous functions, and

$$d(M,N) = \sum_{i=0}^{K} |\mu_M(x_i) - \mu_N(x_i)| \qquad (4.24)$$

for discrete functions. If M and N are continuous functions as shown in Fig. 4.10, then the shaded areas represent the Hamming distance between M and N.

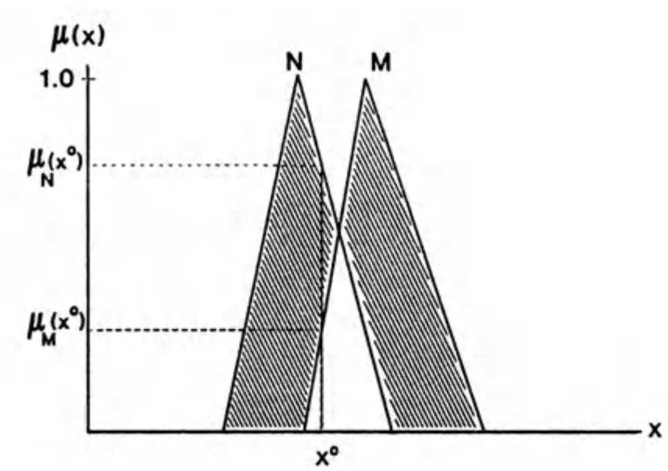


Fig. 4.10 The Hamming distance between fuzzy sets M and N.

Yager [Y11] defines a fuzzy maximum first and then computes the Hamming distance between each fuzzy set and the fuzzy maximum. The fuzzy set with the smallest distance from the fuzzy max is considered the best. Kerre [K17] follows Yager's ranking concept -- except that he defines a fuzzy max differently from Yager. Nakamura [N2] derives fuzzy minimums from the greatest lower set and the greatest upper set, respectively. The Hamming distances for a fuzzy min from the greatest upper set and a fuzzy min from the greatest lower set are computed. A fuzzy order relation, p, between fuzzy sets M and N is then constructed. The membership value $\mu_p(M,N)$ denotes the weighted sum of M's unique superiority over N for the best and worst possible situations. Kolodziejczyk [K27] extends Orlovsky's fuzzy preference relation [O2] and the notion of Hamming distance to construct several preference indices, P_1, P_2, and P_3. These indices follow very closely with Kerre's index and Nakamura's preference relation.

Notice that in Fig. 4.10 the Hamming distance is the summation of the areas not commonly occupied by both M and N. The methods in this category compare fuzzy sets with the fuzzy max (the fuzzy min) using the Hamming distance and the fuzzy set that has shortest (longest) distance to the fuzzy max (the fuzzy min) is considered better. The ideas of comparison are good. However, since the comparison is based solely on area measurement and the fuzzy set's relative location on the x-axis is ignored, the logic of the methods in this category is not sound.

For example, let us consider fuzzy sets M_1, M_2, and M_3 in Fig. 4.11a. Keen observation and common sense indicate $M_3 > M_2 > M_1$. Yager's method, which defines a unique fuzzy max for all comparison cases, would result in: $d(max,M_3) = 0.5$, $d(max,M_2) = .452$, and $d(max,M_1) = .436$. The ranking order is $M_1 > M_2 > M_3$. This ranking order is against human intuition. The problems with Yager's method are that crisp numbers will always be ranked the lowest, and a fuzzy set that overlaps more with Yager's fuzzy max will get a higher rank, regardless of their relative locations on the x-axis.

Kerre's [K17] method defines a fuzzy max which is problem-dependent. The fuzzy max in Fig 4.11a is M_3 (by Kerre's definition).

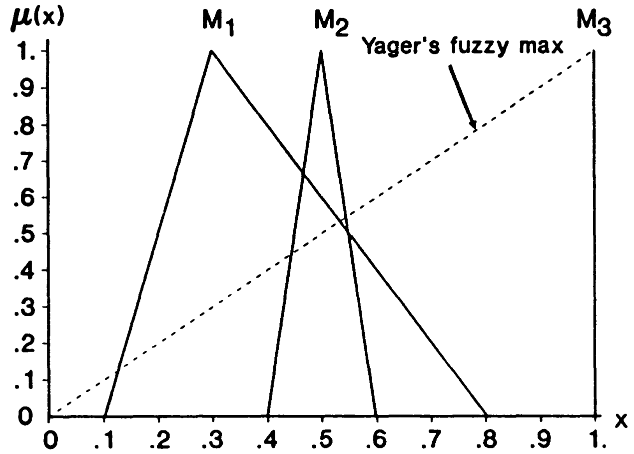


Fig. 4.11a A demonstration of weakness of Yager's method.

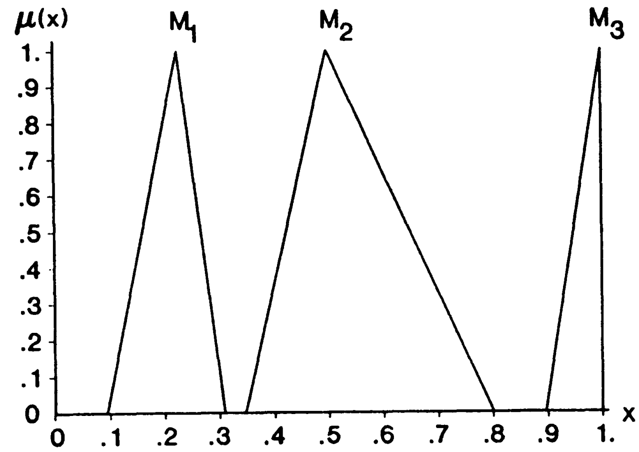


Fig. 4.11b A demonstration of weakness of Kerre's method.

The Hamming distances are: d(max,M_3) = 0; d(max,M_2) = area occupied by M_2; and d(max,M_1) = area occupied by M_1. The ranking order is $M_3 > M_2 > M_1$, which complies with human intuition. Kerre's method for this example looks better than Yager's approach. However, let us consider another set of fuzzy numbers as shown in Fig. 4.11b. Human intuition would favor M_2 over M_1. Kerre's method gives the result $M_1 > M_2$. This result is against our intuition. This counter-intuition case shows that Kerre's method would favor a fuzzy set with smaller area measurement, regardless of its relative location on the x-axis.

The illogicality observed in Yager's and Kerre's methods, i.e., ignoring fuzzy sets' relative location on x-axis, can be addressed to Nakamura's and Kolodziejczyk's methods as well.

4.3.1 Yager's Approach

Yager [Y11] proposed a ranking procedure for fuzzy sets of unit interval. First of all, a fuzzy maximum of the fuzzy sets to be ranked is determined. Then, each fuzzy set is compared with the fuzzy maximum using the Hamming distance measurement. The fuzzy set(s) that have the smallest Hamming distance to the fuzzy max are ranked as the best.

Yager's fuzzy max is defined as $\{(x, \mu_{max}(x))\}$ with membership function

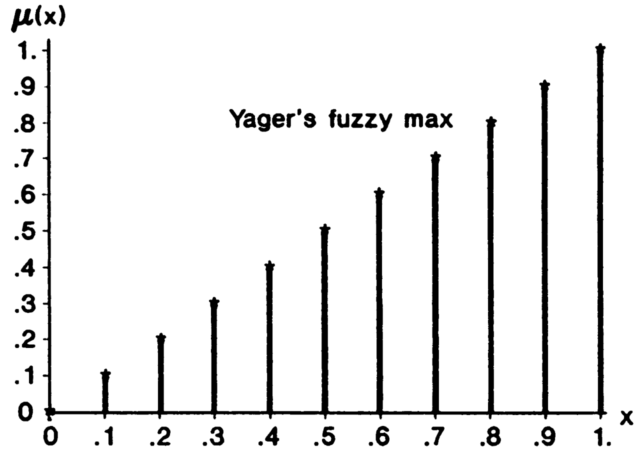
$$\mu_{max}(x) = x, \tag{4.25}$$

where x and $\mu_{max}(x)$ are in [0,1]. Fig. 4.12a illustrates Yager's fuzzy max for discrete functions, and Fig. 4.12b shows Yager's fuzzy max for continuous functions.

Once the fuzzy max is defined, we shall use the Hamming distance to measure the closeness of each fuzzy set to the fuzzy max. For example, given two fuzzy sets M and N as shown in Fig. 4.13, the Hamming distance between the fuzzy max and M is calculated as:

$$\begin{aligned} d(max,M) &= |0 - 0| + |.1 - 0| + |.2 - 0| + |.3 - .3| + |.4 - .8| \\ &+ |.5 - 1| + |.6 - .8| + |.7 - .3| + |.8 - 0| \\ &+ |.9 - 0| + |1 - 0| = 4.5. \end{aligned}$$

Similarly, the Hamming distance between the fuzzy max and N is d(max,N) = 6.0. Since d(max,M) is smaller than d(max,N), M > N.



x	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
$\mu_{max}(x)$	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0

Fig. 4.12a Yager's fuzzy maximum for discrete functions.

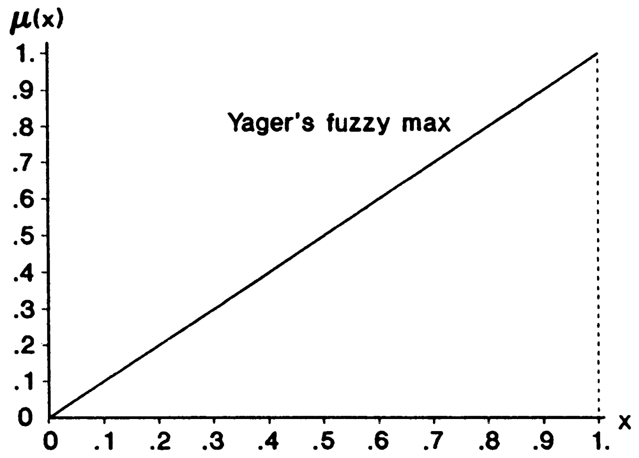
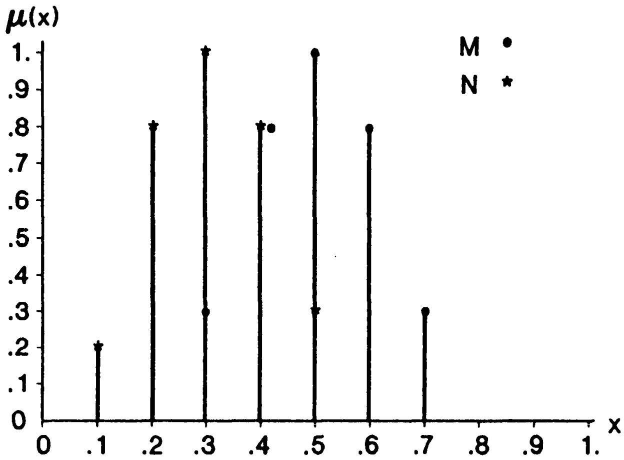


Fig. 4.12b Yager's fuzzy maximum for continuous functions.



x,y	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
$\mu_M(x)$	0	0	0	.3	.8	1	.8	.3	0	0	0
$\mu_N(y)$	0	.2	.8	1	.8	.3	0	0	0	0	0

Fig. 4.13 Discrete fuzzy sets M and N.

This ranking procedure can be applied to fuzzy sets with a piecewise linear membership function. Given fuzzy set M and Yager's fuzzy max as depicted in Fig. 4.14a, the Hamming distance between Yager's fuzzy max and M can be seen as designated by the shaded areas. It is calculated using Eq.(4.23) as:

$$d(\max,M) = \int_0^1 | \mu_{\max}(x) - \mu_M(x) | \, dx$$

$$= \int_0^{.375} x \, dx - \int_{.3}^{.375} (\frac{x-.3}{.2}) \, dx + \int_{.375}^{.5} ((\frac{x-.3}{.2}) - x) \, dx$$

$$+ \int_{.50}^{.58} ((\frac{.7-x}{.2}) - x) \, dx + \int_{.58}^{.70} (x - (\frac{.7-x}{.2})) \, dx$$

$$+ \int_{.7}^{1.0} x \, dx = .433$$

Similarly, the Hamming distance between Yager's fuzzy max and fuzzy set N in Fig. 4.14b is d(max,N) = .50. The result is M > N.

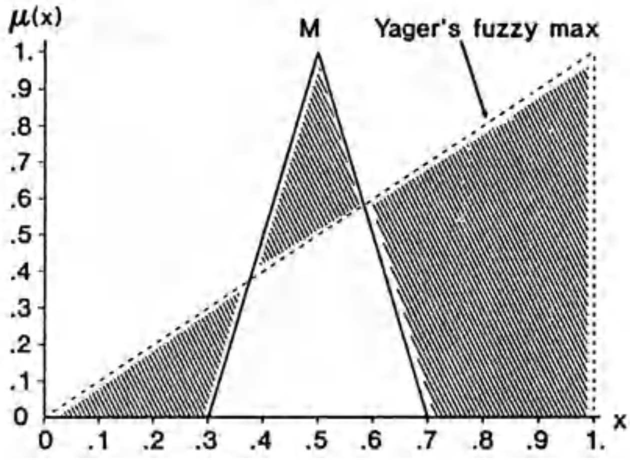


Fig. 4.14a The Hamming distance between Yager's fuzzy max and M.

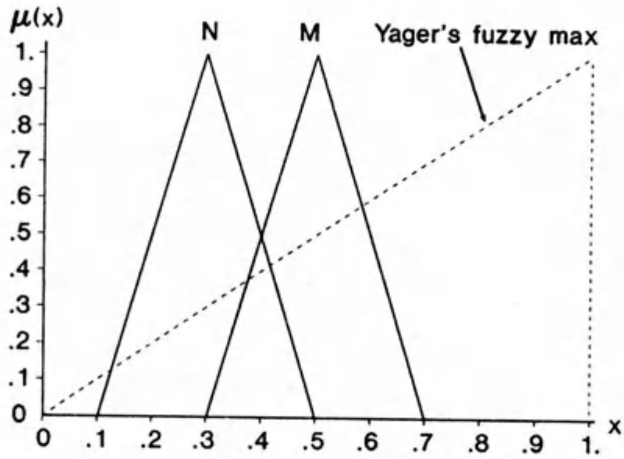


Fig. 4.14b Yager's fuzzy max and fuzzy sets M and N.

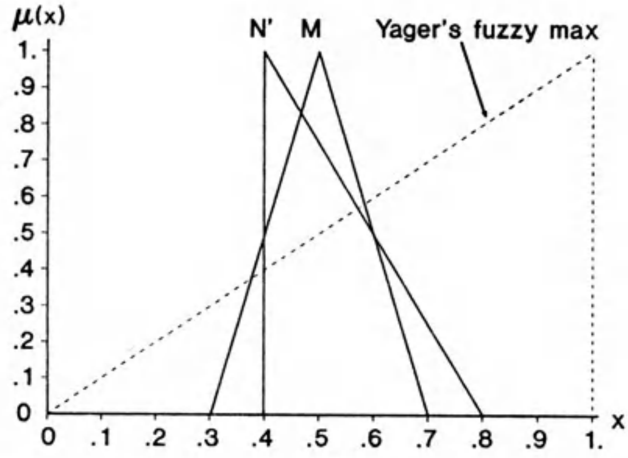


Fig. 4.14c Yager's fuzzy max, M, and N'.

When N becomes N′ as shown in Fig. 4.14c, the Hamming distance of N′ to Yager's fuzzy max is d(max,N′) = .403. Comparing d(max,M) = .433 with d(max,N′) = .403, the smaller Hamming distance to Yager's fuzzy max indicates that N′ > M.

Numerical Example

Given three fuzzy final ratings U_1, U_2, and U_3 as shown in Fig. 4.15, the Hamming distances between Yager's fuzzy max and U_1, U_2, and U_3 are calculated using Eq.(4.23) as:

$$d(max,U_1) = .48, \quad d(max,U_2) = .47, \text{ and } \quad d(max,U_3) = .42.$$

The resulting ranking order is: $U_3 > U_2 > U_1$.

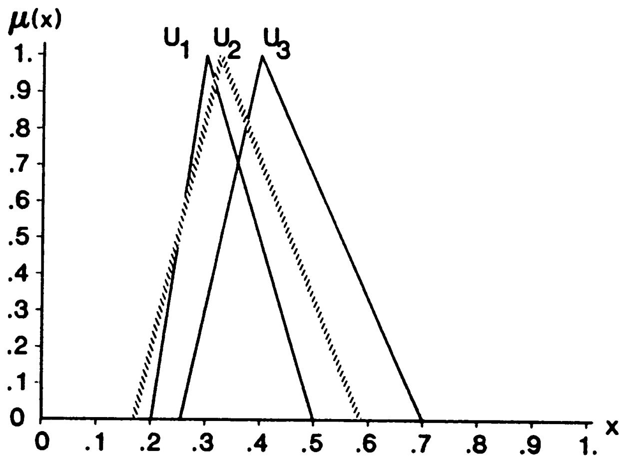


Fig. 4.15 Fuzzy final ratings U_1, U_2, and U_3.

Note

1. Yager's approach is against common sense. Let us consider fuzzy sets with only a single element in them. For example, there are two fuzzy sets M and N defined as:

$$M = \{(.2,.2)\} \qquad N = \{(.2,.8)\}.$$

Our intuition suggests that N > M, because N represents a higher degree of confidence for x = 0.2.

If Yager's method is used, the Hamming distances from Yager's fuzzy max to M and N, respectively, are:

$$d(\max,M) = 5.3, \qquad d(\max,N) = 5.9.$$

The fuzzy set M is preferred. This result is against our intuition.

2. Yager's index is not logically sound. For example, let us consider three fuzzy numbers M_1, M_2, and M_3 as shown in Fig. 4.16. The fuzzy number M_3 is a crisp number 1.0. The Hamming distance from M_3 to Yager's fuzzy max is $d(\max,M_3) = 0.5$. For the fuzzy numbers M_1 and M_2, we get $d(\max,M_1) < d(\max,M_2) < 0.5$. Based on Yager's index, the resulting ranking order is: $M_1 > M_2 > M_3$. This is contrary to the obvious fact that $M_3 > M_2 > M_1$.

Lee and Li [L3] pointed out that human intuition would favor a fuzzy number with the following characteristics: higher mean value and at the same time lower spread. The fuzzy final rating M_1 in Fig. 4.16 possesses none of the two characteristics. Thus, the ranking order $M_1 > M_2 > M_3$ is against human intuition.

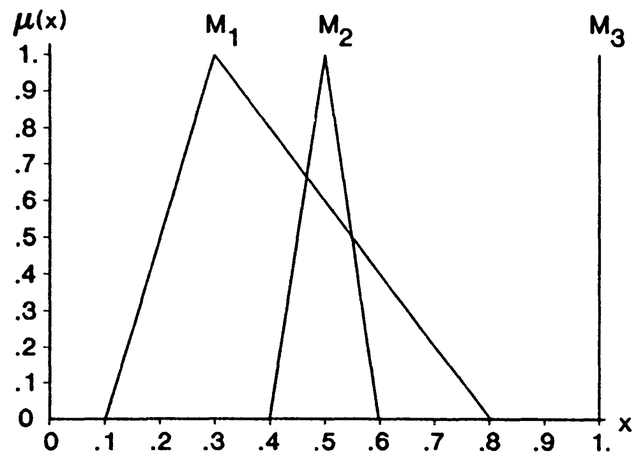


Fig. 4.16 Illustration of the logical weakness of Yager's method.