Institute for Advanced Management Systems Research
Department of Information Technologies
Åbo Akademi University

Aggregation Operators - Tutorial

Robert Fullér

Directory

© 2010
October 2, 2010

rfuller@abo.fi

Table of Contents

1. Averaging Operators

In a decision process the idea of *trade-offs* corresponds to viewing the global evaluation of an action as lying between the *worst* and the *best* local ratings.

This occurs in the presence of conflicting goals, when a compensation between the corresponding compabilities is allowed.

Averaging operators realize trade-offs between objectives, by allowing a positive compensation between ratings.

Definition 1.1. *An averaging (or mean) operator M is a function*

$$M\colon [0,1]\times[0,1]\to[0,1],$$

satisfying the following properties

- $M(x,x)=x,\ \forall x\in[0,1],$ *(idempotency)*
- $M(x,y)=M(y,x),\ \forall x,y\in[0,1],$ *(commutativity)*
- $M(0,0)=0,\ M(1,1)=1,$ *(extremal conditions)*
- $M(x,y)\le M(x',y')$ *if* $x\le x'$ *and* $y\le y'$ *(monotonicity)*
- M *is continuous.*

Lemma 1. *If M is an averaging operator then*

$$\min\{x,y\}\le M(x,y)\le\max\{x,y\},$$

for all $x,y\in[0,1]$.

Proof. From idempotency and monotonicity of M it follows that

$$\min\{x, y\} = M(\min\{x, y\}, \min\{x, y\}) \leq M(x, y)$$

and

$$M(x, y) \leq M(\max\{x, y\}, \max\{x, y\}) = \max\{x, y\}.$$

Which ends the proof. □

An important family of averaging operators is formed by quasi-arithmetic means

$$M(a_1, \dots, a_n) = f^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} f(a_i)\right)$$

This family has been characterized by Kolmogorov as being the class of all decomposable continuous averaging operators.

Example 1.1. *For example, the quasi-arithmetic mean of* a_1 *and* a_2 *is defined by*

$$M(a_1, a_2) = f^{-1}\left[\frac{f(a_1) + f(a_2)}{2}\right].$$

The most often used mean operators.

- harmonic mean: $\dfrac{2xy}{x+y}$
- geometric mean: $\sqrt{xy}$
- arithmetic mean: $\dfrac{x+y}{2}$

- dual of geometric mean: $1-\sqrt{(1-x)(1-y)}$
- dual of harmonic mean: $\dfrac{x+y-2xy}{2-x-y}$
- median

$$\mathrm{med}\,(x,y,\alpha)=\begin{cases} y & \text{if } x\le y\le\alpha \\ \alpha & \text{if } x\le\alpha\le y \\ x & \text{if } \alpha\le x\le y \end{cases}$$

- generalized p-mean:

$$\left[\frac{x^p+y^p}{2}\right]^{1/p},\ p\ge 1.$$

2. OWA operators

The process of information aggregation appears in many applications related to the development of intelligent systems.

One sees aggregation in neural networks, fuzzy logic controllers, vision systems, expert systems and multi-criteria decision aids.

Ronald R. Yager [Ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Trans. on Systems, Man and Cybernetics*, 18(1988) 183-190] introduced a new aggregation technique based on the ordered weighted averaging (OWA) operators.

Definition 2.1. *An OWA operator of dimension* n *is a mapping* $F\colon \mathbb{R}^n \to \mathbb{R}$, *that has an associated vector* $W = (w_1, \dots, w_n)$

such as $w_i \in [0, 1],\ 1 \le i \le n$,

$$\sum_{i=1}^{n} w_i = 1.$$

Furthermore

$$F(a_1, \dots, a_n) = \sum_{j=1}^{n} w_j b_j$$

where b_j *is the* j*-th largest element of the bag*

$$< a_1, \dots, a_n > .$$

Example 2.1. *Assume*

$$W = (0.4, 0.3, 0.2, 0.1)^T$$

then

$$F(.7, 1, .2, .6) = .4 \times 1 + .3 \times .7 + .2 \times .6 + .1 \times .2 = .75.$$

A fundamental aspect of this operator is the re-ordering step, in particular an aggregate a_i is not associated with a particular weight w_i but rather a weight is associated with a particular ordered position of aggregate.

When we view the OWA weights as a column vector we shall find it convenient to refer to the weights with the low indices as weights at the top and those with the higher indices with weights at the bottom.

It is noted that different OWA operators are distinguished by their weighting function.

Yager pointed out three important special cases of OWA aggregations:

- F^*: In this case $W = W^* = (1, 0 \ldots, 0)^T$ and
$$F^*(a_1, \ldots, a_n) = \max\{a_1, \ldots, a_n\},$$
- F_*: In this case $W = W_* = (0, 0 \ldots, 1)^T$ and
$$F_*(a_1, \ldots, a_n) = \min\{a_1, \ldots, a_n\},$$
- F_{mean}: In this case
$$W = W_A = (1/n, \ldots, 1/n)^T$$
and
$$F_{mean}(a_1, \ldots, a_n) = \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

A number of important properties can be associated with the OWA operators. We shall now discuss some of these. For any OWA operator F

$$F_*(a_1, \ldots, a_n) \leq F(a_1, \ldots, a_n) \leq F^*(a_1, \ldots, a_n).$$

Thus the upper an lower star OWA operator are its boundaries.

From the above it becomes clear that for any F

$$\max\{a_1, \ldots, a_n\} \leq F(a_1, \ldots, a_n) \leq \max\{a_1, \ldots, a_n\}.$$

The OWA operator can be seen to be *commutative*.

Let $\{a_1, \ldots, a_n\}$ be a bag of aggregates and let

$$\{d_1, \ldots, d_n\}$$

be any *permutation* of the a_i. Then for any OWA operator

$$F(a_1, \ldots, a_n) = F(d_1, \ldots, d_n).$$

A second characteristic associated with these operators is *monotonicity*. Assume a_i and c_i are a collection of aggregates, $i = 1, \ldots, n$ such that for each i, $a_i \geq c_i$. Then

$$F(a_1, \ldots, a_n) \geq F(c_1, c_2, \ldots, c_n)$$

where F is some fixed weight OWA operator. Another characteristic associated with these operators is *idempotency*.

If $a_i = a$ for all i then for any OWA operator

$$F(a_1, \ldots, a_n) = F(a, \ldots, a) = a.$$

From the above we can see the OWA operators have the basic properties associated with an *averaging operator*.

Example 2.2. *A window type OWA operator takes the average of the m arguments about the center. For this class of operators we have*

$$w_i = \begin{cases} 0 & \textit{if } i < k \\ 1/m & \textit{if } k \le i < k+m \\ 0 & \textit{if } i \ge k+m \end{cases}$$

In order to classify OWA operators in regard to their location between *and* and *or*, a measure of *orness*, associated with any vector W is introduce by Yager as follows

$$\boxed{\mathrm{orness}(W) = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i}$$

$$\textit{orness}(W) = \frac{1}{n-1} \times \left[(n-1)w_1 + \cdots + w_{n-1}\right]$$

$$orness(W) = w_1 + \frac{n-2}{n-1} \times w_2 + \cdots + \frac{1}{n-1} \times w_{n-1}$$

It is easy to see that for any W the *orness*(W) is always in the unit interval. Furthermore, note that the nearer W is to an *or*, the closer its measure is to one; while the nearer it is to an *and*, the closer is to zero.

Lemma 2. *Let us consider the the vectors*

$$W^* = (1, 0 \ldots, 0)^T,$$

$$W_* = (0, 0 \ldots, 1)^T,$$

$$W_{mean} = (1/n, \ldots, 1/n)^T.$$

Then

$$orness(W^*) = 1, orness(W_*) = 0, orness(W_{mean}) = 0.5$$

Proof.

$$orness(W^*) = \frac{1}{n-1}\Big((n-1)w_1 + \cdots + w_{n-1}\Big) =$$

$$\frac{1}{n-1}\Big((n-1) + \cdots + 0\Big) = 1.$$

$$\boldsymbol{orness}(W_*) = \frac{1}{n-1}\Big((n-1)w_1 + \cdots + w_{n-1}\Big)$$

$$= \frac{1}{n-1}\Big(0 + \cdots + 0\Big) = 0.$$

$$orness(W_{mean}) = \frac{1}{n-1}\left[\frac{n-1}{n} + \cdots + \frac{1}{n}\right] = \frac{n(n-1)}{2n(n-1)} = 0.5.$$

□

A measure of *andness* is defined as

$$andness(W) = 1 - orness(W).$$

Generally, an OWA opeartor with much of nonzero weights near the top will be an *orlike* operator,

$$orness(W) \geq 0.5$$

and when much of the weights are nonzero near the bottom, the OWA operator will be *andlike* operator

$$andness(W) \geq 0.5.$$

Example 2.3. *Let* $W = (0.8, 0.2, 0.0)^T$. *Then*

$$orness(W) = \frac{1}{2}(2 \times 0.8 + 0.2) = 0.8 + 1/2 \times 0.2 = 0.9$$

and

$$andness(W) = 1 - orness(W) = 1 - 0.9 = 0.1.$$

This means that the OWA operator, defined by

$$F(a_1, a_2, a_3) = 0.8b_1 + 0.2b_2 + 0.0b_3 = 0.8b_1 + 0.2b_2$$

where b_j *is the* j*-th largest element of the bag*

$$< a_1, a_2, a_3 >,$$

is an orlike *aggregation.*

Suppose we have n appliccants for a Ph.D. program. Each application is evaluated by experts, who provides ratings on each of the criteria from the set

- 3 (high)
- 2 (medium)
- 1 (low)

Compensative connectives have the property that a higher degree of satisfaction of one of the criteria can compensate for a lower degree of satisfaction of another criteria.

Oring the criteria means full compensation and *Anding* the criteria means no compensation.

We illustrate the effect of compensation rate on the overall rating:

Let us have the following ratings $(3, 2, 1)$. If $w = (w_1, w_2, w_3)$ is an OWA weight then

$$orness(w_1, w_2, w_3) = w_1 + \frac{1}{2}w_2.$$

Min operator: the overalling rating is

$$\min\{3, 2, 1\} = 1,$$

in this case there is no compensation, because

$$orness(0, 0, 1) = 0.$$

Max operator: the overalling rating is

$$\max\{3, 2, 1\} = 3$$

in this case there is full compensation, because

$$orness(1, 0, 0) = 1.$$

Mean operator: the overalling rating is $\frac{1}{3}(3 + 2 + 1) = 2$ and in this case the measure of compensation is 0.5, because

$$orness(1/3, 1/3, 1/3) = 1/3 + 1/2 \times 1/3 = 1/2.$$

An andlike operator: the overalling rating is

$$0.2 \times 3 + 0.1 \times 2 + 0.7 \times 1 = 1.5$$

in this case the measure of compensation is 0.25, because

$$orness(0.2, 0.1, 0.7) = 0.2 + 1/2 \times 0.1 = 0.25$$

An orlike operator: the overalling rating is

$$0.6 \times 3 + 0.3 \times 2 + 0.1 \times 1 = 2.5$$

in this case the measure of compensation is 0.75, because

$$orness(0.6, 0.3, 0.1) = 0.6 + 1/2 \times 0.3 = 0.75$$

Yager defined the measure of dispersion (or entropy) of an OWA vector by

$$\boxed{disp(W) = -\sum_i w_i \ln w_i.}$$

We can see when using the OWA operator as an averaging operator $Disp(W)$ measures the degree to which we use all the aggregates equally.

3. Quantifier guided aggregations

An important application of the OWA operators is in the area of quantifier guided aggregations. Assume

$$\{A_1, \dots, A_n\}$$

is a collection of criteria. Let x be an object such that for any criterion A_i, $A_i(x) \in [0, 1]$ indicates the degree to which this criterion is satisfied by x.

If we want to find out the degree to which x satisfies ”*all* the criteria” denoting this by $D(x)$, we get following Bellman and

Zadeh:

$$D(x) = \min\{A_1(x), \ldots, A_n(x)\}$$

In this case we are essentially requiring x to satisfy

$$A_1 \text{ and } A_2 \text{ and } \ldots \text{ and } A_n.$$

If we desire to find out the degree to which x satisfies *"at least one* of the criteria", denoting this $E(x)$, we get

$$E(x) = \max\{A_1(x), \ldots, A_n(x)\}$$

In this case we are requiring x to satisfy

$$A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_n.$$

In many applications rather than desiring that a solution satisfies one of these extreme situations, "all" or "at least one", we may require that x satisfies *most* or *at least half* of the criteria.

Drawing upon Zadeh's concept of linguistic quantifiers we can accomplish these kinds of quantifier guided aggregations.

Definition 3.1. *A quantifier* $Q\colon [0,1] \to [0,1]$ *is called* regular monotonically non-decreasing *if* $Q(0) = 0$ *and* $Q(1) = 1$ *and*

$$\textit{if } r_1 > r_2 \textit{ then } Q(r_1) \geq Q(r_2).$$

With $a_i = A_i(x)$ the overall valuation of x is

$$F_Q(a_1, \dots, a_n)$$

where F_Q is an OWA operator. The weights associated with this quantified guided aggregation are obtained as follows

$$\boxed{w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right),\ i = 1, \dots, n.}$$

The next figure graphically shows the operation involved in determining the OWA weights directly from the quantifier guiding the aggregation.

Theorem 3.1. *If we construct w_i via the method above we always get*

$$\sum w_i = 1, w_i \in [0, 1]$$

for any function

$$Q\colon [0, 1] \to [0, 1]$$

satisfying the conditions of a regular nondecreasing quantifier.

Proof. We first see that from the non-decreasing property

$$Q(i/n) \geq Q(i - 1/n)$$

hence $w_i \geq 0$ and since $Q(r) \leq 1$ then $w_i \leq 1$. Furthermore

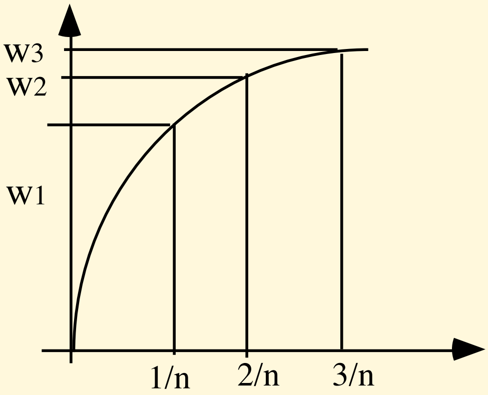

Figure 1: Deriving OWA operator weights from a quantifier.

we see

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \left[Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \right] =$$

$$Q\left(\frac{n}{n}\right) - Q\left(\frac{0}{n}\right) = 1 - 0 = 1.$$

we call any function satisfying the conditions of a regular non-decreasing quantifier an *acceptable OWA weight generating function*. □

Let us look at the weights generated from some basic types of quantifiers. The quantifier, *for all* Q_*, is defined such that

$$Q_*(r) = \begin{cases} 0 & \text{for } r < 1, \\ 1 & \text{for } r = 1. \end{cases}$$

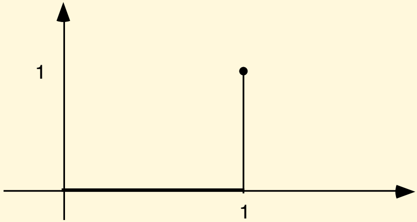

Figure 2: The quantifier *all*.

Using our method for generating weights

$$w_i = Q_*\left(\frac{i}{n}\right) - Q_*\left(\frac{i-1}{n}\right)$$

we get

$$w_i = \begin{cases} 0 & \text{for } i < n, \\ 1 & \text{for } i = n. \end{cases}$$

This is exactly what we previously denoted as W_*.

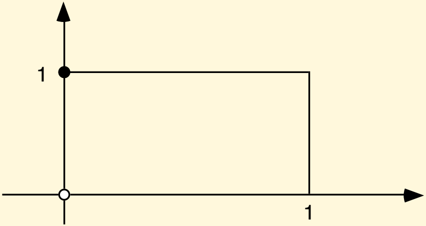

Figure 3: The quantifier *there exists*.

For the quantifier *there exists* we have

$$Q^*(r) = \begin{cases} 0 & \text{for } r = 0, \\ 1 & \text{for } r > 0. \end{cases}$$

In this case we get

$$w_1 = 1, \quad w_i = 0, \text{ for } i \neq 1.$$

This is exactly what we denoted as W^*.

Consider next the quantifier defined by

$$Q(r) = r.$$

This is *an identity* or *linear type* quantifier.

In this case we get

$$w_i = Q(\frac{i}{n}) - Q(\frac{i-1}{n}) = \frac{i}{n} - \frac{i-1}{n} = \frac{1}{n}.$$

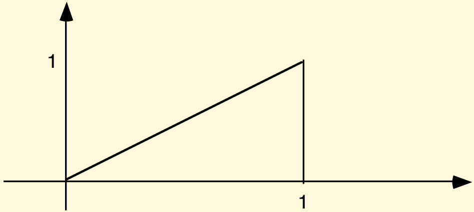

Figure 4: The *identity* quantifier.

This gives us the pure averaging OWA aggregation operator.

Recapitulating using the approach suggested by Yager if we desire to calculate

$$F_Q(a_1, \dots, a_n)$$

for Q being a regular non-decreasing quantifier we proceed as follows:

(1) Calculate:

$$w_i = Q(\frac{i}{n}) - Q(\frac{i-1}{n}),$$

(2) Calculate:

$$F_Q(a_i, \dots, a_n) = \sum_{i=1}^{n} w_i b_i,$$

where b_i is the i-th largest of the a_j.

4. OWA operators: Further issues

In order to classify OWA operators in regard to their location between *and* and *or*, a measure of *orness*, associated with any vector W is introduced by Yager as follows

$$orness(W) = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i.$$

Given a lingusitic quantifier Q if we generate the wights by

$$w_j = Q\left(\frac{j}{n}\right) - Q\left(\frac{j-1}{n}\right)$$

we can associate with this quantifier a degree of orness as

$$orness(Q) = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)\left[Q(\frac{i}{n}) - Q(\frac{i-1}{n})\right]$$

Algebraic manipulation of the formula leads to the form

$$orness(Q) = \frac{1}{n-1}\sum_{i=1}^{n} Q(\frac{i}{n})$$

Furthermore iw we let $b \to \infty$ then we can show that

$$\boxed{orness(Q) = \int_0^1 Q(r)\, dr}$$

The standard degree of orness associated with a Regular Increasing Monotone (RIM) linguistic quantifier is equal to the area under the quantifier.

This standard definition for the measure of orness of quantifier provides a simple useful method for obtaining this measure.

If we consider the quantifier

$$Q(r) = r^{\alpha},\ \alpha \geq 0,$$

then

$$orness(Q) = \int_0^1 r^{\alpha}\, dr = \frac{1}{\alpha + 1}.$$

A number of interesting properties can be associated with this measure of orness.

If Q_1 and Q_2 are two quantifiers such that

$$Q_1(r) \geq Q_2(r)$$

for all r then

$$orness(Q_1) \geq orness(Q_2).$$

In addition since any regular quantifier is normal we see that

$$orness(Q) = 0 \iff Q = Q_*$$

where q_* is the quantifier *for all*.

If we consider again the quantifier

$$Q(r) = r^\alpha, \ \alpha \geq 0,$$

then there are three special cases of these family are worth noting

- For $\alpha = 1$ we get $Q(r) = r$. This called the *unitor* quantifier.
- For $\alpha \to \infty$ we get Q_*, the *universal* quantifier.
- For $\alpha \to 0$ we get Q^*, the *existential* quantifier.

A proportional type quantifier, such as *most*, can be represented as a fuzzy set Q of the segment $[0, 1]$, where Q is the linguistic quantifier.

In this representation, the fuzzy set Q is defined such that for each $r \in [0, 1]$, the membership grade, $Q(r)$, indicates the degree to which the proportion r satisfies the concept, linguistic quantifier, which Q is representing.

Assume $X = \{x_1, \dots, x_n\}$ is some collection of objects and A is some concept expressed as a fuzzy set of X.

A quantified proposition is a statement of the form

$$Q\ X\text{'s are } A$$

Examples of these kinds of statements

Most students are young

The truth value of these quantified propositions can be evaluated using the OWA aggregation operator.

Once having obtained the wights from quantifier Q, we can obtain the truth τ, of the proposition

$$Q\ X\text{'s are } A$$

as

$$\tau = \sum_{i=1}^{n} w_i b_i$$

where b_i is the i-th largest element from the bag

$$< A(x_1), A(x_2) \ldots, A(x_n) > .$$

Let us have five students with ages

$$\{21, 22, 24, 22, 38\},$$

and suppose that the degrees of *youngness* of students obtained from a fuzzy set for *young students*

$$A(u) = \begin{cases} 1 & \text{if } u \leq 22 \\ 1 - \dfrac{u - 22}{8} & \text{if } 22 \leq u \leq 30 \\ 0 & \text{otherwise} \end{cases}$$

That is $A(21) = 1$, $A(22) = 1$, $A(24) = 0.75$, $A(22) = 1$ and $A(38) = 0$.

If we order the bag

$$< 1, 1, 0.75, 1, 0 >$$

in monotone decreasing order then we get

$$(1, 1, 1, 0.75, 0).$$

Let us define the quantifier *most* by the following membership function

$$Q(r) = \begin{cases} 1 & \text{if } r \geq 0.8 \\ 1 - \dfrac{0.8 - r}{0.6} & \text{if } 0.2 \leq r \leq 0.8 \\ 0 & \text{otherwise} \end{cases}$$

So, the OWA weights are

$$\begin{aligned} w_1 &= Q(0.2) - Q(0) = 0, \\ w_2 &= Q(0.4) - Q(0.2) = 1/3, \\ w_3 &= Q(0.6) - Q(0.4) = 1/3, \\ w_4 &= Q(0.8) - Q(0.6) = 1/3, \\ w_5 &= Q(1) - Q(0.8) = 0. \end{aligned}$$

Then the degree of truth (which can be condsidered as the unit score of the bag, if the elements of the bag are criteria satisfactions) is

$$\tau = 0 \times 1 + 1/3 \times 1 + 1/3 \times 1 + 1/3 \times 0.75 + 0 \times 0 = \frac{11}{12}.$$

5. Triangular norms

Triangular norms were introduced by Schweizer and Sklar in 1963 to model distances in probabilistic metric spaces.

In fuzzy sets theory triangular norms are extensively used to model logical connective *and*.

Definition 5.1. *(Triangular norm.) A mapping*

$$T\colon [0,1]\times[0,1]\to[0,1]$$

is a triangular norm (t-norm for short) iff it is symmetric, associative, non-decreasing in each argument and $T(a,1)=a$*, for all* $a\in[0,1]$*. In other words, any t-norm* T *satisfies the properties:*

Symmetricity:

$$T(x, y) = T(y, x),\ \forall x, y \in [0, 1].$$

Associativity:

$$T(x, T(y, z)) = T(T(x, y), z),\ \forall x, y, z \in [0, 1].$$

Monotonicity:

$$T(x, y) \leq T(x', y') \textit{ if } x \leq x' \textit{ and } y \leq y'.$$

One identy:

$$T(x, 1) = x,\ \forall x \in [0, 1].$$

These axioms attempt to capture the basic properties of set intersection. The basic t-norms are:

- minimum: $\min(a, b) = \min\{a, b\}$,

- Łukasiewicz: $T_L(a,b) = \max\{a+b-1, 0\}$
- product: $T_P(a,b) = ab$
- weak:

$$T_W(a,b) = \begin{cases} \min\{a,b\} & \text{if } \max\{a,b\} = 1 \\ 0 & \text{otherwise} \end{cases}$$

- Hamacher:

$$H_\gamma(a,b) = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)}, \ \gamma \geq 0 \qquad (1)$$

- Dubois and Prade:

$$D_\alpha(a,b) = \frac{ab}{\max\{a,b,\alpha\}}, \ \alpha \in (0,1)$$

- Yager:

$$Y_p(a,b) = 1 - \min\{1, \sqrt[p]{[(1-a)^p + (1-b)^p]}\}, \ p > 0$$

- Frank:

$$F_\lambda(a,b) = \begin{cases} \min\{a,b\} & \text{if } \lambda = 0 \\ T_P(a,b) & \text{if } \lambda = 1 \\ T_L(a,b) & \text{if } \lambda = \infty \\ 1 - \log_\lambda \left[1 + \dfrac{(\lambda^a - 1)(\lambda^b - 1)}{\lambda - 1}\right] & \text{otherwise} \end{cases}$$

All t-norms may be extended, through associativity, to $n > 2$ arguments. The minimum t-norm is automatically extended and

$$T_P(a_1, \ldots, a_n) = a_1 \times a_2 \times \cdots \times a_n,$$

$$T_L(a_1, \ldots a_n) = \max\left\{\sum_{i=1}^{n} a_i - n + 1, 0\right\}.$$

Lemma 3. *Let T be a t-norm. Then the following statement holds*

$$T_W(x, y) \leq T(x, y) \leq \min\{x, y\},\ \forall x, y \in [0, 1].$$

Proof. From monotonicity, symmetricity and the extremal condition we get

$$T(x, y) \leq T(x, 1) \leq x,\ T(x, y) = T(y, x) \leq T(y, 1) \leq y.$$

This means that $T(x, y) \leq \min\{x, y\}$. □

Lemma 4. $T(a, a) = a$ *holds for any* $a \in [0, 1]$ *if and only if* T *is the minimum norm.*

Proof. If $T(a,b) = \min(a,b)$ then $T(a,a) = a$ holds obviously.

Suppose $T(a,a) = a$ for any $a \in [0,1]$, and $a \leq b \leq 1$. We can obtain the following expression using monotonicity of T

$$a = T(a,a) \leq T(a,b) \leq \min\{a,b\}.$$

From commutativity of T it follows that

$$a = T(a,a) \leq T(b,a) \leq \min\{b,a\}.$$

These equations show that

$$T(a,b) = \min\{a,b\}$$

for any $a, b \in [0,1]$. □

The operation *intersection* can be defined by the help of triangular norms.

Definition 5.2. *(t-norm-based intersection) Let T be a t-norm. The T-intersection of A and B is defined as*

$$(A \cap B)(t) = T(A(t), B(t)),\ \forall t \in X.$$

Example 5.1. *Let $T(x, y) = T_L(x, y) = \max\{x + y - 1, 0\}$ be the Łukasiewicz t-norm. Then we have*

$$(A \cap B)(t) = \max\{A(t) + B(t) - 1, 0\},\ \forall t \in X.$$

Lemma 5. *The distributive law of t-norm T on the* max *operator holds for any a, b, $c \in [0, 1]$.*

$$T(\max\{a, b\}, c) = \max\{T(a, c), T(b, c)\}.$$

6. Triangular conorms

Triangular conorms are extensively used to model logical connective *or*.

Definition 6.1. *(Triangular conorm.) A mapping*

$$S\colon [0,1]\times[0,1]\to[0,1],$$

is a triangular conorm (t-conorm) if it is symmetric, associative, non-decreasing in each argument and $S(a,0)=a$, for all $a\in[0,1]$. In other words, any t-conorm S satisfies the properties:

$S(x,y)=S(y,x)$ *(symmetricity)*

$S(x,S(y,z))=S(S(x,y),z)$ *(associativity)*

$S(x,y)\le S(x',y')$ *if* $x\le x'$ *and* $y\le y'$ *(monotonicity)*

$S(x,0)=x,\ \forall x\in[0,1]$ *(zero identy)*

If T is a t-norm then the equality

$$S(a,b):=1-T(1-a,1-b),$$

defines a t-conorm and we say that S is derived from T. The basic t-conorms are:

- maximum: $\max(a, b) = \max\{a, b\}$
- Łukasiewicz: $S_L(a, b) = \min\{a + b, 1\}$
- probabilistic: $S_P(a, b) = a + b - ab$
- strong:

$$STRONG(a, b) = \begin{cases} \max\{a, b\} & \text{if } \min\{a, b\} = 0 \\ 1 & \text{otherwise} \end{cases}$$

- Hamacher:

$$HOR_\gamma(a, b) = \frac{a + b - (2 - \gamma)ab}{1 - (1 - \gamma)ab}, \; \gamma \geq 0$$

- Yager:

$$YOR_p(a,b) = \min\{1, \sqrt[p]{a^p + b^p}\},\ p > 0.$$

Lemma 6. *Let S be a t-conorm. Then the following statement holds*

$$\max\{a,b\} \leq S(a,b) \leq STRONG(a,b),\ \forall a,b \in [0,1]$$

Proof. From monotonicity, symmetricity and the extremal condition we get

$$S(x,y) \geq S(x,0) \geq x,\ S(x,y) = S(y,x) \geq S(y,0) \geq y$$

This means that $S(x,y) \geq \max\{x,y\}$. □

The operation *union* can be defined by the help of triangular conorms.

Definition 6.2. *(t-conorm-based union) Let S be a t-conorm. The S-union of A and B is defined as*

$$(A \cup B)(t) = S(A(t), B(t)), \ \forall t \in X.$$

Example 6.1. *Let $S(x, y) = \min\{x + y, 1\}$ be the Łukasiewicz t-conorm. Then we have*

$$(A \cup B)(t) = \min\{A(t) + B(t), 1\}, \ \forall t \in X.$$

In general, *the law of the excluded middle* and the *noncontradiction principle* properties are not satisfied by t-norms and t-conorms defining the intersection and union operations. However, the Łukasiewicz t-norm and t-conorm do satisfy these properties.

Lemma 7. *If $T(x, y) = T_L(x, y) = \max\{x + y - 1, 0\}$ then the law of noncontradiction is valid.*

Proof. Let A be a fuzzy set in X. Then from the definition of t-norm-based intersection we get

$$(A\cap\neg A)(t) = T_L(A(t), 1-A(t)) = (A(t)+1-A(t)-1)\vee 0 = 0, \ \forall t$$

□

Lemma 8. *If* $S(x, y) = S_L(x, y) = \min\{1, x + y\}$, *then* the law of excluded middle *is valid.*

Proof. Let A be a fuzzy set in X. Then from the definition of t-conorm-based union we get

$$(A\cup\neg A)(t) = S_L(A(t), 1-A(t)) = (A(t)+1-A(t))\wedge 1 = 1,$$

for all $t \in X$. □

7. MICA Operators

Let us look at the process for combining the individual ratings

$$(A_1(x), \ldots, A_n(x))$$

of an alternative x.

A basic assumption we shall make is that the operation is *point-wise* and *likewise*.

By *pointwise* we mean that for every alternative x, $A(x)$ just depends upon

$$A_j(x),\ j = 1, \ldots, n.$$

By *likewise* we mean that the process used to combine the A_j is the same for all of the x.

Let us denote the pointwise process we use to combine the individual ratings as

$$A(x) = \mathbf{Agg}\left[A_1(x), \ldots, A_n(x)\right]$$

In the above **Agg** is called the aggregation operator and the $A_j(x)$ are the arguments.

More generally, we can consider this as an operator

$$a = \mathbf{Agg}\left(a_1, \ldots, a_n\right)$$

where the a_i and a are values from the membership grade space, normally the unit interval.

Let us look at the minimal requirements associated with **Agg**. We first note that the combination of of the individual ratings should be independent of the choice of indexing of the criteria.

This implies that a required property that we must associate with the Agg operator is that of commutativity, the indexing of the arguments does not matter.

We note that the commutativity property allows to represent the arguments of the **Agg** operator, as an unordered collection of possible duplicate values; such an object is a **bag**.

For an individual rating, A_j, the membership grade $A_j(x)$ indicates the degree or strength to which this rate suggests that x is the appropriate (best compromise) solution.

In particular if for a pair of elements x_1 and x_2 it is the case that

$$A_j(x_1) \geq A_j(x_2),$$

then we are saying that the j-th rating is preferring x_1 as the output over x_2.

From this we can reasonably conclude that if all criterions prefer x_1 over x_2 as output then the overall rating should prefer x_1 over x_2.

This observation requires us to impose a monotonicity condition on the **Agg** operation.

In particular if $A_j(x_1) \geq A_j(x_2)$, for all j, then

$$A(x_1) \geq A(x_2).$$

There appears one other condition we need to impose upon the aggregation operator.

Assume that there exists some criterion whose value does not matter.

The implication of this is that the criterion provides no information regarding what should be the overall rating.

It should not affect the final A.

The first observation we can make is that whatever rating this criterion provides should not make make any distinction between the potential overall ratings.

Thus, we see that the aggregation operator needs an *identy element*.

In summary, we see that the aggregation operator, **Agg** must satisfy three conditions:

- commutativity,
- monotonicity,
- must contain a fixed identity

These conditions are based on the three requirements: that the indexing of the criterions be unimportant, a positive association between individual criterion rating and total rating, and irrelevant criteria play no role in the decision process.

These operators are called Monotonic Identity Commutative Aggregation (**MICA**) operators.

- R.R.Yager, Aggregation operators and fuzzy systems modeling, *Fuzzy Sets and Systems*, 67(1994) 129-145.

MICA *operators are the most general class for aggregation in fuzzy modeling*.

They include t-norms, t-conorms, averaging and compensatory operators.

Assume X is a set of elements. A bag drawn from X is any collection of elements which is contained in X. A bag is different from a subset in that it allows multiple copies of the same element. A bag is similar to a set in that the ordering of the elements in the bag does not matter. If A is a bag consisiting of a, b, c, d we denote this as

$$A =< a, b, c, d > .$$

Assume A and B are two bags. We denote the sum of the bags

$$C = A \oplus B$$

where C is the bag consisting of the members of both A and B.

Example 7.1. *Let* $A =< a, b, c, d >$ *and*

$$B =< b, c, c >$$

then

$$A \oplus B =< a, b, c, d, b, c, c >$$

In the following we let $\text{Bag}(X)$ indicate the set of all bags of the set X.

Definition 7.1. *A function $F\colon Bag(X) \to X$ is called a bag mapping from $Bag(X)$ into the set X.*

An important property of bag mappings are that they are commutative in the sense that the ordering of the elements does not matter.

Definition 7.2. *Assume $A =< a_1, \ldots, a_n >$ and*

$$B =< b_1, \ldots, b_n >$$

are two bags of the same cardinality n. If the elements in A

and B can be indexed in such way that $a_i \geq b_i$ for all i then we shall denote this

$$A \geq B.$$

Definition 7.3. *A bag mapping*

$$M : Bag([0,1]) \to [0,1]$$

is called MICA operator if it has the following two properties

- *If $A \geq B$ then $M(A) \geq M(B)$*
- *For every bag A there exists an element, $u \in [0,1]$, called the identity of A such that if*

  $$C = A \oplus < u >$$

  then $M(C) = M(A)$.

Thus the MICA operator is endowed with two properties in addition to the inherent commutativity of the bag operator, *monotonicity and identity*:

- The requirement of monotonicity appears natural for an aggregation operator in that it provides some connection between the arguments and the aggregated value.
- The property of identity allows us to have the facility for aggregating data which does not affect the overall result.

  This becomes useful for enabling us to include importances among other characteristics.

8. Linguistic variables

The use of fuzzy sets provides a basis for a systematic way for the manipulation of vague and imprecise concepts. In particular, we can employ fuzzy sets to represent linguistic variables. A linguistic variable can be regarded either as a variable whose value is a fuzzy number or as a variable whose values are defined in linguistic terms.

Definition 8.1. *(linguistic variable) A linguistic variable is characterized by a quintuple*

$$(x, T(x), U, G, M)$$

in which

- x *is the name of variable;*

- *$T(x)$ is the term set of x, that is, the set of names of linguistic values of x with each value being a fuzzy number defined on U;*

- *G is a syntactic rule for generating the names of values of x;*

- *and M is a semantic rule for associating with each value its meaning.*

For example, if *speed* is interpreted as a linguistic variable, then its term set T (speed) could be

T = {slow, moderate, fast, very slow, more or less fast, sligthly slow, ...}

where each term in T (speed) is characterized by a fuzzy set in

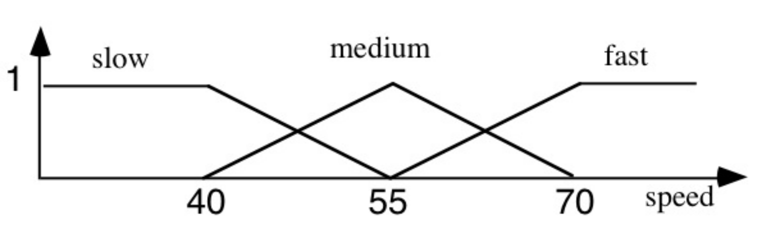

Figure 5: Values of linguistic variable *speed*.

a universe of discourse $U = [0, 100]$. We might interpret

- *slow* as ”a speed below about 40 mph”
- *moderate* as ”a speed close to 55 mph”
- *fast* as ”a speed above about 70 mph”

These terms can be characterized as fuzzy sets whose membership functions are shown in the figure below.

In many practical applications we normalize the domain of inputs and use the following type of fuzzy partition

- NB (Negative Big),
- NM (Negative Medium)
- NS (Negative Small),
- ZE (Zero)
- PS (Positive Small),
- PM (Positive Medium)
- PB (Positive Big)

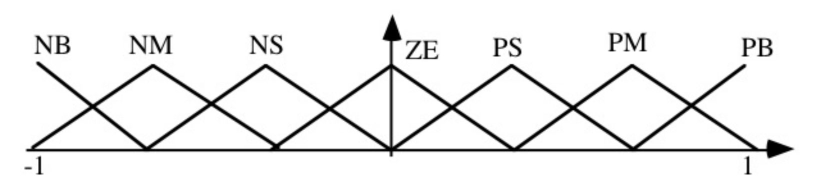

Figure 6: A possible fuzzy partition of $[-1, 1]$.

If A a fuzzy set in X then we can modify the meaning of A with the help of words such as *very, more or less, slightly*, etc. For example, the membership function of fuzzy sets "very A" and "more or less A" can be defined by

$$(\text{very } A)(x) = (A(x))^2,$$

$$(\text{more or less } A)(x) = \sqrt{A(x)},\ \forall x \in X$$

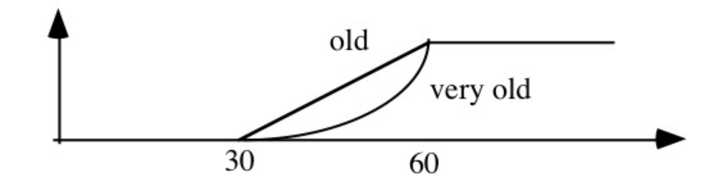

Figure 7: Membership functions of fuzzy sets *old* and *very old*.

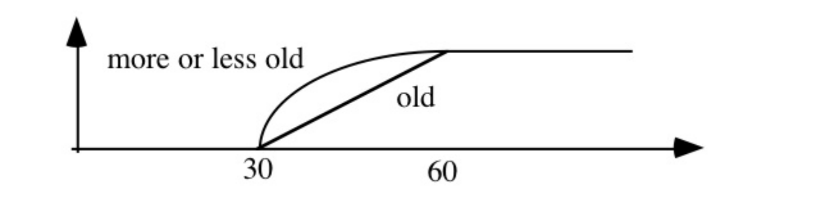

Figure 8: Membership function of fuzzy sets *old* and *more or less old*.

9. The linguistic variable *Truth*

Truth also can be interpreted as linguistic variable with a possible term set

$$T = \{\text{Absolutely false, Very false, False, Fairly true, True, Very true, Absolutely true}\}$$

One may define the membership function of linguistic terms of truth as

$$\mathrm{True}(u) = u, \quad \mathrm{False}(u) = 1 - u$$

for each $u \in [0, 1]$, and

$$\text{Absolutely false}(u) = \begin{cases} 1 & \text{if } u = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Absolutely true}(u) = \begin{cases} 1 & \text{if } u = 1 \\ 0 & \text{otherwise} \end{cases}$$

The words ”Fairly” and ”Very” are interpreted as

$$\text{Fairly true}(u) = \sqrt{u}, \quad \text{Very true}(u) = u^2,$$

$$\text{Fairly false}(u) = \sqrt{1-u}, \quad \text{Very false}(u) = (1-u)^2$$

for each $u \in [0, 1]$.

Suppose we have the fuzzy statement ”x is A”. Let τ be a term of linguistic variable *Truth*. Then the statement ”x is A is τ” is interpreted as ”x is $\tau \circ A$”. Where

$$(\tau \circ A)(u) = \tau(A(u))$$

for each $u \in [0, 1]$. For example, let τ = ”true”. Then ”x is A is true” is defined by ”x is $\tau \circ A$” = ”x is A” because

$$(\tau \circ A)(u) = \tau(A(u)) = A(u)$$

for each $u \in [0, 1]$. It is why "everything we write is considered to be true".

Let τ = "absolutely true". Then the statement "x is A is Absolutely true" is defined by "x is $\tau \circ A$", where

$$(\tau \circ A)(x) = \begin{cases} 1 & \text{if } A(x) = 1 \\ 0 & \text{otherwise} \end{cases}$$

Let τ = "fairly true". Then the statement "x is A is Fairly true" is defined by "x is $\tau \circ A$", where

$$(\tau \circ A)(x) = \sqrt{A(x)}$$

Let τ = "very true". Then the statement "x is A is Fairly true" is defined by "x is $\tau \circ A$", where

$$(\tau \circ A)(x) = (A(x))^2$$

Let τ = "false". Then the statement "x is A is false" is defined by "x is $\tau \circ A$", where

$$(\tau \circ A)(x) = 1 - A(x)$$

Let τ = "absolutely false". Then the statement "x is A is Absolutely false" is defined by "x is $\tau \circ A$", where

$$(\tau \circ A)(x) = \begin{cases} 1 & \text{if } A(x) = 0 \\ 0 & \text{otherwise} \end{cases}$$

10. Linguistic labels

We allow the experts to provide information about satisfactions in the form of a linguistic values such as *high, medium, low*. This ability to perform the necessary operations will only requiring imprecise linguistic preference valuations will enable the experts to comfortably use the kinds of minimally informative sources of information about the objects. These evaluations of alternative satisfaction to criteria will be given in terms of elements from the following scale S:

Excellent (EX)	S_7
Very High (VH)	S_6
High (H)	S_5
Medium (M)	S_4
Low	S_3
Very Low	S_2
None	S_1

The use of such a scale provides a natural ordering, $S_i > S_j$ if $i > j$ and the maximum and minimum of any two scores re defined by

$$\max(S_i, S_j) = S_i \text{ if } S_i \geq S_j,\ \min(S_i, S_j) = S_j \text{ if } S_j \leq S_i$$

The negation operation is defined by

$$Neg(S_i) = S_{7-i+1}$$

For the scale that we are using, we see that the negation operation provides the following:

$$\begin{aligned} Neg(EX) &= N \\ Neg(VH) &= VL \\ Neg(H) &= L \\ Neg(M) &= M \\ Neg(L) &= H \\ Neg(VL) &= VH \\ Neg(N) &= EX \end{aligned}$$

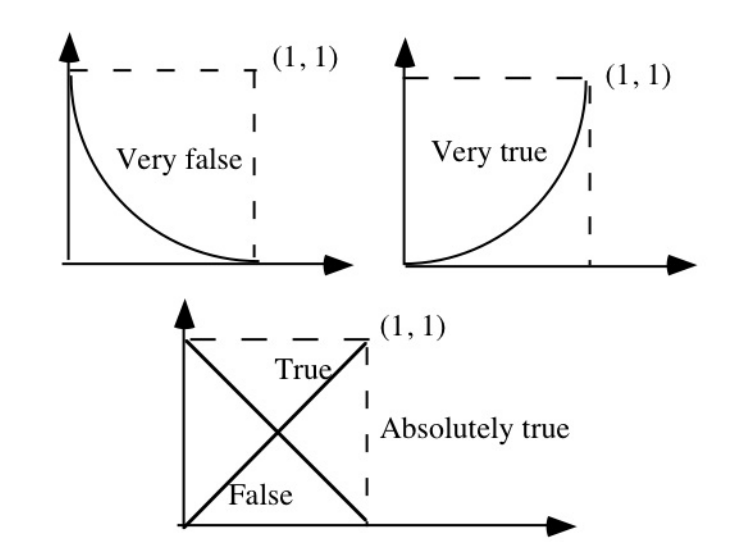

Figure 9: Some values of lingusitic variable *Truth*.

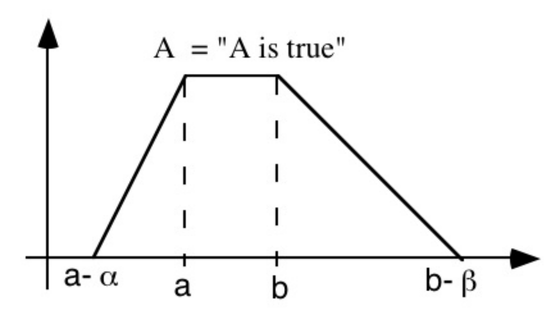

Figure 10: ”A is true”.

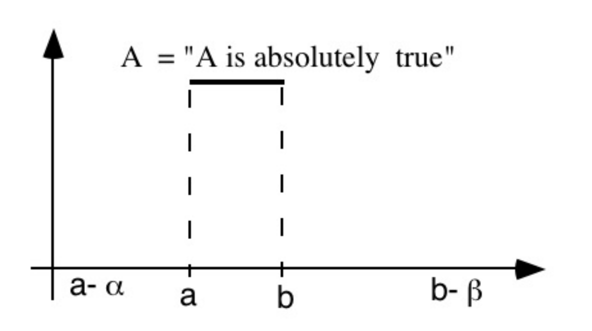

Figure 11: ”A is absolutely true”.

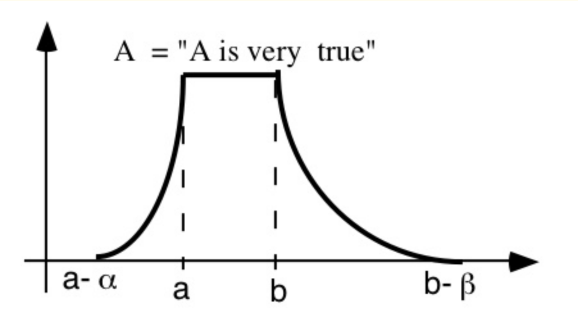

Figure 12: ”A is very true”.

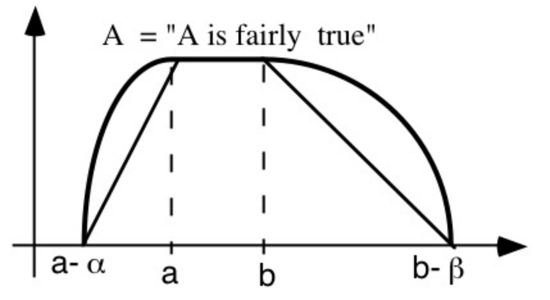

Figure 13: "A is fairly true".

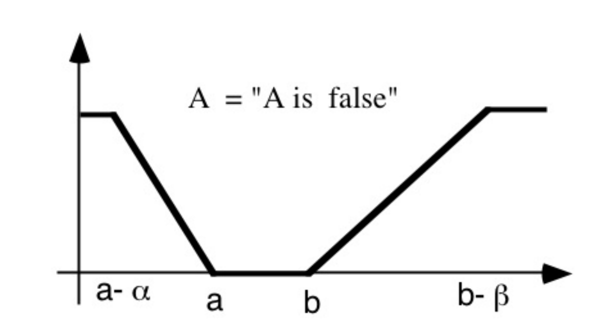

Figure 14: ”A is false”.

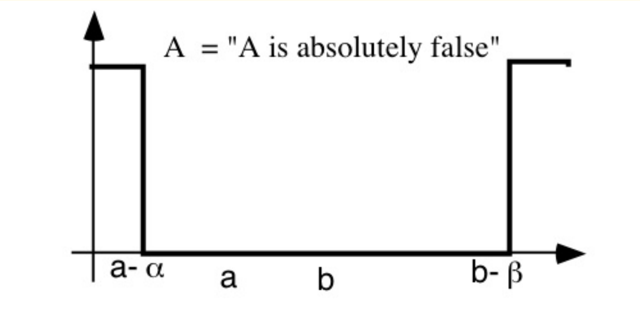

Figure 15: "A is absolutely false".