# Fundamentals on Aggregation Operators*

Marcin DETYNIECKI

Berkeley initiative in Soft Computing
Computer Science Division
University of California, Berkeley
United Sates of America
marcin@eecs.berkeley.edu

* This manuscript can be downloaded from http://www.cs.berkeley.edu/~marcin/agop.pdf. It is based on Detyniecki's doctoral thesis [15]: "Mathematical aggregation operators and their application to video querying", directed by B. Bouchon-Meunier and R. Yager. The complete document is available as a LIP6 research report: 2001-002. Ask for a printed version of the thesis or download it from http://www.lip6.fr/reports/index-eng.html. I also would like to thank Radko Mesiar for all the comments and suggestions.

# Introduction

Aggregation and fusion of information are basic concerns for all kinds of knowledge based systems, from image processing to decision making, from pattern recognition to machine learning. From a general point of view we can say that aggregation has for purpose the simultaneous use of different pieces of information (provided by several sources) in order to come to a conclusion or a decision. Several research groups are directly interested in finding solutions, among them the multi-criteria community, the sensor fusion community, the decision-making community, the data mining community, etc. And each of this groups use or propose some methodologies in order to perform an intelligent aggregation, as for instance the use of rules, the use of neuronal networks, the use of fusion specific techniques, the use of probability theory, evidence theory, possibility theory and fuzzy set theory, etc. But all this approaches are based on some numerical aggregation operator. In other words at some point there is a need of aggregating some numerical values and this is when the numerical aggregation plays a fundamental role.

More generally, the aggregation operators are mathematical objects that have the function of reducing a set of numbers into a unique representative (or meaningful) number. We insist in the mathematical aspect of this aggregation since we are dealing with aggregation of real numbers and not fusion of information at a higher level as for instance the aggregation of rules. But it is important to notice that any aggregation or fusion process done with a computer underlies numerical aggregation. In other words, the mathematical aggregation operators are the key in this kind of processes.

In this manuscript, we try to summarize the different points of view in order to obtain a global vision on the domain. This part will be used as the starting point for the next part.

We start (**part 1**) by adopting a minimal set of mathematical conditions that define an aggregation operator. This is essential, because any mathematical operator that transforms a set of numbers into a unique one does not necessarily give a representative or meaningful final value. More precisely this mathematical axiom set guarantees that we are not obtaining injudicious results.
We proceed by presenting other mathematical properties (not fundamental) that can be expected from any operator. Here, we try to give an intuitive interpretation to the mathematical description.

The second part (**part 2**) presents a catalogue of the existing operators. We portray their characteristics and advantages, but we try to be objective by giving also their disadvantages.

# PART 1 :
# Definition and Properties

## 1.1 Definition

In a rather informal way, the aggregation problem consists in aggregating *n*-tuples of objects all belonging to a given set, into a single object of the same set. In the case of mathematical aggregation operator this set is all the real numbers. In this setting, an aggregation operator is simply a function, which assigns a real number *y* to any *n*-tuple $(x_1, x_2, \ldots, x_n)$ of real numbers :

$$y = Aggreg\ (x_1, x_2, \ldots, x_n) \qquad (1)$$

Naturally, we should impose certain conditions on *Aggreg* to justify the name of "aggregation operator". Several authors [39], [43] have proposed a set of fundamental conditions defining the aggregation operators. It is to notice that these basic definitions are not compatible.
Recently, Mesiar and Komorníková [41] proposed a set of fundamental properties that group all the precedent definitions under weaker conditions. So, we define an aggregation operator as a function $Agreg : \bigcup_{n \in N} [0,1]^n \to [0,1]$ that satisfies :

- $Aggreg\ (x) = x$ — **Identity when unary**
- $Aggreg\ (0,\ldots,0) = 0$ and $Aggreg\ (1,\ldots,1) = 1$ — **Boundary conditions**
- $Aggreg\ (x_1,\ldots, x_n) \leq Aggreg\ (y_1,\ldots, y_n)$ if $(x_1,\ldots, x_n) \leq (y_1,\ldots, y_n)$ — **Non decreasing**

These conditions seem to be recurrent in all other proposed definitions of an aggregation operator. All other properties may come in addition to this fundamental group. In the next section we present an overview of the properties we may expect from an aggregation operator.

# 1.2 Properties of an aggregation operator

We divided the set of properties into two families : the mathematical properties and the behavioral properties. For more details, see Fodor [26] and Grabisch [30].

## Mathematical properties

### 1.2.1 Boundary Conditions

Here we turn our attention to the behavior of the aggregator in the best and in the worst cases. We expect that an aggregation operator satisfies :

$$Aggreg\ (0,0,\ldots,0) = 0 \tag{2}$$

$$Aggreg\ (1,1,\ldots,1) = 1 \tag{3}$$

Condition (2) means that if we observe only completely bad, false or not satisfactory criteria the total aggregation has to be also completely bad, false or not satisfactory. We have that (3) translates that if we observe only true or completely satisfactory criteria then the total aggregation has to be also completely true or satisfactory.
As Mesiar and Komorníková pointed out in [41], this property seems to be fundamental in the definition of aggregation operators.
Extension of this basic condition has been proposed. For example Mayor and Trillas [39] propose as a fundamental condition for an aggregation operator the following :

$$\forall x \in [0,1] \quad Aggreg\ (x,0) = Aggreg\ (1,0) \cdot x \tag{4}$$

$$\forall x \in [0,1] \quad Aggreg\ (x,1) = (1 - Aggreg\ (1,0)) \cdot x + Aggreg\ (1,0) \tag{5}$$

We notice that (4) requires that the value $Aggreg\ (x,0)$ is the weighted arithmetic mean of $x$ and 0 ; in the same way $Aggreg\ (x,1)$ is the weighted arithmetic mean of $x$ and 1 (see (5)). These two conditions constrain a little bit more the group of aggregation operators. In fact (2) and (3) are particular cases for $x$=0 and x=1 respectively for (4) and (5).

### 1.2.2 Monotonicity (non decreasing)

We deal more precisely with a non-decreasingness with respect to each variable. We expect that if an argument increases then the final aggregation increases (or at least not decreases, remaining equal) :

$$y_i \geq x_i \Rightarrow Aggreg(x_1,\ldots,y_i,\ldots,x_n) \geq Aggreg(x_1,\ldots,x_i,\ldots,x_n) \tag{6}$$

Strict monotonicity = cancellativity

### 1.2.3 Continuity

The function *Aggreg* is continuous with respect to each of its variables. This property is a guaranty for certain robustness, for a certain consistency and for a non chaotic behavior.

### 1.2.4 Associativity

An interesting property is to be able to aggregate by packages. We expect that the choice of the packages has no influence on the result. For three arguments the property can be written :

$$\begin{aligned} Aggreg(x_1, x_2, x_3) &= Aggreg(Aggreg(x_1, x_2), x_3) \\ &= Aggreg(x_1, Aggreg(x_2, x_3)) \end{aligned} \quad (7)$$

This property can also be useful if the operator is defined only for two elements, then the Associativity allows extending the definition to $n$ arguments without ambiguity.

### 1.2.5 Symmetry

Also known as commutativity or anonymity : The order of the arguments has no influence on the result. This property is compulsory when the aggregation is made of arguments having the same importance or arises from anonymous experts or sources.
For every permutation $\sigma$ of $\{1,2, \dots , n\}$ the operator satisfies :

$$Aggreg\left(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}\right) = Aggreg\left(x_1, x_2, \dots, x_n\right) \quad (8)$$

### 1.2.6 Bisymmetry

Bisymmetry is a property associated to the aggregation of $n^2$ inputs for a n-ary operators. If we write these inputs in a square matrix, then the bisymmetry translates the fact that it does not matter whether you first aggregate the column vectors and then the outputs of thereof, or vice versa, first you aggregate row vectors and then the relevant outputs. For a binary operator $A$ this means for all $x_{11,}$ $x_{12,}$ $x_{21,}$ $x_{22,}$ that :

$$A(A(x_{11}, x_{12,}), A(x_{21}, x_{22,})) = A(A(x_{11}, x_{21,}), A(x_{12}, x_{22,})) \quad (9)$$

Note : If an operator is commutative and associative then it is necessarily bisymmetric, however, neither commutativity nor associativity is implied by bisymmetry.

### 1.2.7 Absorbent Element

If the aggregation operator has an absorbent element $a$, then it can be used like an eliminating score or like a veto (it can be also considered as a qualifying score) :

$$Aggreg\left(x_1, \dots, a, \dots, x_n\right) = a \quad (10)$$

This element is also called annihilator.

### 1.2.8 Neutral Element

If the operator of aggregation has a neutral element $e$, then it can be used to be associated to an argument that should not have any influence on the aggregation :

$$Aggreg^{[n]}(x_1,\ldots,e,\ldots,x_{n-1}) = Aggreg^{[n-1]}(x_1,\ldots,x_{n-1}) \qquad (11)$$

### 1.2.9 Idempotence

Also known as unanimity or agreement : If we aggregate *n* times the same value, we expect to find the initial value :

$$Aggreg(x, x,\ldots, x) = x \qquad (12)$$

This property and the reinforcement property are incompatible.

### 1.2.10 Compensation

Also known as Pareto property. Here we expect that the result of the aggregation is lower than the highest element aggregated (the maximum) and higher than the lowest one (the minimum) :

$$\min_{i=1}^{n}(x_i) \leq Aggreg(x_1, x_2,\ldots, x_n) \leq \max_{i=1}^{n}(x_i) \qquad (13)$$

This property is not to be confused with the counterbalance property.

### 1.2.11 Counterbalancement

Also called by some authors compensation. This means that some confusion may appear with the previous property.

We call the counterbalance property, the behavior of an operator that decreases the final result if there are arguments that go into an opposite direction.

$$\forall t \in \left]0,1\right[,\ \forall (x_1,\ldots x_n)\ \exists (y_1,\ldots y_m) \text{ so that } Aggreg(x_1,\ldots x_n, y_1,\ldots y_m) = t \qquad (14)$$

### 1.2.12 Reinforcement

One characteristic of many types of human information processing, which was strikingly pointed out by is what Yager and Rybalov [51] called **full reinforcement**. By this property we mean to indicate the tendency, on the one hand, of a collection of high scores to reinforce each other to give a resulting score more affirmative then any of the individual scores alone and on the other hand the tendency of a collection of low scores to reinforce each other to give a resulting score more "disfirmative" than any of the individual scores. The first concept is called upward reinforcement and the second concept is called downward reinforcement. Yager shows [51] that the t-norms have only a downward reinforcement behavior, while the t-conorms have only an upward reinforcement behavior. He also shows that the uninorms have a full reinforcement behavior.

This property can be very interesting. For example, in medical diagnosis the appearance of a number of symptoms indicative of a disease will make us more confident in diagnosing a patient as having the disease than any symptoms alone while the lack of appearance of this symptoms will make us more confident diagnosing a patient as not having the disease.

### 1.2.13 Stability for a linear function

This property translates a stability of the operator for a change of measurement scale :

$$Aggreg\left(r\cdot x_1+t, r\cdot x_2+t, \ldots, r\cdot x_n+t\right)=r\cdot\left(Aggreg\left(x_1, x_2, \ldots, x_n\right)\right)+t \qquad (15)$$

When $r \geq 0$ then we speak of stability for positive linear transformation.
A widely study particular case is the self-duality (see [47] and [16]). It corresponds to the stability for the linear function where $r$ = -1 and $t$ = 1.

### 1.2.14 Invariance

When aggregating numbers ($x_1, x_2, \ldots, x_n$) represent measurement of certain criteria, we should specify a scale in which these measurements were performed. Moreover, we may want the aggregation function *Aggreg* to respect a meaningful relation with respect to the given scale. The notion of meaningfulness is formalized in the representational theory of measurement [37] as the invariance property :
For any admissible transformation *f*, we have :

$$Aggreg\left(f(x_1), f(x_2), \ldots, f(x_n)\right)=f\left(Aggreg\left(x_1, x_2, \ldots, x_n\right)\right) \qquad (16)$$

*Note :* the only aggregation operator totally invariant (invariant for any bijection $f$) is the projection.

## Behavioral properties

### 1.2.15 Decisional behavior

It is useful to have the possibility to express the behavior of the decision-maker. For example : tolerant, optimistic, pessimistic or strict. These behaviors are in multi-criteria usually named **disjunctive** and **conjunctive** behaviors.

### 1.2.16 Interpretability of the parameters

It is to be hoped that the parameters have almost evident semantic interpretation. This property forbids the use of a black box methodology.

### 1.2.17 Weights on the arguments

It is crucial to have the possibility to express weights on the arguments. This can be understood as privileging some of them.

# PART 2 :
# The Aggregation Operators

In this second part, we present an overview of the existing mathematical operators. We explain their main properties and particularities. We also present some notable particular cases.

We start by presenting some of most often used aggregation operators. We call them the basic ones. In this group we find the prototype of an aggregation operator, *the average*, but also we find the median, the minimum and the maximum, as well as some classical generalizations like the weighted mean and the k-order statistics.

We proceed by presenting the quasi-arithmetic means, a large useful family built on a transformation of average operator. Then we continue by presenting a generalization of the weighted mean, the ordered weighted average (OWA), which in an additive form includes the minimum and the maximum as particular cases. This leads us to the discrete fuzzy integrals : Choquet and Sugeno. The Choquet integral generalizes "additive" operators as the OWA or the weighted mean, while the Sugeno generalizes "min-max" operators. All these operators give a representative value "in the middle" of the aggregated set.

After the precedent, we present two families specialized on the aggregation under uncertainty : the t-norms and the t-conorms. These operators do not look for giving a "middle value", but instead they compute the intersection and union (respectively) of fuzzy sets. These operators are often used, since they can also be seen as a generalization of the logical aggregation operators : AND (t-norms) and OR (t-conorms).

Some research works revealed that human do not aggregate "logically" as the t-norms and t-conorms do. These works stressed on the fact that the operators classically used do not compensate "low" with "high" values. We present some of the proposed solutions that were based on t-norms and t-conorms : the compensatory operators.

Other kind of operators appeared when relaxing the axiom that differentiates the t-norm and t-conorm : the uninorms. These operators solve another problem of the t-norms and t-conorms, which is the lack of full (downwards and upwards) reinforcement.

We try here to present an objective overview of the domain, by presenting the characteristics, the advantages and disadvantages of each operator. A very good overview is also available in the form of a book [5] edited by Bouchon-Meunier. See also [] by Detyniecki, were you can find new interesting tools for aggregation of uncertain information and its relationship to the application.

## 2.1 Basic Operators

### 2.1.1 The arithmetic mean

The simplest and most common way to aggregate is to use a simple arithmetic mean (also know as the average). Mathematically we have :

$$\begin{aligned} M(x_1, x_2, ..., x_n) &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ &= \sum_{i=1}^{n}\left(\frac{1}{n}\cdot x_i\right) \end{aligned} \qquad (17)$$

This operator is interesting because it gives an aggregated value that is smaller than the greatest argument and bigger than the smallest one. So, the resulting aggregation is "a middle value". This property is known as the compensation property (see section 1.2.10). The average is often used since it is simple and satisfies the properties of monotonicity, continuity, symmetry, associativity, idempotence and stability for linear transformations. But it has neither absorbent nor neutral element and has no behavioral properties.

### 2.1.2 The weighted mean

There exists a classical extension, the weighted mean, which allows placing weights on the arguments. But we loose the property of symmetry. It is expressed mathematically by :

$$M_{w_1,\ldots,w_n}(x_1, x_2, ..., x_n) = \sum_{i=1}^{n}(w_i \cdot x_i) \qquad (18)$$

Where the weights are non negative and $\sum_{i=1}^{n} w_i = 1$.

### 2.1.3 The median

Another operator that follows the idea obtaining "a middle value" is the median. It consists in ordering the arguments from the smallest one to the biggest one. And then taking the element in the middle. If the cardinality of the set of arguments is not odd then there is not a middle argument but a pair. We take then the mean of the middle pair.
This aggregation operator satisfies the boundary conditions, the monotonicity, the symmetry, the idempotence and evidently the compensation behavior.
There exists a generalization of this operator : **the k-order statistic**, with which we can choose the element on the $k^{th}$ position on the ordered list (from the smallest to the biggest element). Recent works present even more general median-based operators, see Calvo and Mesiar [6].

### 2.1.4 The minimum and the maximum

The minimum and the maximum are also basic aggregation operators. The minimum gives the smallest value of a set, while the maximum gives the greatest one.
They do not give a representative "middle value", but they can be very meaning full in different contexts. For instance in a decision making context the minimum operator translates a conjunctive attitude (note that it is a particular t-norm (see 2.7)). And the maximum is a t-conorm (see 2.7) and has a disjunctive behavior.
As aggregation operators they satisfy the axioms of the definition (identity when unary, boundary conditions, non decreasing. Besides this basic properties this two operators are interesting because they are monotone, symmetric, associative, idempotent. Mathematically speaking they have a compensation behavior, but these are the limit cases. As we already said, using these operators we will never obtain an aggregated value "in the middle". For this reason, we do not consider that we can talk about compensation behavior in this case.
If we work in a restricted interval [*a*,*b*] the minimum has for absorbent element *a* and for neutral element *b*, while for the maximum it will be the opposite : *a* will be the neutral element and *b* the absorbent one.

### 2.1.5 The weighted minimum and the weighted maximum

It may be interesting to introduce weights for these basic operators (the same way as for the simple mean). But here there is no standard solution. For instance Yager [53] (see also Dubois and Prade [19]) proposed the following weighted minimum and weighted maximum, where the weights are non negative and $\max_{i=1}^{n}(w_i) = 1$ :

$$\min^{\oplus}_{w_1,\ldots,w_n}(x_1, x_2, \ldots, x_n) = \min_{i=1}^{n}\left[\max\left(1 - w_i, x_i\right)\right] \quad (19)$$

$$\max^{\oplus}_{w_1,\ldots,w_n}(x_1, x_2, \ldots, x_n) = \max_{i=1}^{n}\left[\min\left(w_i, x_i\right)\right] \quad (20)$$

From the weighting point of view, these two operators have interesting properties. For instance, if a weight $w_i$ equals zero then the argument $x_i$ will not be taken into account in the aggregation. Also if all weights are equal, then we obtain the minimum and the maximum respectively.
These operators have been criticized because it is possible to increase one of the weights (i.e. the importance of an argument) without having any change in the result. This is translated mathematically by the fact that these operators are not *strictly* monotone with respect to the weights (they are of course monotone).

Fagin and Wimmers [25] proposed a general method for incorporating weights to any aggregation operators. With their method we obtain the following for the weighted minimum and weighted maximum respectively :

$$\min^{\otimes}_{w_1,\ldots,w_n}(x_1, x_2, ..., x_n) = \sum_{i=1}^{n}\left[ i \cdot (w_{\sigma(i)} - w_{\sigma(i+1)}) \cdot \min(x_{\sigma(1)}, \ldots, x_{\sigma(i)}) \right] \qquad (21)$$

$$\max^{\otimes}_{w_1,\ldots,w_n}(x_1, x_2, ..., x_n) = \sum_{i=1}^{n}\left[ i \cdot (w_{\sigma(i)} - w_{\sigma(i+1)}) \cdot \max(x_{\sigma(1)}, \ldots, x_{\sigma(i)}) \right] \qquad (22)$$

Where the weights are non negative, $\sum_{i=1}^{n} w_i = 1$, $\sigma$ is a permutation that orders the weights as follows $w_{\sigma(1)} \geq w_{\sigma(2)} \geq \cdots \geq w_{\sigma(n)}$ and $w_{\sigma(n+1)} = 0$.
These operators verify the same properties than the announced one for the precedent two, but they are strictly monotone with respect to the weights. In other words, a variation on any of the weights produces a change in the result. The authors note that these operators are stable for any positive linear transformation (see 1.2.13), the same way as the minimum and maximum. Note that this is not the case for (19) or (20).

## 2.2 Quasi-arithmetic means

Many extensions of the simple arithmetic mean have been introduced such as the geometric mean :

$$M_{geometric}(x_1, x_2, \ldots, x_n) = \left( \prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \tag{23}$$

and the harmonic mean :

$$M_{harmonic}(x_1, x_2, \ldots, x_n) = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} \tag{24}$$

In fact all these common means belong to the family of the quasi-arithmetic means. This family has been studied in detail by Kolmogorov [36] and by Aczel [1],[2] and is defined as follows :

$$\begin{aligned} M_f(x_1, x_2, \ldots, x_n) &= f^{-1}\left[ \frac{1}{n} \sum_{i=1}^{n} f(x_i) \right] \\ &= f^{-1}\left[ \sum_{i=1}^{n} \left( \frac{1}{n} \cdot f(x_i) \right) \right] \end{aligned} \tag{25}$$

Where $f$ is a strictly monotone continuous function in the extended real line. We should notice that a generator $f$ is not unique. Think for instance on the linear transformation of a generator : $f'(x) = a \cdot f(x) + b$ , where $a \neq 0$ .

We notice that the geometric mean (23) is the particular case of (25) with $f(x) = \log x$ and the harmonic mean (24) is the particular case of (25) with $f(x) = \frac{1}{x}$ .

A particular attention should be taken in the case, when there exist arguments $x_i$ and $x_j$ that have for images $f(x_i) = -\infty$ and $f(x_j) = +\infty$. In this case the convention for the addition of minus infinity and plus infinity should be specified.

A very notable particular case, studied in detail by Dujmovic [23] and by Dyckhoff [24] corresponds to the function $f$ is defined by $f : x \to x^{\alpha}$ . We obtain then a quasi arithmetic mean of the form :

$$\begin{aligned} M(x_1, x_2, \ldots, x_n) &= \left[ \frac{1}{n} \sum_{i=1}^{n} x_i^{\alpha} \right]^{\frac{1}{\alpha}} \\ &= \left[ \sum_{i=1}^{n} \left( \frac{1}{n} \cdot x_i^{\alpha} \right) \right]^{\frac{1}{\alpha}} \end{aligned} \tag{26}$$

This family is particularly interesting, because it generalizes a group of common means, only by changing the value of $\alpha$:

- For $\alpha = 1$ we obtain the arithmetic mean
- For $\alpha = 2$ we obtain the quadratic mean (also called the Euclidean mean)
- For $\alpha = -1$ we obtain the harmonic mean
- When $\alpha$ tends to $-\infty$, formula (26) tends to the maximum operator.
- When $\alpha$ tends to $+\infty$, formula (26) tends to the minimum operator.
- When $\alpha$ tends to 0, formula (26) tends to the geometric mean.

## 2.3 Symmetric Sum

We call a symmetric sum a continuous self-dual aggregation operator *S*. We recall the self-duality is defined by :

$$S(x_1, x_2, \ldots, x_n) = 1 - S(1-x_1, 1-x_2, \ldots, 1-x_n) \qquad (27)$$

This operator was studied in detail by Silvert [47]. In particular he showed that the symmetric sum of two arguments can be written under the form :

$$S(x, y) = \frac{G(x, y)}{G(x, y) + G(1-x, 1-y)} \qquad (28)$$

Where *G* is a continuous, increasing, positive function satisfying *G*(0,0)=0. It is to notice there is not a unique function *G* characterizing each symmetric sum. It is also important to notice that we use the convention $\frac{0}{0} = \frac{1}{2}$.

We remark that symmetric sums are in general not symmetric or commutative. A good example of symmetric sums is the weighted mean (18).

An interesting particular case is the additive generated aggregation :

$$S_f(x_1, x_2, \ldots, x_n) = f^{-1}\left[\sum_{i=1}^{n} f(x_i)\right] \qquad (29)$$

Where the generator *f* is a strictly monotone continuous function in the extended real line and satisfying :

$$f(x) + f(1-x) = 0 \qquad (30)$$

In this case, we obtain an associative symmetric sum. If the range of *f* is [-∞,+∞], then we obtain the associativity on $[0,1]^2 \setminus \{(0,1),(1,0)\}$. In other words it is associative as long as we do not aggregate the values 0 and 1.

## 2.4 Ordered Weighted Averaging Operators

The Ordered Weighted Averaging Operators (OWA) were originally introduced by Yager in [55] to provide a means for aggregating scores associated with the satisfaction of multiple criteria, which unifies in one operator the conjunctive and disjunctive behavior :

$$OWA(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} w_j x_{\sigma(j)} \qquad (31)$$

Where $\sigma$ is a permutation that orders the elements : $x_{\sigma(1)} \le x_{\sigma(2)} \le \cdots \le x_{\sigma(n)}$. The weights are all non negative ($w_i \ge 0$) and their sum equals one ($\sum_{i=1}^{n} w_i = 1$).

This operator has been proved to be very useful, because of his versatility, and it is the object of a book edited in 1997 by Yager and Kacprzyk [50].

The OWA operators provide a parameterized family of aggregation operators, which include many of the well-known operators such as the maximum, the minimum, the k-order statistics, the median and the arithmetic mean. In order to obtain these particular operators we should simply choose particular weights (see Table 1).

| | **OWA** |
|---|---|
| Minimum | $\begin{cases} w_1 = 1 \\ w_i = 0 \quad \text{if } i \ne 1 \end{cases}$ |
| Maximum | $\begin{cases} w_n = 1 \\ w_i = 0 \quad \text{if } i \ne n \end{cases}$ |
| Median | $\begin{cases} w_{\frac{n+1}{2}} = 1 & \text{if } n \text{ is odd} \\ w_{\frac{n}{2}} = \frac{1}{2} \text{ and } w_{\frac{n}{2}+1} = \frac{1}{2} & \text{if } n \text{ is even} \\ w_i = 0 & \text{else.} \end{cases}$ |
| k-order statistics | $\begin{cases} w_k = 1 \\ w_i = 0 \quad \text{if } i \ne k \end{cases}$ |
| Arithmetic mean | $w_i = \frac{1}{n} \quad \text{for } \forall i$ |

**Table 1.** Particular cases of OWA

The Ordered Weighted Averaging operators are commutative, monotone, idempotent, they are stable for positive linear transformations and they have a compensatory behavior. This last property translates the fact that the aggregation done by an OWA operator always is between the maximum and the minimum. Since this operator generalizes the minimum and the maximum, it can be seen as a parameterized way to go from the *min* to the *max*. In this context Yager introduced in [55] a degree of maxness (initially called orness), defined by :

$$maxness(w_1, w_2, \ldots, w_n) = \sum_{j=1}^{n} w_{n-j+1} \cdot \frac{n-j}{n-1} = \sum_{j=1}^{n} w_j \cdot \frac{j-1}{n-1} \qquad (32)$$

We see that for the minimum, we have that *maxness*(1,0,…,0)=0 and for the maximum *maxness*(0, …,0,1)=1.

Another operator was introduced by Yager [55] and used by O'Hagan in [42] to characterize a particular OWA. This degree describes the dispersion of the weights and it is based on the idea of entropy :

$$dispersion(w_1, w_2, \ldots, w_n) = -\sum_{j=1}^{n} w_j \cdot \ln(w_j) \qquad (33)$$

One issue of considerable interest related to the use of these operators is the development of an appropriate methodology for the derivation of the weights used in the OWA aggregation. Two of the main approaches used are the following :

The first one introduced by O'Hagan [42] makes use of the measure of maxness and the measure of dispersion. In this approach we only required that the user provides a value $\alpha \in [0,1]$ corresponding to the degree of maxness suited. The idea is to maximize the dispersion of the weights under the constraint of a fixed maxness. The following mathematical programming problem computes the weights, for a given $\alpha$ :

| |
|---|
| Maximize $-\sum_{j=1}^{n} w_j \cdot \ln(w_j)$ (dispersion) |
| Under the constraints :<br>- $maxness(w_1, w_2, \ldots, w_n) = \sum_{j=1}^{n} w_j \cdot \frac{j-1}{n-1} = \alpha$<br>- $\sum_{j=1}^{n} w_j = 1$<br>- $0 \le w_j \le 1$ |

**Table 2.** Mathematical program that computes OWA weights

The second approach makes use of the knowledge of a linguistic quantifier to guide the aggregation [56]. We are interested in regular increasing monotone quantifiers defined by :

- $Q(0) = 0$ and $Q(1) = 1$.
- if $x \leq y$ then $Q(x) \leq Q(y)$ .

These quantifiers translate notions like *most, almost all, many, at least half* and *some*.

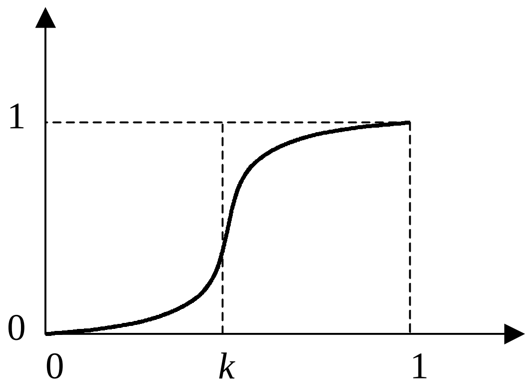


**Figure 1.** Regular increasing monotone quantifier "at least *k* %"

On the basis of this kind of quantifiers Yager in [55] proposed to compute the weights using the formula :

$$w_j = Q\left(\frac{n-j+1}{n}\right) - Q\left(\frac{n-j}{n}\right) \qquad (34)$$

Using this approach we can define goal functions under the constraint :

*Q criteria should be satisfied*

To illustrate this approach, let us consider one limit case. For instance if we want that "at least 100%" of the criteria to be satisfied, then we observe that the OWA operators is the *minimum.* And when the minimum is satisfied all other criteria are satisfied.

# 2.5 Choquet & Sugeno discrete Fuzzy Integrals

### 2.5.1 Definitions

The fuzzy integral with respect to a fuzzy measure has mainly been studied in a multicriteria decision making framework (see [30] and [28]). It is based on the notion of a fuzzy measure, which can be viewed as the weight of importance of a set. Mathematically we define the **fuzzy measure** as follows :
Let us denote by $C = \{c_1,\ldots, c_n\}$ the set of criteria, and $P(C)$ the power set of $C$, i.e. the set of all subsets of $C$. A **fuzzy measure** on $C$ is a set function $\mu : P(C) \to [0,1]$, satisfying the following axioms.

- $\mu(\varnothing) = 0$ and $\mu(C) = 1$. Boundary conditions
- for $A,B \in P(C)$, if $A \subset B$ then $\mu(A) \le \mu(B)$. Monotonicity

This kind of measure is more flexible than a probability, which is constrained by its additivity property. In fact, the importance of two criteria in the probability framework cannot be anything else than the sum of the individual importances, while the fuzzy measures can provide a greater (superadditive measure) or a lower (subadditive measure) value. This allows us to model interaction between criteria.
Now, using a fuzzy measure we introduce fuzzy integrals :
**The discrete Sugeno integral** [48] of scores $x_1,\ldots, x_n$ for criteria $c_1,\ldots, c_n$ with respect to a fuzzy measure $\mu$, is defined by :

$$Sugeno_\mu(x_1, x_2, \ldots, x_n) = \max_{i=1}^{n}\left(\min\left(x_{\sigma(i)}, \mu(C_{\sigma(i)})\right)\right) \qquad (35)$$

Where $\sigma$ is a permutation that orders the elements : $x_{\sigma(1)} \le x_{\sigma(2)} \le \cdots \le x_{\sigma(n)}$, where and $C_{\sigma(i)}=\{c_{\sigma(i)},\ldots, c_{\sigma(n)}\}$.
**The discrete Choquet integral** [8] of scores $x_1,\ldots, x_n$ for criteria $c_1,\ldots, c_n$ with respect to a fuzzy measure $\mu$, is defined by :

$$Choquet_\mu(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} (x_{\sigma(i)} - x_{\sigma(i-1)}) \cdot \mu(C_{\sigma(i)}) \qquad (36)$$

with the same notation as above, and $x_{\sigma(0)} = 0$.

An equivalent expression of (36) is

$$Choquet_\mu(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_{\sigma(i)} \cdot (\mu(C_{\sigma(i)}) - \mu(C_{\sigma(i+1)})) \qquad (37)$$

with $C_{\sigma(n+1)} = \varnothing$.

### 2.5.2 Properties

Sugeno and Choquet integrals are interesting since they are monotone, continuous, idempotent operators, with a compensation behavior. The Choquet integral is stable under positive linear transformation, while the Sugeno integral is stable under a similar transformation with minimum and maximum replacing the product and the sum respectively. This last property points out that the Sugeno integral is more suitable for ordinal aggregation (where only the order of the elements is important) while the Choquet integral is suitable for cardinal aggregation (where the distance between the numbers has a meaning).
The commutativity is only obtained when the fuzzy measure just depends on the cardinality of the sets, i.e. $\mu(A) = \mu(B)$ if card($A$) = card($B$). The associativity is usually not satisfied.

The generalization capability of the Choquet and Sugeno integrals is remarkable. Both contain the order statistics and in particular the minimum and the maximum. The Choquet integral generalizes the weighted means and the OWA operator, while the Sugeno integral generalizes the weighted minimum (see formula (20)) and the weighted maximum (see formula (21)). In the following table we present the corresponding measures in order to get a particular operator (for details see [28],[30],[48]).

| | **Sugeno integral** |
|---|---|
| Minimum | $\begin{cases}\mu(A)=1 & \text{if } A=C\\ \mu(A)=0 & \text{otherwise}\end{cases}$ |
| Maximum | $\begin{cases}\mu(A)=0 & \text{if } A=\varnothing\\ \mu(A)=1 & \text{otherwise}\end{cases}$ |
| k-order statistics | $\begin{cases}\mu(A)=0 & \text{if } card(A)\le n-k\\ \mu(A)=1 & \text{otherwise}\end{cases}$ |
| Weighted minimum | $\mu(A)=1-\max_{x_i\notin A}\left[\mu(\{x_i\})\right]$ and $\mu(\{x_i\})=w_i$ for all $i$ |
| Weighted maximum | $\mu(A)=\max_{x_i\in A}\left[\mu(\{x_i\})\right]$ and $\mu(\{x_i\})=w_i$ for all $i$ |

**Table 3.** Particular cases of the Sugeno Integral

| | Choquet integral |
|---|---|
| Minimum | $\begin{cases}\mu(A)=1 & \text{if } A=C\\ \mu(A)=0 & \text{otherwise}\end{cases}$ |
| Maximum | $\begin{cases}\mu(A)=0 & \text{if } A=\emptyset\\ \mu(A)=1 & \text{otherwise}\end{cases}$ |
| k-order statistics | $\begin{cases}\mu(A)=0 & \text{if } card(A)\le n-k\\ \mu(A)=1 & \text{otherwise}\end{cases}$ |
| Arithmetic mean | $\mu(A)=\dfrac{card(A)}{card(C)}$ |
| Weighted mean | $\mu(A)=\sum_{x_i\in A}\mu(\{x_i\})$ and $\mu(\{x_i\})=w_i$ for all $i$ |
| OWA | $\mu(A)=\sum_{j=0}^{card(A)-1} w_{n-j}$ |

**Table 4.** Particular cases of the Choquet Integral

The main problem in the use of Choquet or Sugeno integral (besides the fact that they are not associative and commutative) is the number $2^n$ of weights to be provided, for a simple $n$ criteria aggregation. These weights being nothing else than the characterization of the fuzzy measure. A main aspect of the actual research is based on the methods to determine or reduce the number of these weights. Some solutions have been proposed :
A first approach is to work on the measure, by defining (for instance) decomposable measures. An interesting approach was proposed by Grabisch in [29], where he suggests to use $k$-additive fuzzy measures. The idea is to define measures that are multilinear of degree $k$, i.e. if $card(A) > k$ then $m(A) = 0$. This approach allows to model the strength of small coalitions and reduces the number of weights to $\sum_{i=1}^{k} C_n^i$ instead of $2^n$. The case of 2-additive measure has been pointed out as particularly interesting. The number of weights to establish is $\dfrac{n^2+n}{2}$, the measure being defined by :

$$\mu(A)=\sum_{\{x_i,x_j\}\subset A}\mu(x_i,x_j)-(card(A)-2)\cdot\sum_{\{x_i\}\in A}\mu(x_i) \qquad (38)$$

Another approach is to determine the weights by learning on examples. The data being a set of *n*-dimensional vectors : $(x_1^k, x_2^k, \ldots, x_n^k)$ and their corresponding aggregated values : $y_k$. Then we identify the fuzzy measure by minimizing the following error :

$$E^2 = \sum_{k=1}^{l} \left( Choquet_\mu (x_1^k, x_2^k, \ldots, x_n^k) - y_k \right)^2 \qquad (39)$$

It can be shown [30] that (39) can be put under a quadratic program form, that is

| |
|---|
| minimize $\frac{1}{2} \cdot u^t D u + c^t u$ |
| under the constraint $Au + b \geq 0$ |

**Table 5.** Quadratic program computing the weights for a Choquet Integral

Where *u* is the vector containing all the weights of the fuzzy measure.
It appears as an interesting solution, but if there is not enough data the matrices may be ill conditioned. In fact there must be at least $\frac{n!}{[(n/2)!]^2}$ training vectors.

# 2.6 Fusion Operators

This section gives an overview of the group of specific operators dealing with fusion. The authors of these operators were particularly aware of the problem of "reliability of the sources".

### 2.6.1 The Bayesian Approach

The idea in the Bayesian approach is to estimate the most probable "$x$" knowing that we observed "$x_1$" from source 1 and "$x_2$" from source 2. Mathematically we are looking for the maximum of the a posteriori probability $P(x|x_1, x_2)$.

The a posteriori probability can be derived using the Bayes theorem and knowing the value of the a priori probability $P(x_1, x_2|x)$ :

$$P(x|x_1, x_2) = \frac{P(x_1, x_2|x) \cdot P(x)}{P(x_1, x_2)} \qquad (40)$$

If we assume that the sources are independent, we obtain the most usual Bayesian fusion formula :

$$P(x|x_1, x_2) = \frac{P(x_1|x) \cdot P(x_2|x) \cdot P(x)}{P(x_1, x_2)} \qquad (41)$$

Note that the a priori probabilities $P(x_1|x)$ and $P(x_2|x)$ can be seen as the credibility of the source (expert). They actually translate the probability of source 1 (and source 2) to observe $x_1$ (and $x_2$), given the "real value" $x$.

The main default of this method is that the a priori probabilities are not easy to obtain. Also some critics come from the independence assumption. Some solutions are reviewed in [9].

### 2.6.2 Possibilistic approaches

The first idea is to modify directly the information provided by the source given its reliability. Let $\alpha$ be the degree of certainty that a given source is reliable, then Yager [53] and Prade [44] propose to modify the possibility distribution $\pi$ provided by the source, using the operator :

$$\pi' = \max(\pi, 1 - \alpha) \qquad (42)$$

When $\alpha$=1 (fully reliable source), $\pi'=\pi$, and when $\alpha$=0 (absolutely unreliable source), then $\forall x \ \pi'(x) = 1$ (total ignorance). Note that 0 does not mean that the source lies, but that it is impossible to known whether its advice is good or not.

There exist other proposals for certainty qualification that behave similarly in the limit cases ($\alpha$=1 and $\alpha$=0). Yager [53] suggested an expression of the form :

$$\pi' = \pi * \alpha + 1 - \alpha \qquad (43)$$

where $*$ stands for minimum, product or linear product (max(0,a+b-1).

Another point of view is to consider that the reliability of the sources imply priorities in the aggregation. The idea of a prioritized fusion is to accept the conjunctive merging of information from a reliable source 1 and source 2 as long as the data coming from the second is consistent with the former. In case of inconsistency, the information given by the less reliable is simply discarded. If $\pi_1$ is obtained from source 1 and $\pi_2$ from source 2, the degree of consistency of $\pi_1$ and $\pi_2$ is defined by :

$$h(\pi_1, \pi_2) = \sup_x \big(\min(\pi_1(x), \pi_2(\mathrm{x}))\big) \qquad (44)$$

and the following **prioritized conjunction** has been proposed by Dubois and Prade [21] (see also Yager [54]) :

$$\pi_{1\wedge 2} = \min(\pi_1, \max(\pi_2, 1 - h(\pi_1, \pi_2))) \qquad (45)$$

Note that when $h(\pi_1, \pi_2)$ =0, source 1 contradicts source 2 and the only opinion of source 1 is retained (i.e. $\pi_1$), while if $h(\pi_1, \pi_2)$ =1 we have the minimum, which is a conjunction.

The **disjunctive** counterpart of this **prioritized** operator has been proposed by Dubois and Prade [22] :

$$\pi_{1\vee 2} = \max(\pi_1, \min(\pi_2, h(\pi_1, \pi_2))) \qquad (46)$$

The effect of this operator is to truncate the information supplied by the less prioritary source, while disjunctively combined with source 1. Again if the two sources disagree ( $h(\pi_1, \pi_2)$ =0) then we have $\pi_1$; if $h(\pi_1, \pi_2)$ =1 then we have the maximum of $\pi_1$ and $\pi_2$ .

A very complete overview of the possibilistic fusion is offered in [20].

# 2.7 T-norms and t-conorms

The concept of a triangular norm was introduced by Menger [40] in order to generalize the triangular inequality of a metric. The current notion of a t-norm and its dual operator (t-conorm) is due to Schweizer and Sklar [46] [45]. Both of these operations can also be used as a generalization of the Boolean logic connectives to multi-valued logic. The t-norms generalize the conjunctive 'AND' operator and the t-conorms generalize the disjunctive 'OR' operator. This situation allows them to be used to define the intersection and union operations in fuzzy logic. This possibility was first noted by Hohle [31], Klement [34], Dubois and Prade [17] and Alsina, Trillas, and Valverde [3] very early appreciated the possibilities of this generalization. Bonissone [4] investigated the properties of these operators with the goal of using them in the development of intelligent systems. t-norm and t-conorms have been well-studied and very good overviews and classifications of these operators can be found in the literature, see [35],[18]. A particular complete work is presented in a book [32] explicitly dedicated to these operators.

### 2.7.1 Definitions

**t-norm :** A t-norm is a function $T : [0,1]\text{x}[0,1] \to [0,1]$, having the following properties

- $T(x,y) = T(y,x)$ (T1) **Commutativity**
- $T(x,y) \leq T(u,v)$, if $x \leq u$ and $y \leq v$ (T2) **Monotonicity (increasing)**
- $T(x,T(y,z)) = T(T(x,y),z)$ (T3) **Associativity**
- $T(x,1) = x$ (T4) **One as a neutral element**

A well known property of t-norms is :

- $T(x,y) \leq \min(x,y)$ (47)

**Proof :** Using the monotonicity (T2) and axiom T4, we have $T(x,y) \leq T(x,1) = x$ and using the commutativity we have $T(x,y) \leq T(1,y) = y$. So, $T(x,y) \leq \min(x,y)$.

**t-conorm :** Formally, a t-conorm is a function $S : [0,1]\text{x}[0,1] \to [0,1]$, having the following properties :

- $S(x,y) = S(y,x)$ (S1) **Commutativity**
- $S(x,y) \leq S(u,v)$, if $x \leq u$ and $y \leq v$ (S2) **Monotonicity (increasing)**
- $S(x,S(y,z)) = S(S(x,y),z)$ (S3) **Associativity**
- $S(x,0) = x$ (S4) **Zero as a neutral element**

A well known property of t-conorms is :

- $S(x,y) \geq \text{Max}(x,y)$ (48)

It is actually a consequence of axioms (S1, S2, S4).

### 2.7.2 Dual t-norms and t-conorms

We say that a t-norm and a t-conorm are dual (or associated) if they satisfy the DeMorgan law.

- $\overline{T(x,y)} = S\left(\bar{x}, \bar{y}\right)$ (the DeMorgan law)

Where the line over the expression means the negation of the expression. We will use the most typical negation defined by :

- $\bar{x} = 1 - x$ (negation)

### 2.7.3 Examples

The definitions of t-norms and t-conorms are always given for only two elements, since these operators are by definition associative and in this case the generalization to *n* elements is trivial. The more common t-norms and their dual t-conorms are :

| | **t-norm** | **t-conorm** |
|---|---|---|
| **Min-Max** | $\min(x, y)$ | $\max(x, y)$ |
| **Probabilistic** | $x \cdot y$ | $x + y - x \cdot y$ |
| **Lukasiewicz** | $\max(x + y - 1, 0)$ | $\min(x + y, 1)$ |
| **Drastic** | $\begin{cases} x & \text{if } y = 1 \\ y & \text{if } x = 1 \\ 0 & \text{anywhere else} \end{cases}$ | $\begin{cases} x & \text{if } y = 0 \\ y & \text{if } x = 0 \\ 1 & \text{anywhere else} \end{cases}$ |

**Table 6.** Common t-norms and their dual t-conorms

We would like to insist here in some important particularities of these t-norms and t-conorms :
The minimum is the biggest t-norm (i.e. when using the *min*, we obtain a higher value than when using any other t-norm). It is also the only idempotent t-norm. Its dual is also idempotent and it is the smallest t-conorm.
The probabilistic case has the nice property to be "smooth". This can be translated mathematically through a continuous derivative.
The Lukasiewicz t-norm satisfies the classical logical law of non-contradiction (i.e. $T\left(x, \bar{x}\right) = 0$). And its dual the Lukasiewicz t-conorm satisfies the classical logical law of the excluded middle (i.e. $S\left(x, \bar{x}\right) = 1$).
The drastic case is interesting only from the mathematical point of view. These operators give the values 0 or 1 except when they are forced by the axioms. They are not

continuous. The main interesting aspect is that the drastic t-norm is the smallest t-norm and the drastic t-conorm is the biggest t-conorm.
A main result is that we can construct any continuous t-norm by using the precedent t-norms and the equivalent result exists for the continuous t-conorms.

### 2.7.4 Parameterized t-no rms and t-conorms

We have also parameterized t-norms and t-conorms. As special cases we obtain some of the precedent t-norms and t-conorms :

| | t-norm | t-conorm |
|---|---|---|
| **Hamacher ($\gamma \geq 0$)** | $\frac{x \cdot y}{\gamma + (1-\gamma) \cdot (x + y - x \cdot y)}$ | $\frac{x + y - x \cdot y - (1-\gamma) \cdot x \cdot y}{1 - (1-\gamma) \cdot x \cdot y}$ |
| **Yager ($p$ > 0)** | $\max\left(1 - \left[(1-x)^p + (1-y)^p\right]^{1/p} , \; 0\right)$ | $\min\left(\left[x^p + y^p\right]^{1/p} , \; 1\right)$ |
| **Weber-Sugeno ($\lambda_T , \lambda_S$ >-1)** | $\max\left(\frac{x + y - 1 + \lambda_T \cdot x \cdot y}{1 + \lambda_T} , \; 0\right)$ | $\min\left(x + y + \lambda_S \cdot x \cdot y , \; 1\right)$ |
| **Schweizer & Sklar (q>0)** | $1 - \left[(1-x)^q + (1-y)^q - (1-x)^q (1-y)^q\right]^{\frac{1}{q}}$ | $\left[x^q + y^q - x^q y^q\right]^{\frac{1}{q}}$ |
| **Frank ($s$ > 0, $s \neq 1$)** | $\log_s\left[1 + \frac{(s^x - 1) \cdot (s^y - 1)}{s - 1}\right]$ | $1 - \log_s\left[1 + \frac{(s^{1-x} - 1) \cdot (s^{1-y} - 1)}{s - 1}\right]$ |

**Table 7.** Parameterized t-norms and t-conorms.

It is to notice that the precedent t-norms and t-conorms are dual, besides in the Weber-Sugeno case. In this last case, the duality is satisfied if the parameters satisfy $\lambda_S = \frac{\lambda_T}{1 + \lambda_T}$ .

### 2.7.5 The Archimedean t-n orms and t-conorms

A t-norm $T$ is called Archimedean if for each $(x,y) \in ]0,1[^2$ there is an number $n$ so that :

$$T(\underbrace{x, \ldots, x}_{n-\text{times}}) < y \qquad (49)$$

The subset of continuous Archimedean t-norms (and t-conorms) is particularly interesting because they can be represented by means of a single function that we will call the additive generator. It can be shown that for every continuous Archimedean t-norm $T$, there exists a continuous decreasing function $f$ such that :

$$T(x_1,\ldots,x_n) = f^{(-1)}\left(\sum_{i=1}^{n} f(x_i)\right) \qquad (50)$$

with $f:[0,1] \to [0,+\infty]$ satisfying $f(1)=0$ and $f^{(-1)}$ is the pseudo inverse of $f$, defined by :

$$f^{(-1)}(z) = \begin{cases} f^{-1}(z) & \text{if } z \in [0, f(0)] \\ 0 & \text{if } z \in \,]f(0),+\infty] \end{cases} \qquad (51)$$

An equivalent theorem exists for the t-conorms.
In Table 7 we present the most common continuous Archimedean t-norms and t-conorms and the corresponding additive generators. We present two simple t-norms and t-conorms and two parameterized families. For a complete overview see [35].

| | | **Usual representation** | **Additive generator:** $f(u)$ |
|---|---|---|---|
| **Probabilistic** | t-norm | $x \cdot y$ | $-\ln(u)$ |
| | t-conorm | $x+y-x\cdot y$ | $-\ln(1-u)$ |
| **Lukasiewicz** | t-norm | $\max(x+y-1,0)$ | $1-u$ |
| | t-conorm | $\min(x+y,1)$ | $u$ |
| **Hamacher** | t-norm | $\dfrac{x\cdot y}{\gamma+(1-\gamma)\cdot(x+y-x\cdot y)}$ | for $\gamma>0$ : $-\dfrac{1}{\gamma}\cdot\ln\left(\dfrac{u}{\gamma+(1-\gamma)\cdot u}\right)$ |
| | | | for $\gamma=0$ : $\dfrac{1-u}{u}$ |
| | t-conorm | $\dfrac{x+y-x\cdot y-(1-\gamma)\cdot x\cdot y}{1-(1-\gamma)\cdot x\cdot y}$ | for $\gamma>0$ : $-\dfrac{1}{\gamma}\cdot\ln\left(\dfrac{1-u}{\gamma+(1-\gamma)\cdot(1-u)}\right)$ |
| | | | for $\gamma=0$ : $\dfrac{u}{1-u}$ |
| **Yager** | t-norm | $\max\left(1-\left[(1-x)^p+(1-y)^p\right]^{1/p},\ 0\right)$ | $(1-u)^p$ |
| | t-conorm | $\min\left(\left[x^p+y^p\right]^{1/p},\ 1\right)$ | $u^p$ |

**Table 8.** T-norms and t-conorms and their additive generators.

**Note :** the minimum and the maximum are not Archimedean, but they can be limit cases of Archimedean parameterized cases.

## 2.8 Compensatory Operators

Several authors noticed that t-norms and t-conorms lack of compensation behavior and that this particular property seems crucial in the aggregation process. One of the first authors to notice this were Zimmermann and Zysno [57]. They discover that in a decision making context humans do not follow exactly the behavior of a t-norm (nor of a t-conorm) when aggregating. In order to get closer to the human aggregation process, they proposed an operator on the unit interval based on t-norms and t-conorms :

$$Z_\gamma(x_1,\ldots,x_n)=\left(\prod_{i=1}^{n} x_i\right)^{1-\gamma}\cdot\left(1-\prod_{i=1}^{n}(1-x_i)\right)^{\gamma} \quad (52)$$

Here the parameter $\gamma$ indicates the degree of compensation. This operator is a particular case of the **exponential compensatory operators** [49] :

$$E_{\gamma}^{T,S}(x_1,\ldots,x_n)=(T(x_1,\ldots,x_n))^{1-\gamma}\cdot(S(x_1,\ldots,x_n))^{\gamma} \quad (53)$$

Where $T$ is a t-norm and $S$ a t-conorm.
It is important to notice that the exponential compensatory operators are not associative for $\gamma$ different from 0 or 1.
Another class of non-associative t-norm and t-conorm-based compensatory operator is the **convex-linear compensatory operator** [49], [38] :

$$L_{\gamma}^{T,S}(x_1,\ldots,x_n)=(1-\gamma)\cdot T(x_1,\ldots,x_n)+\gamma\cdot S(x_1,\ldots,x_n) \quad (54)$$

Setting the value of the parameter $\gamma$ is a delicate issue. Zimmerman and Zysno calculated the best $\gamma$ to match the human behavior. Yager and Rybalov proposed in [51] a method based on fuzzy modeling techniques to compute the parameter $\gamma$ :

$$\gamma=\frac{T(x_1,\ldots,x_n)}{T(x_1,\ldots,x_n)+T(1-x_1,\ldots,1-x_n)} \quad (55)$$

Where $T(x_1,\ldots,x_n)$ is called the highness and $T(1-x_1,\ldots,1-x_n)$ the lowness.
Another approach to the construction of compensatory operators based on t-norms and t-conorms was taken by Klement, Mesiar and Pap [33]. They based their construction on the additive generators of continuous Archimedean t-norms and t-conorms. Their **associative compensatory operator** is defined by :

$$C(x,y)=f^{-1}(f(x)+f(y)) \quad (56)$$

Where the function $f$ is defined by :

$$f(x) = \begin{cases} -g\left(\dfrac{x}{e}\right) & if \quad x \leq e \\ h\left(\dfrac{x-e}{1-e}\right) & if \quad x \geq e \end{cases} \tag{57}$$

Where $g$ is an additive generator of a t-norm, $h$ is an additive generator of a t-conorm and $e$ is a neutral element. It is to notice that this operator is a particular case of uninorms (see next section 2.9).

# 2.9 Uninorms

T-norms and t-conorms play a notable role in fuzzy logic theory, unfortunately these operators do not admit a compensating behavior. In fact t-norms do not allow low values to be compensated by high values and t-conorms do not allow high values to be compensated by low values (see [51]). For this reason Fodor, Yager and Rybalov introduced in [27] (see also [52]) the family of uniform aggregation operators (uninorm), as a generalization of both t-norm and t-conorm. This operator has a neutral element laying anywhere in the unit interval rather than at one or zero as for the t-norms and t-conorms respectively.

### 2.9.1 Definition

Formally, a uninorm is a function $U$: [0,1]x[0,1] $\rightarrow$ [0,1], having the following properties :

- $U(x,y) = U(y,x)$ (U1) **Commutativity**
- $U(x,y) \leq U(u,v)$, if $x \leq u$ and $y \leq v$ (U2) **Monotonicity (increasing)**
- $U(x,U(y,z)) = U(U(x,y),z)$ (U3) **Associativity**
- $\exists e \in [0,1] \quad \forall x \in [0,1] \; U(x,e) = x$ (U4) ***e* is the neutral element**

We see that the first three properties (U1, U2, U3) are common to uninorms, t-norms and t-conorms, but the fourth condition U4 is more general in the case of uninorms, in that it allows any value for the identity. These properties seem to be interesting for aggregation purposes. In fact, the commutativity translates the property of an operator to give the same result independently of the order of the values to be aggregated. The monotonicity translates the fact that if one of the aggregated element augments its value, then the aggregated value at least does not decrease. The associativity imposes to the operator the property of accepting the aggregation by groups. With an associative operator we can aggregate by groups and then aggregate all the groups and obtain the same result as when aggregating all the elements directly. The associativity is also interesting when aggregating new information, in that case we would not need to re-compute the aggregation with all the arguments, but simply aggregate the old calculated value with the new one. Finally, the neutral element is interesting, because it can be considered as the score that we would give to an argument, which should not have any influence in the aggregation. It is somehow a null vote.

### 2.9.2 Other Properties

One characteristic of many types of human information processing is what we shall call **full reinforcement** (see 1.2.12). Yager shows in [51] that the uninorms have a full reinforcement behavior, if the neutral element $e$ is different to zero or one. In fact, a uninorm having $e$=1 as a neutral element is a t-norm, and a t-conorm for $e$=0.
If we now take a look more precisely at the uninorms we discover that a uninorm behaves as a t-norm in the square $[0,e]^2$ and as a t-conorm on the square $[e,1]^2$. In fact, De Baets

and Fodor showed in [10] that to any uninorm with neutral element $e \in ]0,1[$, there corresponds a t-norm $T$ and a t-conorm $S$ such that :

$$\forall\ (x,y) \in [0,e]^2 \qquad U(x,y) = e \cdot T\left(\frac{x}{e}, \frac{y}{e}\right) \tag{58}$$

$$\forall\ (x,y) \in [e,1]^2 \qquad U(x,y) = e + (1-e) \cdot S\left(\frac{x-e}{1-e}, \frac{y-e}{1-e}\right) \tag{59}$$

Another interesting property is the **compensation behavior**. Neither the t-norms nor the t-conorms present a compensation behavior. De Baets and Fodor showed in [10] that on $[0,e[ \times\ ]e,1] \cup\ ]e,1] \times [0,e[$ any uninorm $U$ satisfies :

$$\min(x,y) \le U(x,y) \le \max(x,y), \tag{60}$$

In other words uninorms show partially a compensating behavior. This property is particularly interesting when we think about the fact that if we use a t-norm the occurrence of an input small positive value would mean that the result is small, no matter what the other inputs are. They can for instance all equal one. A uninorm will allow to compensate these low values with other high values. We notice that as for the t-norms, the t-conorms do not have a compensating behavior. The appearance of a high value (i.e. close to one) will not be compensated using a t-conorm.

### 2.9.3 Minimal and maximal uninorms

We first note that for $e = 1$ or $e = 0$, there exists a large class of such uninorms (corresponding to the t-norms and t-conorms respectively). However, for the purpose of finding full reinforcement operators we need uninorms with neutral element other than 1 or 0. In [27] Fodor, Yager and Rybalov introduced two general classes of uninorms for any $e$. We will call them the minimal uninorms and the maximal uninorms.

The **minimal uninorm** is the weakest uninorm $U$ given a t-norm $T$, a t-conorm $S$ and a neutral element $e$. This operator will be defined by :

$$U_{\min}(x,y) = \begin{cases} e \cdot T\left(\frac{x}{e}, \frac{y}{e}\right) & \textit{for } x \le e \textit{ and } y \le e \\ e + (1-e) \cdot S\left(\frac{x-e}{1-e}, \frac{y-e}{1-e}\right) & \textit{for } x \ge e \textit{ and } y \ge e \\ \min(x,y) & \textit{elsewhere} \end{cases} \tag{61}$$

The **maximal uninorm** is the strongest uninorm $U$ given a t-norm $T$, a t-conorm $S$ and a neutral element $e$. This operator will be defined by :

$$U_{\max}(x,y)=\begin{cases} e\cdot T\left(\frac{x}{e},\frac{y}{e}\right) & \textit{for } x\le e \textit{ and } y\le e \\ e+(1-e)\cdot S\left(\frac{x-e}{1-e},\frac{y-e}{1-e}\right) & \textit{for } x\ge e \textit{ and } y\ge e \\ \max(x,y) & \textit{elsewhere} \end{cases} \quad (62)$$

We observe that the two families satisfy all the properties announced before, besides the compensation behavior, because we have a min or a max operator elsewhere. Another great disadvantage of these uninorms is that they have discontinuities around the neutral element. In fact for the maximal uninorm for any $a < e$, we have the following :
when we approach with values smaller than the neutral element, we have :

$$U_{\max}\left(a, x\to e^{-}\right)=e\cdot T\left(\frac{a}{e},\frac{x}{e}\to 1\right)=e\cdot\frac{a}{e}=a \quad (63)$$

and when we approach with values greater than the neutral element we have :

$$U_{\max}\left(a, x\to e^{+}\right)=\max\left(a, x\to e^{+}\right)=e \quad (64)$$

Comparing (63) and (64) we see that the value of the uninorm springs from $a<e$ to $e$, when we go from a little bit smaller than $e$ to a little bit bigger than $e$.
In an analogous way, we have for the conjunctive uninorm for any $a > e$ :
when we approach with values greater than the neutral element :

$$U_{\min}\left(a, x\to e^{+}\right)=e+(1-e)\cdot S\left(\frac{a-e}{1-e},\frac{x-e}{1-e}\to 0\right)=e+(1-e)\cdot\frac{a-e}{1-e}=a \quad (65)$$

and when we approach with values smaller than the neutral element we have :

$$U_{\min}\left(a, x\to e^{-}\right)=\min\left(a, x\to e^{-}\right)=e \quad (66)$$

Comparing (65) and (66) we see that the value of the uninorm springs from $a > e$ to $e$, when we go from a little bit bigger than $e$ to a little bit smaller than $e$.

### 2.9.4 Generated uninorms

In [27] Fodor, Yager and Rybalov showed the following additive generated representation theorem, which gives an almost continuous uninorm :
Suppose $U$ is continuous on $[0,1]^2 \setminus \{(0,1),(1,0)\}$ with neutral element $e\in\,]0,1[$. Then there exists a strictly increasing continuous function $g:[0,1]\to[-\infty,+\infty]$, with $g(e)=0$ such that the representation

$$U(x,y) = g^{(-1)}(g(x) + g(y)) \tag{67}$$

holds if and only if the following two conditions are satisfied :

- *U* is strictly increasing on the open unit square.
- *U* is self-dual with respect to a strong negation *N* with fixed point *e*.

In this case $g(0)= -\infty$, $g(1)= +\infty$ and $g^{(-1)} = g^{-1}$.

We remark that generated uninorms were already introduced as an interesting class of aggregation operators by Klement, Mesiar and Pap in [32], and were called the **associative compensatory operator** (see compensatory operators in section 2.8). Also Dombi [16] arrived to the same construction when presenting his **aggregative operator**.

### 2.9.5 Nullnorms

Nullnorms were found as solutions of the Frank equation for uninorms [7] :

$$U(x,y) + N(x,y) = x + y \tag{68}$$

From this follows that a nullnorm *N* is a commutative, associative and increasing operator, with an absorbent element $a \in [0,1]$ and that satisfies $\forall x \in [0,a] \quad N(x,0) = x$ and $\forall x \in [a,1] \quad N(x,1) = x$.

It can be shown that a nullnorm can be written under the following form :

$$N(x,y) = \begin{cases} a \cdot S\left(\dfrac{x}{a}, \dfrac{y}{a}\right) & \textit{for } x \le a \textit{ and } y \le a \\ a + (1-a) \cdot T\left(\dfrac{x-a}{1-a}, \dfrac{y-a}{1-a}\right) & \textit{for } x \ge a \textit{ and } y \ge a \\ a \quad \textit{elsewhere} & \end{cases} \tag{69}$$

From this it is clear that this class contains t-norms (for *a*=0) and t-conorms (for *a*=1) as special cases.

# Reference s


[1]. Aczel, J., *Lectures on Functional Equations and Applications,* Academic Press : New York, 1966.

[2]. Aczel, J., On Mean Values, *Bulletin of the American Mathematical Society 54*, 392-400, 1948.

[3]. Alsina, C., Trillas, E. and Valverde, L., On some logical connectives for fuzzy set theory, *Journal of Mathematics Analysis & Applications 93*, 15-26, 1983.

[4]. Bonissone, P.P., Selecting uncertainty calculi and granularity : An experiment in trading off precision and complexity, *Proceedings of the first Workshop on Uncertainty in Artificial Intelligence,* Los Angeles, 57 - 66, 1985.

[5]. Bouchon-Meunier B., Ed., *Aggregation and Fusion of Imperfect Information*, Physica-Verlag, Heidelberg, Germany, 1997.

[6]. Calvo T. and Mesiar R., Generalized medians. *Proceedings of AGGREGATION'99, Palma de Mayorca*, 159-165, 1999.

[7]. Calvo T., De Baets B. and Fodor J., The functional equations of Frank and Alsina for uninorms and nullnorms, *Fuzzy Sets and Systems,* to appear.

[8]. Choquet, G., Theory of capacities. *Annales de l'Institut Fourier 5*, 131-295, 1953.

[9]. Cooke R. M., *Experts in Uncertainty*, Oxford Press, Oxford, UK, 1991.

[10]. De Baets B. and Fodor J., On the structure of uninorms and their residual implicators, *Proceedings of the Eighteenth Linz Seminar on Fuzzy Set Theory,* Linz, Austria, 81-87, 1997.

[11]. Detyniecki M., Yager R. and Bouchon-Meunier B., Specifying t-norms based on the value of T(1/2,1/2), *Mathware & Soft Computing vol.VII*, number 1, 2000.

[12]. Detyniecki M. and Bouchon-Meunier B., Aggregating Truth and Falsity Values, *FUSION'2000*, Paris, France, 18-24, July 2000.

[13]. Detyniecki M. and Bouchon-Meunier B., Building an Aggregation Operator with a Balance, *IPMU'2000*, Madrid, Spain, 686-692, July 2000.

[14]. Detyniecki M., Bouchon-Meunier B., and Yager R., Balance Operator: A new

vision on Aggregation Operators, *Proceedings of Eurofuse-Sic'99*, 241-246, Budapest, Hungary, May 1999.

[15]. Detyniecki, M., Mathematical aggregation operators and their application to video querying, *Doctoral thesis - research report 2001-002*, Laboratoire d'Informatique de Paris 6, 2000.

[16]. Dombi, J., Basic concepts for the theory of evaluation : The aggregative operator, European Journal of Operational Research 10, 282-293, 1982.

[17]. Dubois, D. and Prade H., Triangular norms for fuzzy sets, *Proceedings of the 2nd International Symposium of Fuzzy Sets*, Linz, 39 - 68, 1981.

[18]. Dubois, D. and Prade, H., A review of fuzzy sets aggregation connectives, *Information Sciences 36*, 85 - 121, 1985.

[19]. Dubois, D. and Prade, H., Weighted minimum and maximum operations in fuzzy set theory, *The Reliability of Expert Systems*, Hollnagel, E. (ed.), Ellis Horwood Limited : Chichester, 64-118, 1986.

[20]. Dubois, D. and Prade, H., *Possibility theory - An approach to the computerized Processing of Uncertainty*, Plenum press, New York, 1988.

[21]. Dubois, D. and Prade, H., Default reasoning and possibility theory, *Artificial Intelligence 35*, 243-257, 1988.

[22]. Dubois, D. and Prade, H., On the combination of uncertain or imprecise pieces of information in ruled based systems, *International Journal of approximate Reasoning 2*, 65-87, 1988.

[23]. Dujmovic, J. J., Evaluation, Comparison and Optimization of hybrid computers using the theory of complex criteria, *Simulation of Systems*, Dekker, L. (Eds.), North Holland : Amsterdam, 553-566, 1976.

[24]. Dyckhoff, H. and Pedrycz, W., Generalized means as model of compensative connectivities, *Fuzzy Sets and Systems 14*, 143-154, 1984.

[25]. Fagin, R. and Wimmers, E., A formula for incorporating weights into scoring rules, *Theoretical Computer 239*, 309-338, 2000.

[26]. Fodor, J. and Roubens, M., Fuzzy Preference Modeling and Multicriteria Decision Support, Kluwer Academic Publisher : Dordrecht, 1994.

[27]. Fodor, J.C., Yager, R.R. and Rybalov, A., Structure of Uninorms, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 5*, 411-427, 1997.

[28]. Grabisch, M., Fuzzy integrals in multicriteria decision making, *Fuzzy Sets and Systems 69*, 279-298, 1995.

[29]. Grabisch, M., k-additive fuzzy measures, *6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (*IPMU*), Granada, Spain, July 1996.

[30]. Grabisch, M., Nguyen, H.T. and Walker, E.A., Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference, *Kluwer Academics Publishers* : Dordrecht, 1995.

[31]. Hohle, U., Probabilistic uniformization of fuzzy topologies, *Fuzzy Sets and Systems 1*, 1978.

[32]. Klement E. P., Mesiar R. and Pap, Triangular norms, *Trends in Logic, Studia Logica Library, Vol.8*, Kluwer Academic Publishers, 2000.

[33]. Klement, E. P., Mesiar, R. and Pap, E. On the relationship of associative compensatory operators to triangular norms and conorms. *International Journal of Uncertainty, Fuzziness and Knowledge based Systems, Vol. 4, No. 2*, 129-144, 1996.

[34]. Klement, E. P., A theory of fuzzy measures : a survey, *Fuzzy Information and Decision Processes*, edited by Gupta, M. and Sanchez, E., North Holland : Amsterdam, 59-66, 1982.

[35]. Klir, G. J. and Folger, T. A., Fuzzy Sets, Uncertainty and Information, *Prentice-Hall :* Englewood Cliffs, N.J., 1988.

[36]. Kolmogorov, A., Sur la notion de moyenne, *Atti delle Reale Accademia Nazionale dei Lincei Mem. Cl. Sci. Mat. Natur. Sez. 12*, 323-343, 1930.

[37]. Luce, R.D. and Krantz, D.H., Suppes, P., and Tversky, A., Foundations on measurements, *Academic Press*, New York, 1990.

[38]. Luhandjula, M. K., Compensatory operators in fuzzy linear programming with multiple objectives, *Fuzzy Sets and Systems 8*, 245-252, 1982.

[39]. Mayor, G. and Trillas E., On the representation of some Aggregation functions, *Proceeding of ISMVL*, 111-114, 1986.

[40]. Menger, K., Statistical metrics, *Proc. Nat. Acad. Sci. U.S.A. 8*, 535-537, 1942.

[41]. Mesiar R. and Komorníková M., Aggregation Operators, *Proceeding of the XI Conference on applied Mathematics PRIM' 96*, Herceg D., Surla K. (eds.), Institute of Mathematics, Novi Sad, 193-211, 1997.

[42]. O'Hagan, M., Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic, *Proceedings of the 22nd Annual IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, Ca., 681-689, 1988.

[43]. Ovchinnikov, S., On Robust Aggregation Procedures, *Aggregation Operators for Fusion under Fuzziness.* Bouchon-Meunier B. (eds.), 3-10, 1998.

[44]. Prade H., Reasoning with fuzzy default values, *Proceedings of the 5th International Symposium on Multiple-Valued Logic*, Kingston, Ont., May, 191-197, 1985.

[45]. Schweizer B. and Sklar A., Probabilistic metric spaces, *North Holland*, New York, 1983.

[46]. Schweizer, B. and Sklar, A., Statistical metric spaces, *Pacific J. Math. 10*, 313-334, 1960.

[47]. Silvert, W., Symmetric summation : a class of operations on fuzzy sets, *IEEE Transactions on Systems, Man and Cybernetics 9*, 659-667, 1979.

[48]. Sugeno, M., Theory of fuzzy integrals and its application, *Doctoral thesis*, Tokyo Institute of Technology, 1974.

[49]. Turksen, I. B., Interval-valued fuzzy sets and 'compensatory AND', *Fuzzy Sets and Systems 51*, 295-307, 1992.

[50]. Yager, R. R. and Kacprzyk J. (Eds.), *The Ordered Weighted Averaging Operators, Theory and Applications*, Kluwer Academic Publisher : Boston, Dordrecht, London, 1997.

[51]. Yager, R. R. and Rybalov, A., Full reinforcement operators in aggregation techniques, *IEEE Transactions on Systems, Man and Cybernetics 28*, 757-769, 1998.

[52]. Yager, R. R., and Rybalov, A., Uninorm Aggregation Operators, *Fuzzy Sets and Systems 80*, 111-120, 1996.

[53]. Yager, R. R., A new methodology for ordinal multiple aspect decisions based on fuzzy sets, *Decision Sciences 12,* 589-600, 1981.

[54]. Yager, R. R., Non-monotonic set-theoretic operators, *Fuzzy Sets and Systems 42*, 173-190, 1996.

[55]. Yager, R. R., On ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE transactions on Systems, Man and Cybernetics 18*, 183-190, 1988.

[56]. Zadeh, L. A., Fuzzy Logic = Computing with Words, *IEEE Transactions on Fuzzy Systems* 59, 125-148, 1993.

[57]. Zimmermann, H.-J. and Zysno, P., Latent connectives in human decision making, *Fuzzy Sets and Systems 4*, 37-51, 1980.

# Table of Contents