Knowledge-Based Systems 38 (2013) 3–13

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Information aggregation in intelligent systems: An application oriented approach

Imre J. Rudas [a], Endre Pap [a,b,c,*], János Fodor [a]

[a] Óbuda University, Institute of Intelligent Engineering Systems, Becsi út 96/B, H-1034 Budapest, Hungary
[b] University of Novi Sad, Trg Dositeja Obradovica 4, 21000 Novi Sad, Serbia
[c] Educons University, Vojvode Putnika 87, 21208 Sremska Kamenica, Serbia

ARTICLE INFO

Article history:
Available online 24 August 2012

Keywords:
Aggregation function
Mean
Triangular norm
Uninorm
Decision making
Fuzzy inference system
Multisensor data fusion
Image approximation

ABSTRACT

This paper offers a comprehensive study of information aggregation in intelligent systems prompted by common engineering interest. After a motivating introduction we consider aggregation functions and their fundamental properties as a basis for further development. Four main classes of aggregation functions are identified, and important subclasses are described and characterized as prototypes. For practical purposes, we outline two procedures to identify aggregation function that fits best to empirical data. Finally, we briefly recall some applications of aggregation functions in decision making, utility theory, fuzzy inference systems, multisensor data fusion, image processing, and their hardware implementation.

1. Introduction

The demand for appropriate tools to combine information arising from different sources into a single value has been of central interest for centuries (think on probability and statistics). The advent and fast development of computers just accelerated and widened the groups of potential applications. This process inspired also more intensive theoretical research, although the study of aggregation operations has always been an active field of interest. Mathematical investigations have especially been motivated by the development of fuzzy set theory and fuzzy logics [57], but also by some non-classical approaches to decisions (e.g. multicriteria decision theory).

Information fusion in intelligent systems is also a fundamental problem, and its use is rapidly increasing as more complex systems are being developed. For example, fields such as robotics (e.g., fusion of data provided by sensors), vision (e.g., fusion of images), knowledge based systems (e.g., decision making in a multicriteria framework, integration of different kinds of knowledge, and verification of knowledge-based systems correctness) and data mining (e.g., ensemble methods) are well known. According to Dubois and Prade [26] the aims of information fusion are one or several of the following: to improve the available knowledge, to update the current information, to lay bare a consensual opinion, to derive or improve generic knowledge by means of data. As it is emphasized in [50], recent advances in multiagent systems extend the range of information fusion applications in systems where an agent needs to consider the behavior of other agents to make decisions on the basis of distributed information.

The process of combining several (numerical) values into a single representative one is called *aggregation*, and the function performing this process is called *aggregation function* [30]. In this paper we consider aggregation functions as mappings that assign a single output in the closed unit interval $[0,1]$ to several inputs from the same interval. At first this may seem to be a restrictive consideration because aggregation functions may depend on the nature of the items to be aggregated and also on the representation framework. Nevertheless, it is still possible to offer a joint approximate interpretation as follows. If X denotes a set (of possible worlds, of states, of alternatives) and L is an evaluation scale then a *profile* is defined as a function μ from X to L. This is an extension of the classical notion of a membership function [57], where $L = [0,1]$ is the range of membership functions. As it is claimed in [26], items of information can be interpreted, at least approximately, in this form. By adopting this approach, the aggregation problem finally can be modeled by suitable *profile aggregation* operations.

The aim of the present paper is to provide an opportunity for interested readers to update and deepen their knowledge on information aggregation in intelligent systems. We remark that some of aggregation functions are used in the so called pseudo-analysis which is a generalization of the classical analysis, where instead of the field of real numbers a semiring is taken on a real interval $[a,b] \subset [-\infty,\infty]$ endowed with pseudo-addition $\oplus$ and

* Corresponding author at: University of Novi Sad, Trg Dositeja Obradovica 4, 21000 Novi Sad, Serbia.
E-mail addresses: rudas@uni-obuda.hu (I.J. Rudas), pape@eunet.rs (E. Pap), fodor@uni-obuda.hu (J. Fodor).

0950-7051/$ - see front matter
http://dx.doi.org/10.1016/j.knosys.2012.07.025

with pseudo-multiplication $\odot$, see [45,47], which is useful in modeling uncertainty and nonlinearity. Some applications are given in Section 5.6.

To this end, the paper is organized as follows. In Section 2 we define aggregation functions and study their fundamental properties as a basis for further development. We shall use the notations from Ref. [30]. In Section 3 we identify four main classes: conjunctive, disjunctive, internal, and mixed aggregation functions. In each class we identify and characterize prototypical subclasses. In Section 4, for practical purposes, we outline two procedures to identify aggregation function that fits best to empirical data. In Section 5 we sketch some applications of information aggregation in decision making, fuzzy inference systems, multisensor data fusion, image processing, and hardware implementation of parametric operations in fuzzy logic. At the end our conclusion is drawn.

2. Aggregation functions

We start with some well known simple examples, which are often used in many fields. One of the most often used aggregation function is the arithmetical mean. For two real numbers a and b we are finding a number $\frac{a+b}{2}$, which is exactly the middle point between a and b. But if we consider the arithmetical mean of a and b with the third number c, we have to calculate from the beginning as $\frac{a+b+c}{3}$. Namely, we have to take care that we cannot obtain this value taking simple the arithmetical mean of $\frac{a+b}{2}$ and c! Therefore, we calculate the arithmetical mean of n numbers $a_1, a_2, \ldots, a_n$ by

$$\mathsf{AM}(a_1, a_2, \ldots, a_n) = \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{1}{n}\sum_{i=1}^{n} a_i.$$

We shall use in the rest of paper the notation AM for the arithmetical mean as a function of n-variables. We are very often use the arithmetical mean, but the question is when it is meaningfull to use it. For example, if we are looking for a mean amount of rainfall, then we note that the total amount of rain, which affects crop growth, etc., is obtained by adding the daily numbers. Therefore if we add them up and divide by the number of days, the resulting arithmetic mean is the amount of rain we could have had on each of those days, to get the same total. That means the arithmetical mean is used when we are adding inputs, e.g., it cannot be used averaging the rates of investments return over many years.

When the data are in some finite interval usually with the normalization procedure we can switch the considerations to numbers in the unit interval $[0,1]$. Therefore we shall consider the arithmetical mean as the function $\mathsf{AM}: [0,1]^n \to [0,1]$. Since it works for any natural number n we sometimes use also the notation $\mathsf{AM}^{(n)}$ to stress that it is a function of exactly n-variables. Important property of the arithmetical mean is the monotonicity. Namely, if in the arithmetical mean $\frac{a+b+c}{3}$ we take instead of the number c a greater number c', then obviously that the new arithmetical mean $\frac{a+b+c'}{3}$ will be greater than the previous one. This would be true if we done the same on any coordinate, and also true for general case of n-dimension. The second important property of the arithmetical mean we obtain if we take the border numbers *0* and 1. Namely, the arithmetical mean of n zeros is again zero and the arithmetical mean of n ones is again one (boundary conditions).

As we already said, in this paper we only consider aggregation functions that take arguments from the closed unit interval $[0,1]$ and yield a value also in $[0,1]$. This is usually denoted as $\mathsf{A}: [0,1]^n \to [0,1]$ when there are n components to aggregate.

The inputs that aggregation functions should combine are typically interpreted as degrees of membership in fuzzy sets, degrees of preference, strength of evidence, etc. For example, a rule based system contains rules of the following form:

IF 't_1 is A_1' AND ... AND 't_n is A_n' THEN 'v is B'.

If $x_1, \ldots, x_n$ denote the degrees of satisfaction of the rule predicates 't_1 is A_1', ..., 't_n is A_n' then the overall degree of satisfaction of the combined predicate of the rule antecedent can be calculated as $\mathsf{A}(x_1, \ldots, x_n)$. The input value 0 means no satisfaction, and input 1 is interpreted as full satisfaction. As a consequence, *aggregation functions should preserve the bounds 0 and 1*. This inevitable property is supported also by examples from preference modeling and multicriteria decision making [28].

If in the rule base example we imagine another input vector $(t'_1, \ldots, t'_n)$ with the corresponding degrees of satisfaction $x'_1, \ldots, x'_n$ such that $x_i \leqslant x'_i$ for all $i = 1, \ldots, n$. Then it is natural to expect $\mathsf{A}(x_1, \ldots, x_n) \leqslant \mathsf{A}(x'_1, \ldots, x'_n)$, that is, the overall degree of satisfaction of the combined predicate of the rule antecedent cannot be smaller in the second case than in the first one. In other words: *aggregation functions should be monotone non-decreasing*.

Having in mind these two inevitable properties of aggregation functions, we are ready to give the general definition of aggregation functions, see also [11,16,30,50]: A function $\mathsf{A}: [0,1]^n \to [0,1]$ is called an *aggregation function* in $[0,1]^n$ if

(i) A is nondecreasing: for $x_1 \leqslant x'_1, \ldots, x_n \leqslant x'_n$ we have
$$\mathsf{A}(x_1, x_2, \ldots, x_n) \leqslant \mathsf{A}(x'_1, x'_2, \ldots, x'_n);$$
(ii) A fulfills the boundary conditions
$$\mathsf{A}(0, 0, \ldots, 0) = 0 \quad \text{and} \quad \mathsf{A}(1, 1, \ldots, 1) = 1.$$

Just to illustrate the generality of this definition, we list several aggregation functions in $[0,1]^n$ that – otherwise – behave quite differently.

Example 2.1. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in [0,1]^n$. It is easy to justify that the following functions are aggregation functions.

- *The smallest aggregation function*: $\mathsf{A}_\perp(\mathbf{x}) = 1$ if $(x_1, x_2, \ldots, x_n) = (1, 1, \ldots, 1)$, and $\mathsf{A}_\perp(\mathbf{x}) = 0$ otherwise.
- *The greatest aggregation function*: $\mathsf{A}_\top(\mathbf{x}) = 0$ if $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$, and $\mathsf{A}_\top(\mathbf{x}) = 1$ otherwise.
- *Arithmetic mean*: $\mathsf{AM}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i$.
- *Geometric mean*: $\mathsf{GM}(\mathbf{x}) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$.
- *Minimum*: $\mathsf{Min}(\mathbf{x}) = \min\{x_1, \ldots, x_n\}$.
- *Maximum*: $\mathsf{Max}(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$.
- *Product*: $\Pi(\mathbf{x}) = \prod_{i=1}^{n} x_i$.
- *Bounded sum*: $\mathsf{S}_{\mathbf{L}}(\mathbf{x}) = \mathsf{Min}\left(1, \sum_{i=1}^{n} x_i\right)$.
- k-Order statistics: $\mathsf{OS}_k(\mathbf{x})$ is defined for any $k \in \{1, \ldots, n\}$ as the kth smallest argument of $\mathbf{x}$. Note that $\mathsf{OS}_1 = \mathsf{Min}$ and $\mathsf{OS}_n = \mathsf{Max}$.
- *kth Projection*: $\mathsf{P}_k: [0,1]^n \to [0,1]$ is defined as the kth argument of $\mathbf{x}$, $\mathsf{P}_k(x_1, x_2, \ldots, x_n) = x_k$.
- 3-Π-*Operator* [55]: $\mathsf{E}(\mathbf{x}) = \frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} x_i + \prod_{i=1}^{n}(1-x_i)}$, with the convention $\frac{0}{0} = 0$.

For example, we can apply geometric mean only to positive numbers and we can model with it the proportional growth, both exponential growth and compound interest growth rate, e.g., averaging the rates of investments return over many years. Geometric mean we can also apply when a rectangle is given and we have to find a square equal to the area of the rectangle, e.g., for a rectangle with the length 28 and width 7 the corresponding area is $28 \cdot 7 = 196$ we find the length of the side of a square which equals the area of the rectangle as the square root of 196 what is 14 (the numbers 7, 14, 28 form geometric progression and square root of $28 \cdot 7$ is the middle number 14).

Now we go through elementary properties of aggregation functions. Some of them are more or less known properties of functions, the others are linked to some algebraic expressions.