INFORMATION FUSION IN INTELLIGENT SYSTEMS

Endre Pap

Singidunum University,
32 Danijelova Street, Belgrade, Serbia

Abstract:

Information fusion in intelligent systems is one of the basic problems. This issue is rapidly increasing since more complex systems are being developed, *e.g.*, robotics, vision, knowledge-based systems and data mining. The process of combining several values into a single representative one is called aggregation, and we present two applications of aggregation functions in multisensor data fusion and network aggregation in sensor networks. Special attention is devoted to the situation of pairs of aggregation functions under the semiring structure (pseudo-analysis). The author shall present two applications of it in image processing.

Key words:

aggregation function, multisensor, semiring, pseudo-integral.

Acknowledge:

The author wishes to acknowledge the financial support of MPNTRS (the project 174009) and of PSNTR of Vojvodina.

1. INTRODUCTION

Information fusion is a fundamental problem in intelligent systems and its importance is rapidly increasing since many complex systems are being developed. There are many areas where there is a need to develop theoretical background for information fusion, *e.g.*, fields such as robotics (fusion of data provided by sensors), vision (fusion of images), knowledge based systems (decision making in a multicriteria framework, integration of different kinds of knowledge, and verification of knowledge-based systems correctness) and data mining (ensemble methods) are well known. The aims of information fusion include 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.

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* [8]. 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.

In order to deal with the problems involving interaction between

Correspondence:
Endre Pap

e-mail:
epap@singidunum.ac.rs

DOI: 10.15308/Sinteza-2016-100-105

criteria various classes of nonadditive set functions have been introduced. Based on them Choquet and Sugeno (see [14]), idempotent [9], universal, pseudo-integral [14] and other integrals [14], [15], [23] were introduced. The pseudo-integral is part of the pseudo-analysis, where a real interval $[a, b]\subset[-\infty,\infty]$ equipped insted of the usual addition and product, with two operations: the pseudo-addition $\oplus$ and the pseudo-multiplication $\odot$. Based on $\oplus$-measure (pseudo-additive measure), pseudo-integral is defined. Then properties of the pseudo-integral including inequalities for this type integral have been investigated, see [14], [19]. The pseudo-analysis is used in many applications, modeling uncertainty, nonlinearity and optimization, see [9], [12], [13], [16], [17].

This paper is organized as follows. Section II presents some basic facts on aggregation functions. Applications of aggregation functions in multisensor data fusion and sensor networks are presented in Section III. Some elements of pseudo-analysis are presented in Section IV. Two applications of pseudo-analysis are given in Section V.

2. AGGREGATION FUNCTIONS

One of the mainly used aggregation function is the arithmetical mean AM: $[0,1]^n \to [0,1]$. of n numbers $x1$, $x2,\dots, xn \in [0,1]$, given by

$$\mathsf{AM}(x_1, a_2, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Important property of the arithmetical mean is monotonicity, *i.e.*, if in the arithmetical mean $(x+y+z)/3$ we take instead of the number z a greater number z', then obviously that the new arithmetical mean $(x+y+z)/3$ will be greater than the previous one. This would be true if we did 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; *i.e.*, the arithmetical mean of n zeros is again zero and the arithmetical mean of n ones is again one.

The inputs that aggregation functions combine are 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 An*' THEN 'v is B'.

If $x_1,\dots, x_n$ 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 $A(x_1,\dots, x_n)$. If all input values are 0, it implies lack of satisfaction, and if all inputs are 1 then this is interpreted as full satisfaction. As a consequence, aggregation functions should preserve the bounds 0 and 1.

The general definition of aggregation functions, see [8] is the following.

Definition 1:

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 $x_1 \le x_1', \dots, x_n \le x_n'$ we have

$$\mathsf{A}(x_1, x_2,\dots, x_n) \le \mathsf{A}(x_1', x_2',\dots, x_n');$$

(ii) A fulfills the boundary conditions

$$\mathsf{A}(0, 0,\dots, 0) = 0 \text{ and } \mathsf{A}(1,1,\dots, 1) = 1.$$

We list only few aggregation functions in $[0,1]^n$, for pleanty of them see [8]:

(i) Geometric mean: $\mathsf{GM}(\mathbf{x}) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}.$

(ii) Harmonic mean: $\mathsf{HM}(\mathbf{x}) = \dfrac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}.$

(iii) Minimum: $\mathsf{Min}(\mathbf{x}) = \min\{x_1, \ldots, x_n\}.$

(iv) Maximum: $\mathsf{Max}(\mathbf{x}) = \max\{x_1, \ldots, x_n\}.$

We frequently use the arithmetical mean, but the question is when it is meaningfull to use it. In some practical situations, the appropriate mean is uniquely determined by the situation itself.

Example 2:

(i) Connecting n resistors with resistances $x_1,\dots, x_n$ in series, then the average resistance is the arithmetic mean of $x_1,\dots, x_n$. On the other hand, if the resistors are connected in parallel then the effect is the same as if one had used n resistors with the same resistance, all equal to the harmonic mean of $x_1,\dots, x_n$.

(ii) In information retrieval, where precision and recall are two widely used metrics for evaluating the correctness of a pattern recognition algorithm. The harmonic mean of the precision and the recall is often used as an aggregated performance score.

(iii) In hydrology, the arithmetic mean is used to average hydraulic conductivity values for flow that is parallel to layers, while flow perpendicular to layers uses the harmonic mean.

Clasification of aggregation functions, see [8]:

(i) Conjunctive $0 \leq A \leq \text{Min}$;

(ii) Internal (means) $\text{Min} \leq A \leq \text{Max}$;

(iii) Disjunctive $\text{Max} \leq A \leq 1$;

(iv) Mixed.

3. APPLICATION OF AGGREGATION FUNCTIONS

A. Multisensor data fusion

Multisensor data fusion refers to the synergistic combination of sensory data from multiple sensors and related information to provide more reliable and accurate information than could be achieved using a single, independent sensor. The important advantages of multisensor fusion are redundancy, complementarity, timeliness and cost of the information. The fusion of redundant information can reduce the overall uncertainty and thus serve to increase the accuracy with which the features are perceived by the system. Complementary information from multiple sensors allows features in the environment to be perceived that are impossible to perceive using just the information from each individual sensor operating separately [11]. Multisensor data fusion strategies are developed for advanced driver assistance systems (ADAS), see [22], where there are given data fusion concepts, an applicable model, paradigm of multisensor fusion algorithms, current sensor technologies and some applications such as object tracking, identification and classification and a providence view on next-generation car safety and driver assistance systems. The fuzzy logic approach is used in the model as sensor fusion. To follow a vehicle by an adaptive cruise control (ACC) system one must keep safe distance by measuring the front vehicle distance to the host vehicle. The host vehicle is equipped with four types and five sensors among total 16 sensors that should be considered, each of which with different coverage area and may be infected by some environments noise, consequently with deferent measurements regarding the position of front vehicle. All sensor data are fuzzified in order to determine a near real distance. After several modifications and improvements of membership functions with Min-Max aggregation operator of FuzzyTECH Simulator, a satisfactory following by the host vehicle is extracted.

Sensor fusion is also used for classification of objects and for pattern recognition. Let X be a set of objects under consideration (such as airplanes, cars, animals, flowers, *etc.*), and $\{C_1,..., C_p\}$ are predefined classes. Assume that each object in X is measured by a set of sensors $S_1,..., S_n$. Each sensor provides a partial description of objects. For example, a plane can be detected by several radars and observed by other equipment. In a military situation, only two classes of planes count: friend or foe, see [8]. We remark that the aggregation function may depend on the particular class considered. Some sensors may be more discriminative for certain classes, but unable to distinguish some others. This means that the weight assigned to sensors may depend on the classes. In such a situation, internal weighted aggregation functions can be a good choice.

B. In-network aggregation in sensor networks

A sensor network consists of a large number of individual devices called sensor motes. Each device produces a data stream through sensing modalities. Because in many cases there is no need to report the total data stream, and because individual observations may be noisy or missing, it is necessary to aggregate information collected by sensors. A spanning tree is used, where each sensor combines its own observations with those received from its children. This is an effective procedure for aggregation functions such as Min, Max, AM when there is no failure. In general, one can use so-called decomposable aggregation functions, where the value of the aggregation function can be computed for disjoint subsets of variables, and then in the second step, these value can be aggregated to obtain the aggregate of the whole set of variables. Node or link failures, or packet losses may cause significant change in the aggregated value. For instance, some aggregation functions such as AM are duplicate-sensitive: incorrect aggregated value is resulted when the same value is counted multiple times. In [6] an approximate method is proposed to overcome these difficulties of duplicate-sensitive aggregation functions in faulty sensor networks. In a similar way, a flexible on-board stream processing method of sensor data of a vehicle is introduced in [21].

4. PSEUDO-ANALYSIS

Let $\preceq$ be full order on $[a, b] \subseteq [-\infty, \infty]$.

The pseudo-addition $\oplus : [a, b]^2 \to [a, b]$ is a function that is commutative, non-decreasing with respect to $\preceq$, associative and with a zero (neutral) element denoted by 0: Let $[a, b]_+ = \{x \mid x \in [a, b], 0 \preceq x\}$. The pseudo-multiplication $\odot : [a, b]^2 \to [a, b]$ is a function that is commutative, positively non-decreasing, *i.e.*, $x \preceq y$ implies $x \odot z \preceq y \odot z$ for all $z \in [a, b]_+$, associative and with a unit element $1 \in [a, b]$, *i.e.*, for each $x \in [a, b]$, $1 \odot x = x$. The pseudo-multiplication $\odot$ is distributive over the pseudo-addition $\oplus$, *i.e.*, $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$, and $0 \odot x = 0$. The interval $[a, b]$ equipped with pseudo-multiplication $\odot$ and pseudo-addition $\oplus$ is a semiring, denoted by $([a, b], \oplus, \odot)$ (see [10], [14]).

We list three characteristic cases:

Case I: (Idempotent $\oplus$ and non-idempotent $\odot$)

$x \oplus y = sup(x, y)$, is an arbitrary not idempotent pseudo-multiplication on the interval $[a, b]$. A full order is induced by the idempotent operation sup as follows: $x \preceq y$ if and only if $sup(x, y) = y$. The pseudo-multiplication $\odot$ is generated by an increasing bijection $g : [a, b] \to [0, 1]$, $x \odot y = g^{-1}(g(x) \cdot g(y))$. It holds $0 = a$ and $1 = g^{-1}(1)$.

Case II: (*g-semiring*) The pseudo-operations are defined by $x \oplus y = g^{-1}(g(x) + g(y))$ and $x \odot y = g^{-1}(g(x) \cdot g(y))$, where $g : [a, b] \to [0, 1]$ is continuous and monotone.

Case III: (Idempotent $\oplus$ and idempotent $\odot$)

$x \oplus y = sup(x, y)$, $x \odot y = inf(x, y)$, on the interval $[a, b]$. Here is $0 = a$, $1 = b$ and the pseudo-addition induces the usual order. Let X be a non-empty set and A be a σ-algebra of subsets of X, *i.e.*, (X, A) be a measurable space.

Definition 3:

A σ-$\oplus$-measure is a set function $m : A \to [a, b]_+$ such that the following conditions are fulfilled:

(i) $m(\emptyset) = 0$ (for non-idempotent $\oplus$)

(ii) for any sequence $(A_i)_{i \in N}$ of mutually disjoint sets from A, we have: $m\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigoplus_{i=1}^{\infty} m(A_i)$

The pseudo-characteristic function of a set A is defined by

$$\chi_A(x) = \begin{cases} \mathbf{0} & , \quad x \notin A, \\ \mathbf{1} & , \quad x \in A. \end{cases}$$

A step (measurable) function is a mapping $e : X \to [a, b]$ that has the following representation $e = \bigoplus_{i=1}^{n} a_i \odot \chi_{A_i}$ for $\alpha_i \in [a, b]$ and sets $A_i \in A$ are pairwise disjoint if $\oplus$ is nonidempotent.

Definition 4:

Let $m : A \to [a, b]$ be a σ-$\oplus$-measure.

(i) The pseudo-integral of a step function $e : X \to [a, b]$ is defined by

$$\int_X^{\oplus} e \odot dm = \bigoplus_{i=1}^{n} a_i \odot m(A_i).$$

(ii) The pseudo-integral of a measurable function $f : X \to [a, b]$, (if $\oplus$ is not idempotent we suppose that for each $\varepsilon > 0$ there exists a monotone ε-net in $f(Z)$) is defined by

$$\int_X^{\oplus} f \odot dm = \lim_{n \to \infty} \int_X^{\oplus} e_n(x) \odot dm,$$

where $(e_n)_{n \in N}$ is a sequence of step functions such that $d(e_n(x), f(x)) \to 0$ uniformly as $n \to \infty$.

For more details see [9], [14].

We shall consider the semiring $([a, b], \oplus, \odot)$ for three cases, namely I, II and III. If the pseudo-operations are generated by a monotone and continuous function $g : [a, b] \to [0, \infty]$, then the pseudo-integral for a measurable function $f : X \to [a, b]$ is given by,

$$\int_X^{\oplus} f \odot dm = g^{-1}\left(\int_X (g \circ f)\, d(g \circ m)\right), \tag{1}$$

where the integral on the right side is the Lebesgue integral. When $X = [c, d]$, $A = B(X)$ and $m = g^{-1} \circ \lambda$, λ the Lebesgue measure on $[c, d]$, then we use notation

$$\int_{[c,d]}^{\oplus} f(x)\, dx = \int_X^{\oplus} f \odot dm. \tag{2}$$

This form of pseudo-integral is known as the g-integral, see [14]. When the semiring is of the form $([a, b], \sup, \odot)$, cases I and III and function $\psi : X \to [a, b]$ defines σ-sup-measure m by $m(A) = \sup_{x \in A} \psi(x)$, then the pseudo-integral for a function $f : X \to [a, b]$ is given by

$$\int_X^{\oplus} f \odot dm = \sup_{x \in X} \left(f(x) \odot \psi(x)\right). \tag{3}$$

5. APPLICATION OF PSEUDO-ANALYSIS

Many applications can be found in [2], [9], [12], [13], [14], [16]. We shall give here only two applications.

A. Image approximation

If one replaces the sum with the maximum $\bigvee$, the general form of a max-product approximation operator is obtained:

$$P_n(f, x) = \bigvee_{i=0}^{n} K_n(x, x_i) \cdot f(x_i), \tag{4}$$

where $K_n(\cdot, x_i) : X \to [0, \infty[$, $i = 0, \ldots, n$, are a given continuous functions. Here $f : X \to [0, \infty[$ is a continuous function on a compact metric space (X, d). Let also, $x_i \in X$, $i \in \{0,..., n\}$, $n \geq 1$ be fixed sampled data. Such an operator was used in image approximation can be found in [7] where max-product Shepard operator was defined and studied. In several cases, max-product approximation outperforms classical linear approximation in the sense that they lead to essentially better error estimates, if for example Bernstein basis is used in the construction of max-product operators.

As second case we consider a pair of operators $\oplus$ and $\odot$ from case II. We use these pairs of operations instead of the standard addition and multiplication of the reals as follows:

$$P_n(f, x) = \bigoplus_{i=0}^{n} K_{n,i}(x) \odot f_i, \qquad (5)$$

where $K_{n,i}(x) : X^2 \to [0,1]$ are given continuous functions, $i=1,..., n$. The operator P_n is continuous and pseudo-linear. In the formula (5) f_i denotes the classical discrete cosine transform of the target function f, and K where g^{-1} is the inverse of the generator function g, and $A_{n,i}$ is a function defined through the cosine function for all $i=1,..., n$. Then $P_n(f, x)$ in (5) is called the pseudo-linear inverse DCT, see [3]. The pseudo-linear DCT has very good reconstruction property, *i.e.*, in image compression.

B. *Pseudo-linear superposition principle for Perona and Malik equation*

Partial differential equators are successfully applied to the relevant problem of image processing, see [1], [4], [20]. In that method, a restored image can be seen as a version of the initial image at a special scale. An image u is embedded in an evolution process, denoted by $u(t,\cdot)$. The original image is taken at time $t=0$, $u(0,\cdot)=u_0(\cdot)$ and is then transformed. The idea is to construct a family of functions $\{u(t, x)\}_{t>0}$ representing successive versions of $u_0(x)$. As t increases $u(t, x)$ changes into a more and more simplified image. We would like to attain two goals. The first is that $u(t, x)$ should represent a smooth version of $u_0(x)$, where the noise has been removed. The second is to be able to preserve some features such as edges, corners, which may be viewed as singularitis. The heat equation has been (and is) successfully applied to image processing but it has some drawbacks. It is too smoothing and because of that, edges can be lost or severely blurred. In [1] authors consider models that are generalizations of the heat equation. The domain image will be a bounded open set Ω of $\mathbb{R}^2$. The following equation is initially proposed by Perona and Malik [20]:

$$\begin{cases} \frac{\partial u}{\partial t} = \operatorname{div}\left(c\left(|\nabla u|^2\right)\nabla u\right) & \text{in }]0,T[\times\Omega, \\ \frac{\partial u}{\partial N} = 0 \text{ on }]0,T[\times\partial\Omega, & \\ u(0,x) = u_0(x) \text{ in } \Omega & \end{cases} \qquad (6)$$

where: c: $[0,\infty[\to]0,\infty[$. If we choose $c\equiv1$, then it is reduced on the heat equation. If we assume that $c(s)$ is a decreasing function satisfying $c(0)=1$ and $\lim_{s\to\infty}c(s)=0$, then inside the regions where the magnitude of the gradient of u is weak, equation (6) acts like the heat equation and the edges are preserved.

We have proved in [18] that the pseudo-linear superposition principle holds for Perona and Malik equation.

Theorem 5:

If $u_1=u_1(t, x)$ and $u_2=u_2(t, x)$ are solutions of the equation

$$\frac{\partial u}{\partial t} - \operatorname{div}\left(c\left(|\nabla u|^2\right)\nabla u\right) = 0, \qquad (7)$$

then $(\lambda_1 \odot u_1) \oplus (\lambda_2 \odot u_2)$ is also a solution of (7) on the set

$$D = \{(t,x) \,|\, t \in]0,T[, x \in \mathbb{R}^2, u_1(t,x) \neq u_2(t,x)\},$$

with respect to the operations $\oplus = \min$ and $\odot = +$.

The obtained results will serve for further investigation of weak solutions of the equation (7) in the sense of Maslov [9], [12], [16], [17], with further important applications.

REFERENCES

[1] Aubert, G., & Kornprobst, P. (2002). Mathematical Problems in Image Processing, Springer-Verlag.

[2] Baccelli, F., Cohen, G., Olsder G.J., & Quadrat J. (1992). Synchronization and Linearity: an Algebra for Discrete Event Systems. New York, John-Wiley and Sons.

[3] Bede, B., Nobuhara, H., Rudas, I. J., & Fodor, J. (2008). Discrete Cosine Transform based on Uninorms and Absorbing Norms, Proceedings of the IEEE International Conference on Fuzzy Systems, Hong Kong, 2008, 1982–1986.

[4] Catte, F., Lions, P.L., Morel, J. M., & Coll, T. (1992). Image selective smoothing and edge detection by nonlinear diffusion, SIAM Journal of Numerical Analysis, 29(1), 182-193.

[5] Bede, B., Coroianu, L., & Gal, SG. (2010). Approximation and shape preserving properties of the nonlinear Meyer-Konig and Zeller operator of max-product kind, Numerical Functional Analysis and Optimization 31, 232–253.

[6] Considine, J., Li, F., Kollios, G., & Byers, J. (2004). Approximate aggregation techniques for sensor databases, Proceedings of the 20th International Conference on Data Engineering, ICDE 2004, IEEE Computer Society, 2004, 30 March - 2 April, 449–460.

[7] Bede, B., Coroianu, L., & Gal, SG. (2010). Approximation and shape preserving properties of the nonlinear Meyer-Konig and Zeller operator of max-product kind, Numerical Functional Analysis and Optimization 31, 232–253.

[8] Grabisch, M., Marichal, J.-L. Mesiar, R., & Pap, E. (2009). Aggregation Functions, Cambridge University Press, Encyclopedia of Mathematics and its Applications 127.

[9] Kolokoltsov, V. N., & Maslov, V. P. (1997). Idempotent Analysis and Its Applications, Kluwer Academic Publishers, Dordrecht, Boston, London.

[10] Kuich, W. (1986). Semirings, Automata, Languages, Berlin, Springer-Verlag.

[11] Luo, R. C. and Chou, Y. and Chen, O. (2007). Multisensor fusion and integration: algorithms, applications, and future research directions, Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation, 2007, August 5-8, 1986–1991, Harbin, China.

[12] Maslov, V. P. (1987). Asymptotic methods for solving pseudo-differential equations, (in Russian), Nauka, Moscow.

[13] Pap, E. (1997). Pseudo-analysis as a mathematical base for soft computing. Soft Computing 1, 61-68.

[14] Pap, E. (1995).Null-Additive Set Functions, Kluwer Academic Publishers, Dordrecht- Boston-London.

[15] Pap E.(ed.) (2002). Handbook of Measure Theory, Elsevier Science, Amsterdam.

[16] Pap, E. (2002). Pseudo-Additive Measures and Their Applications, Handbook of Measure Theory In: E. Pap, editor, Handbook of Measure Theory, Vol II, Elsevier Science, 1403-1465.

[17] Pap, E. (2008). Pseudo-analysis approach to nonlinear partial differential equations, Acta Polytechnica Hungarica 5, 31-45.

[18] Pap, E., & Strboja, M. (2007). Image processing based on a partial differential equation model satisfying the superposition principle, Proc. of International Workshop Idempotent and Tropical Mathematics and Problems of Mathematical Physics, Volume II, Moscow, August 25-30, 2007, 38-42.

[19] Pap, E., & Strboja, M. (2013). Generalizations of Integral Inequalities for Integrals Based on Nonadditive Measures, (Ed. E. Pap) Intelligent Systems: Models and Application, Springer, 3-22.

[20] Peron, P., & Malik, J. (1900).Scale-space and edge detection using anisotropic diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7): 629-639.

[21] Schweppe, H. Zimmermann, A., & Grill, D. (2010). Flexible on-board stream processing for automotive sensor data, IEEE Transactions on Industrial Informatics 6, 81–92.

[22] Rezaei, G., & Sabzevari, R. (2009). Multisensor Data Fusion Strategies for Advanced Driver Assistance Systems, (Ed. N. Milisavljevic)´ Sensor and Data Fusion, In-Teh, 141–166.

[23] Wang, Z., & Klir, G. J. (2009). Generalized Measure Theory, Springer, Boston.