Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Approximate Reasoning
41 (2006) 43–58
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

The interpretation of fuzzy integrals and their application to fuzzy systems ☆

Vicenç Torra [a,*], Yasuo Narukawa [b]

[a] Institut d'Investigació en Intel·ligència Artificial, Campus de Bellaterra, E-08193 Bellaterra, Catalonia, Spain
[b] Toho Gakuen, 3-1-10 Naka, Kunitachi, Tokyo 186-0004, Japan

Available online 29 August 2005

Abstract

Fuzzy integrals, in general, and Sugeno integrals, in particular, are well known aggregation operators. They can be used in a great variety of decision making applications. Nevertheless, their use is not easy as their interpretation is not straightforward. In this paper we study the interpretation of fuzzy integrals, focusing on Sugeno ones, and we show their application to fuzzy inference systems when the rules are not independent.

Keywords: Fuzzy integrals; Sugeno integral; Choquet integral; Twofold integral; Fuzzy inference system

1. Introduction

Within aggregation operators, fuzzy integrals are known to be one of the most powerful and flexible functions as they permit the aggregation of information under different assumptions on the independence of the information sources. In particular,

☆ Work partly funded by the AGAUR—Generalitat de Catalunya (2004XT 00004) and the Spanish Ministry of Education and Science under project SEG2004-04352-C04-02 "PROPRIETAS".
* Corresponding author. Tel.: +34 93580 9570; fax: +34 93580 9661.
E-mail addresses: vtorra@iiia.csic.es (V. Torra), narukawa@d4.dion.ne.jp (Y. Narukawa).

0888-613X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2005.08.001

they can be used to model situations in which sources are independent as well as in situations in which such independence cannot be assured.

The most well-known integrals are the Choquet [4] and the Sugeno [14] integrals. Generalizations of such integrals include e.g. the t-conorm integral [12], the twofold integral [16] or the Composition aggregation operators in [3]. See also [9].

The flexibility of such operators is also due to the fact that they generalize several of the most widely-known and used aggregation functions. In particular, they generalize the arithmetic and weighted mean, as well as the median and linear combination of order statistics (or OWA [19]).

The flexibility of such operators is tightly related with the difficulties of using them in practical applications.

Fuzzy integrals combine the data supplied by several information sources according to a fuzzy measure. This fuzzy measure, that (using Artificial Intelligence terminology) represents the background knowledge on the information sources, is a set function from the set of information sources into an appropriate domain (e.g. the $[0,1]$ interval or an ordered set D). Typically, this fuzzy measure represents the importance or relevance of the sources when computing the aggregation.

For building a real system, several difficulties arise. One of them is that the set function needs to be defined, and this requires 2^{n-1} values, where n is the number of information sources. Thus, there is a curse of dimensionality. Another difficulty for the use of fuzzy integrals in real applications is that their interpretation is not easy. In this respect, an interpretation was introduced in [10] for the Choquet integral. Instead, for the Sugeno integral, although there exist some work on the mathematical properties of such integrals, there is not yet a clear interpretation of their operational principles. Let alone, about the meaning of the fuzzy measures they need to operate.

In this paper we study the interpretation of fuzzy integrals, focusing on Sugeno integrals. We give some examples and describe an application that corresponds to their use in fuzzy systems.

The structure of the paper is as follows. In Section 2 we review some results that are needed in the rest of the paper. Then, in Section 3 we study the interpretation of the Sugeno integral. Section 4 is devoted to the use of Sugeno integrals for fuzzy inference systems. Section 5 gives interpretation of other integrals. Finally, the paper finishes with some conclusions.

2. Preliminaries

Basic definitions of aggregation operators are described in this section. We focus on fuzzy integrals (Choquet, Sugeno and twofold integrals). Weighted minimum and weighted maximum are also reviewed as they will also be used in the rest of the paper. See [11,1] for details on fuzzy measures and fuzzy integrals and [2] for a broader view of the field of aggregation operators. Definitions given below are based on a set X that corresponds to the set of information sources. In this paper, we assume that X is finite.

Definition 1. A set function $\mu : 2^X \rightarrow [0,1]$ is a fuzzy measure if it satisfies the following axioms:

(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$ (boundary conditions).
(ii) $A \subseteq B$ implies $\mu(A) \leqslant \mu(B)$ (monotonicity) for $A, B \in 2^X$.

Definition 2 [4]. Let μ be a fuzzy measure on $(X, 2^X)$. Then, the Choquet integral $C_\mu(f)$ of $f : X \rightarrow [0,1]$ with respect to μ is defined by

$$C_\mu(f) := \sum_{j=1}^{n} f(x_{s(j)})(\mu(A_{s(j)}) - \mu(A_{s(j+1)})),$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that

$$0 \leqslant f(x_{s(1)}) \leqslant \cdots \leqslant f(x_{s(n)}) \leqslant 1, \quad A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(n)}\}, A_{s(n+1)} = \emptyset.$$

Definition 3 [14]. Let μ be a fuzzy measure on $(X, 2^X)$. Then, the Sugeno integral $S_\mu(f)$ of a function $f : X \rightarrow [0,1]$ with respect to μ is defined by

$$S_\mu(f) := \bigvee_{j=1}^{n} f(x_{s(j)}) \wedge \mu(A_{s(j)}),$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that

$$0 \leqslant f(x_{s(1)}) \leqslant \cdots \leqslant f(x_{s(n)}) \leqslant 1, \quad A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(n)}\}, A_{s(n+1)} = \emptyset.$$

Here $\wedge$ denotes the minimum and $\vee$ denotes the maximum.

Definition 4 ([16,13]). Let μ_C and μ_S be two fuzzy measures on $(X, 2^X)$. Then, the *twofold integral* of a function $f : X \rightarrow [0,1]$ with respect to the fuzzy measures μ_S and μ_C is defined by

$$\mathrm{TI}_{\mu_S,\mu_C}(f) := \sum_{i=1}^{n} \left(\left(\bigvee_{j=1}^{i} f(x_{s(j)}) \wedge \mu_S(A_{s(j)}) \right) \left(\mu_C(A_{s(i)}) - \mu_C(A_{s(i+1)}) \right) \right),$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that

$$0 \leqslant f(x_{s(1)}) \leqslant \cdots \leqslant f(x_{s(n)}) \leqslant 1, \quad A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(n)}\}, A_{s(n+1)} = \emptyset.$$

Definition 5. Let $\mathbf{u}$ be a n-dimensional vector, that is, $\mathbf{u} := (u_1, \ldots, u_n) \in R^n$.

(1) $\mathbf{u}$ is a *possibility distribution* or a *possibilistic weighting vector* of dimension n if and only if $u_i \in [0,1]$ for all $i \in \{u_1, \ldots, u_n\}$ and $\max_i u_i = 1$.
(2) [5] Let $\mathbf{u}$ be a possibilistic weighting vector of dimension n, then a mapping $\mathrm{WMin} : [0,1]^n \rightarrow [0,1]$ is a *weighted minimum* of dimension n if

$$\mathrm{W}\,\mathrm{Min}_{\mathbf{u}}(a_1, \ldots, a_n) = \min_i \max(1 - u_i, a_i).$$

(3) [5] Let **u** be a possibilistic weighting vector of dimension n, then a mapping WMax: $[0,1]^n \rightarrow [0,1]$ is a *weighted maximum* of dimension n if

$$\mathrm{W}\,\underset{\mathbf{u}}{\mathrm{Max}}(a_1,\ldots,a_n) = \max_i \min(u_i, a_i).$$

3. Interpretation of the Sugeno integral

Although in Definition 3 the Sugeno integral takes the range $[0,1]$ for the input values and also $[0,1]$ for the range of fuzzy measures μ, the definition would be valid for other domains. In particular, any linearly ordered scale D (i.e., D is a linearly ordered set of categories) is suitable for computing

$$S_\mu(f) := \bigvee_{j=1}^{n} f(x_{s(j)}) \wedge \mu(A_{s(j)}).$$

This is so, because this expression only involves operations that are consistently defined in an ordinal scale. Note that $\vee(a,b) = a$ if and only if $a \geqslant b$ and that $\wedge(a,b) = a$ if and only if $a \leqslant b$. At the same time, the ordering s in $f(x_{s(j)})$ can be computed as long as there is an order on $f(x_j)$.

Taking into account this ordinal setting, it is clear that given a set X, both $f(x)$ (for $x \in X$) and $\mu(A)$ (for $A \subseteq X$) should be into the same codomain D. Otherwise, the integral cannot be applied because the minimum cannot be applied to combine $f(x)$ and $\mu(A)$. So, in some sense, both μ and f should denote the same concept.

As μ denotes some *importance*, *reliability*, *satisfaction* or similar concepts, the same should apply to f. Accordingly, the Sugeno integral combines a kind of e.g. importance or reliability leading to another value for importance or reliability. In fact, there are some applications of Sugeno integral in the literature (e.g. [14,18]) that fit with this perspective.

This situation is illustrated in the following example, where the values to be aggregated and the measure (the values in the codomain D) are related with *reliability*.

Example 1. Let X be a set of experts $X = \{x_1, x_2, \ldots, x_n\}$ that evaluate the reliability of a given machine. Say, a new Japanese copy machine. Then, we consider $f(x_i)$ as the reliability of such copy machine according to expert x_i. At the same time, we also consider the reliability of subsets of experts. So, $\mu(A)$ is the reliability of experts in A all together.

For expressing the reliabilities, any ordered set D is appropriate. Nevertheless, for applying the Sugeno integral we need to consider the same set D to express both the reliability of experts and the reliability of copy machines. In other words, we have that for all $A \subseteq X$ it holds $\mu(A) \in D$ and that for all $x_i \in X$ it holds $f(x_i) \in D$.

Making the example concrete, let $X = \{x_1, x_2, x_3\}$ be a set of three experts, then with $\mu(\{x_1\}) = 0.2$, $\mu(\{x_2\}) = 0.3$ and $\mu(\{x_3\}) = 0.4$ we express that the expert x_3 is more reliable than the expert x_2 and that x_2 is more reliable than x_1. Let $\mu(\{x_1, x_2\}) = 0.3$ and $\mu(\{x_1, x_3\}) = 0.4$ represent that joining x_1 to x_2 or to x_3 does not imply a larger reliability than the one of x_2 or x_3 alone. Instead, with $\mu(\{x_2, x_3\}) = 0.8$ we express that joining both x_2 and x_3 their reliability is greatly increased. Finally, according to the boundary conditions, we set $\mu(\emptyset) = 0$ and $\mu(X) = 1$. This is, the set of all experts has the maximum reliability (equal to 1), and that the empty set is not *reliable*.

Then, when experts assign a particular reliability to the copy machine, we can compute an overall reliability using the Sugeno integral. For example, let $f(x_1) = 0.3$, $f(x_2) = 0.7$ and $f(x_3) = 0.6$ be experts' opinions on the reliability of the copy machine. In this case, the Sugeno integral leads to 0.6, that corresponds in this case to the value of one of the most relevant experts.

Another example follows. In this case, the codomain D stands for *satisfaction*. We use then this example to show the application of the Sugeno integral to alternative selection.

Example 2. Let us consider a traveler in Japan that intends to visit Tokyo, Kyoto and Nagano and considers several alternative places for staying. Then, let $X = \{x_1, x_2, x_3\}$ denote the three mentioned cities. That is, x_1 corresponds to Tokyo, x_2 to Kyoto and x_3 to Nagano. Then, we consider the degree of satisfaction of the traveler visiting such cities. Such degree is expressed with the fuzzy measure $\mu(A)$ described in Table 1. The measure of satisfaction is monotonic increasing (the more cities are visited, the greater the satisfaction) and is bounded. Here, the boundary conditions mean that no visit implies no satisfaction, and that visiting all cities has maximum satisfaction (equal to 1).

The degree of satisfaction with respect to visiting certain towns, will be combined with the degree of satisfaction of doing the travel itself to the particular towns. Such degree of satisfaction will be in proportion to the accessibility of such towns from the particular location of the traveler. This degree of satisfaction will be expressed by a function $f: X \rightarrow D$. Assuming that the traveler is located at Tsukuba, the most accessible town is Tokyo. Then, the second accessible town is Nagano and, finally, Kyoto. Therefore, $f(x_1) > f(x_3) > f(x_2)$. Table 2 gives measures for such accessibility from Tsukuba. The values for the measure are expressed using the same terms than the degree of satisfaction. This is, $f(x)$ is comparable with $\mu(A)$.

Using μ and f, we can define $\mu_f(x_i) := \mu(\{x \mid f(x) \geqslant f(x_i)\})$. This expression stands for the degree of satisfaction of visiting x_i and all those cities that are at least as

Table 1
Fuzzy measure for the traveler example: satisfaction degree for visiting cities in $X = \{x_1, x_2, x_3\}$

Set	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_2, x_3\}$	$\{x_1, x_3\}$	X
μ	0.7	0.5	0.2	0.9	0.6	0.8	1

Here, x_1 corresponds to Tokyo, x_2 to Kyoto and x_3 to Nagano.

Table 2
Accessibility degrees from Tsukuba

Set	x_1	x_2	x_3
f	0.8	0.4	0.5

Table 3
Satisfaction degree for each city for the traveler example

Set	x_1	x_2	x_3
μ_f	0.7	1	0.8

accessible as x_i. Roughly speaking, in what accessibility concerns, if we visit a place x_i with a given accessibility $f(x_i)$, then we assume that all places with a greater accessibility than x_i will be also visited. So, in fact, we are considering each $f(x_i)$ as a threshold for selecting the visits. Formally speaking, if we visit x_i, it is possible to visit as well all those cities with a greater degree of possibility than x_i. This is, all those cities in the set $\{x \mid f(x) \geqslant f(x_i)\}$. Accordingly, $\mu_f(x_i)$ is the degree of satisfaction the traveler achieves when decides to visit x_i. Table 3 gives the values for μ_f for each of such towns. The next step for computing a degree of satisfaction of being located in a particular town (in this case Tsukuba), we need to consider the combination of the degrees of satisfaction of visiting some towns and of the degrees of accessibility of visiting such towns.

To do such combination, we first consider for each city x_i, both degrees of accessibility $f(x_i)$ and satisfaction $\mu_f(x_i)$. Two cases can be considered with respect to such $f(x_i)$ and $\mu_f(x_i)$:

- Case $f(x_i) \geqslant \mu_f(x_i)$: In this case, as it is relatively easy to access town x_i, much concern is given to the *satisfaction* $\mu_f(x_i)$. Accordingly, the degree of x_i cannot be larger than $\mu_f(x_i)$.
- Case $\mu_f(x_i) \geqslant f(x_i)$: In this case, the traveler must give special importance to the physical accessibility of x_i, and this constraints the degree of x_i. Thus, the combination is $f(x_i)$.

To take both aspects into account, the degrees are combined by means of the $\wedge$ (the minimum) operator. Thus, $f(x_i) \wedge \mu_f(x_i)$ is the evaluation of visiting x_i.

Taking everything into account, the place x_i with the largest evaluation $f(x_i) \wedge \mu_f(x_i)$ stands for the evaluation of staying in Tsukuba. This largest evaluation corresponds to the Sugeno integral of f with respect to μ, that in this example is equal to $\mathrm{SI}_\mu(f) = \max_{x_i} f(x_i) \wedge \mu_f(x_i) = 0.7$.

In this latter example, we have shown the use of the Sugeno integral to evaluate the satisfaction of the traveler of staying in a particular town. This interpretation of the Sugeno integral can be used in an alternative selection problem. For example, to determine which is the most suitable town for the traveler to stay. In this situation, several towns $\{t_i\}_i$ will be considered and for each one a function f_{t_i} should be defined. At the same time, the satisfaction of visiting certain towns will be expressed (as in Example 2) in terms of a fuzzy measure μ. This fuzzy measure will be constant through the whole computation. Then, the integration of f_{t_i} with respect to μ for each t_i will give the degree of satisfaction of finally staying in town t_i. The town with the best evaluation will be the one selected.

Table 4
Accessibility degrees from Osaka

Set	x_1	x_2	x_3
g	0.4	1	0.7

Table 5
Satisfaction degree for each city for the traveler example when staying in Osaka

Set	x_1	x_2	x_3
μ_g	1	0.5	0.6

Example 3. Let us consider two alternatives for the traveler's staying: Tsukuba and Osaka. Then, given the conditions of Example 2, the satisfaction degree of staying in Tsukuba is 0.7 (see Example 2) and the one of staying in Osaka is $\mathrm{SI}_\mu(g) = 0.6$ when the accessibility degrees from Osaka, denoted by g, are the ones displayed in Table 4. Table 5 gives function μ_g.

As, $\mathrm{SI}_\mu(g) < \mathrm{SI}_\mu(f)$, it means that the traveler will chose to stay in Tsukuba instead of staying in Osaka.

From an operational perspective, it can be considered that the Sugeno integral proceeds like by "saturation". It selects the importance that overcomes (saturates) a certain degree or threshold. In fact, as the threshold is decreasing while the inputs are increasing, it finds a tradeoff (or compromise) between the importance or reliability of the set and the importance that the members of the set have assigned. This follows from the graphical interpretation of the integral (see, e.g., [21]).

4. Sugeno integral for fuzzy inference systems

In this section we consider a different scenario where the central role is played by certainty degrees. This is, we will describe an application of the Sugeno integral where certainty degrees will be aggregated. The application scenario consists on a set of fuzzy rules (although any knowledge based system would lead to a similar scenario) where each rule assigns degrees of satisfaction to particular output values.

Before going into details, we will first review fuzzy inference systems.

4.1. Fuzzy inference systems

Fuzzy inference systems are knowledge based systems defined in terms of sets of rules that include fuzzy sets in their definition. In this section we will only review those elements that are needed latter on in this paper. In particular, we will focus on standard (flat or one-stage) fuzzy inference systems. See e.g. [6,20] for more details and e.g. [17,8] for some recent state-of-the-art applications.

Fuzzy rules R_i (for i in $\{1,\ldots,n\}$) in fuzzy inference systems have the following structure:

$$R_i : \textbf{IF } x^1 \text{ is } A_i^1 \textbf{ and } \ldots \textbf{ and } x^m \text{ is } A_i^m \textbf{ THEN } y \text{ is } B_i.$$

Here, A_i^j and B_i denote fuzzy terms, defined by fuzzy sets. For example, A_i^j might correspond to a fuzzy set expressing that x^j is *low*, *medium* or *high*.

For simplicity, we will only consider here systems that have a single input variable x. Therefore, the structure of the rules is as follows:

$$R_i : \textbf{IF } x \text{ is } A_i \textbf{ THEN } y \text{ is } B_i.$$

Given a set of fuzzy rules $\{R_i\}_i$, and given a particular value for variable x, say x_0, the system computes the output value for variable y. At present, there are two main approaches for computing such distribution. They correspond to two interpretations of the set of rules: conjunctive and disjunctive rules. We will study them below. Nevertheless, in both cases, the output is computed through the obtention of the so-called possibility distribution on the range of y, a function from the range of y into $[0,1]$.

The possibility distribution of the system is built, in an element-wise manner, from similar distributions obtained for each rule R_i. This is, the final possibility, or certainty, degree that y takes a particular value y_0 is defined as the aggregation of the degrees that y equals y_0 according to each rule R_i. Denoting by $f_i(y_0)$ the certainty degree that rule R_i assigns to y_0 and denoting by $\mu(A)$ the certainty degree of rules $R_i \in A$, we will show that the outcome of a fuzzy inference system for y_0 (for both conjunctive and disjunctive interpretations) can be expressed in terms of a Sugeno integral.

This construction using Sugeno integral links, as it will be shown below, fuzzy inference with Weighted Minimum and Weighted Maximum [22], and the latter operators with the Sugeno integral (one of their generalizations).

The next two sections focus on the two types of fuzzy systems. First we consider the case of disjunctive rules and, then, the case of conjunctive rules. As we see it, the examples considered validate our interpretation of the Sugeno integral. The use of Sugeno integral for disjunctive rules was previously suggested in [15].

4.2. The case of disjunctive rules

Let us consider a fuzzy inference system with n rules interpreted in a disjunctive manner. Then, the output of the system when $x = x_0$ is computed as follows:

(1) Compute the satisfaction degree for the antecedent of all rules R_i. This corresponds to compute α_i, where α_i corresponds to the degree of satisfaction of "x_0 is A_i". In our case, as there is a single condition in the antecedent, $\alpha_i = A_i(x_0)$.
(2) Compute the conclusion of rule R_i. For systems defined in terms of disjunctive rules, the output of a rule is often computed using Mamdani's approach. This approach is equivalent to compute the output for $A' = \{x_0\}$ as either $\cup_j(A' \circ R_j)$ or $A' \circ (\cup_j R_j)$ with $\circ$ being a max-min composition and R_j being

the intersection of A_j and B_j. See e.g. [7] for a proof. We apply Mamdani's approach computing $\cup_j(A' \circ R_i)$.From an operational point of view, Mamdani's approach is as follows: for each rule R_i, its output fuzzy set B_i is clipped according to the degree of satisfaction α_i. According to this, the output of such fuzzy rule R_i is $B_i \wedge A_i(x_0)$.

(3) Compute the output of the set of rules $\{R_i\}_i$. This is, the fuzzy output $\tilde{B}$ (for the whole system) is computed as the union of the outputs of each rule R_i. Using, maximum for the union (the most usual operator) we obtain

$$\tilde{B} = \bigvee_{i=1}^{n} (B_i \wedge A_i(x_0)).$$

(4) Finally, the output fuzzy set $\tilde{B}$ is usually defuzzified. In what follows, we skip the defuzzificatioin stage as it is not relevant for our study.

Let us now consider the membership of the fuzzy output $\tilde{B}$ for a given element y_0 in Y. This is, let us consider the computation

$$\tilde{B}(y_0) = \bigvee_{i=1}^{n} (B_i(y_0) \wedge A_i(x_0)).$$

On the light of the weighted maximum (see Definition 3) this expression can be rewritten as

$$\tilde{B}(y_0) = \mathrm{W}\,\underset{\mathbf{u}}{\mathrm{Max}}(B_1(y_0), \ldots, B_n(y_0)), \tag{1}$$

where the weighting vector $\mathbf{u}$ is defined as $\mathbf{u} = (A_1(x_0), \ldots, A_n(x_0))$ or, in general, $\mathbf{u} = (\alpha_1, \ldots, \alpha_n)$. Note that the weighting vector is independent of the value y_0. Thus, for a given x_0, the same aggregation operator with the same weights is applied to all y_0 in Y.

4.3. The case of conjunctive rules

Let us now consider the case of conjunctive rules. In this case there are two alternative expressions for computing the output of the system for a particular input A'. Such expressions are: $\cap_j(A' \circ R_j)$ and $A' \circ (\cap_j R_j)$. Although these two expressions do not lead, in general, to the same output, they are equal when A' is a single value. As this is the case here, we will use $\cap_j(A' \circ R_j)$ for convenience.

In this case, the output of the system when $x = x_0$ is computed as follows:

(1) Compute $A' \circ R_j$ for each rule R_j.
(2) Compute the intersection of all such outputs.

Using the minimum (denoted $\wedge$) to compute the intersection, we have that the output is

$$\tilde{B} = \bigwedge_{i=1}^{n} (A' \circ R_i).$$

This can be rewritten for all $y_0 \in Y$ as follows:

$$\tilde{B}(y_0) = \bigwedge_{i=1}^{n}((A' \circ R_i)(y_0)),$$

where R_i is the relation built from A_i and B_i using an implication function I (see [7] for details on implication functions and a description of several of their families). Formally speaking, R_i is defined as $R_i = I(A_i, B_i)$. Now, due to the fact that $A' = \{x_0\}$, we have that the expression above of $A' \circ R_i$ corresponds to $I(A_i(x_0), B_i(y))$. Considering this latter expression, we can rewrite $\tilde{B}(y_0)$ as follows:

$$\tilde{B}(y_0) = \bigwedge_{i=1}^{n}(I(A_i(x_0), B_i(y_0))).$$

Selecting an appropriate implication I, the expression above for $\tilde{B}(y_0)$ can be expressed in terms of a weighted minimum (see Definition 3). In particular, the Kleene–Dienes implication $I(a,b) = \max(1 - a, b)$ is the one that makes this correspondence possible. Note that with this implication, $\tilde{B}(y_0)$ can be rewritten as

$$\tilde{B}(y_0) = \bigwedge_{i=1}^{n}(I(A_i(x_0), B_i(y_0))) = \bigwedge_{i=1}^{n} \max(1 - A_i(x_0), B_i(y_0)). \tag{2}$$

Therefore, when the weighting vector $\mathbf{u}$ is defined as $\mathbf{u} = (A_1(x_0), \ldots, A_n(x_0))$, the following equality holds:

$$\tilde{B}(y_0) = \mathrm{W}\,\mathrm{Min}_{\mathbf{u}}(B_1(y_0), \ldots, B_n(y_0)). \tag{3}$$

4.4. Using the Sugeno integral

The last two sections have shown that inference in fuzzy rule based systems can be formalized in terms of aggregation operators. We have seen that weighted max was used when rules are interpreted in a disjunctive manner, and that weighted min was used when rules are interpreted in a conjunctive manner. As the Sugeno integral is known to generalize both W Min and W Max, the output of a fuzzy inference system can be generally understood as the integration of the values $B_i(y_0)$ with respect to a fuzzy measure constructed from the values in the weighting vector $\mathbf{u} = (A_1(x_0), \ldots, A_n(x_0))$.

In particular, let $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$ and $\mu_{\mathbf{u}}^{\mathrm{w\,min}}$ be fuzzy measures with $\mu_{\mathbf{u}}^{\mathrm{w\,max}}(Z) = \max_{i \in Z} u_i$ and $\mu_{\mathbf{u}}^{\mathrm{w\,min}}(Z) = 1 - \max_{i \notin Z} u_i$ where $Z \subset X$ and $X := \{1, 2, \ldots, n\}$. Then, since $\mathrm{W\,Max}_{\mathbf{u}}(f) = \mathrm{SI}_{\mu_{\mathbf{u}}^{\mathrm{w\,max}}}(f)$ and $\mathrm{W\,Min}_{\mathbf{u}}(f) = \mathrm{SI}_{\mu_{\mathbf{u}}^{\mathrm{w\,min}}}(f)$, we can rewrite expressions (1) and (3), respectively, as follows:

$$\tilde{B}(y_0) = \mathrm{SI}_{\mu_{\mathbf{u}}^{\mathrm{w\,max}}}(B_1(y_0), \ldots, B_n(y_0)), \tag{4}$$

$$\tilde{B}(y_0) = \mathrm{SI}_{\mu_{\mathbf{u}}^{\mathrm{w\,min}}}(B_1(y_0), \ldots, B_n(y_0)), \tag{5}$$

where $\mathbf{u} = (A_1(x_0), \ldots, A_n(x_0))$.

The consequences of this formalization is that both conjunctive and disjunctive fuzzy rule based systems are expressed in a unified way in terms of the Sugeno integral. The solely difference between the two approaches is how the fuzzy measure is defined from **u**. In the case of disjunctive rules, a possibility measure $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$ is used. Instead, in the case of disjunctive rules, a necessity measure $\mu_{\mathbf{u}}^{\mathrm{w\,min}}$ is used.

Moreover, as it is known (see [7]) that $\cap_j(A' \circ R_j) \subseteq \cup_j(A' \circ R_j)$, it is easy to see that the two Sugeno integrals defined above (or, more precisely, the two fuzzy measures $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$ and $\mu_{\mathbf{u}}^{\mathrm{w\,min}}$ in conjunction with the Sugeno integral) define an interval for each y_0. It is clear that the use of other fuzzy measures would lead to other values for the certainty degree of y_0 (in, or around, the same interval).

An important aspect that cannot be skipped is that the weighting vectors used above are not, strictly speaking, possibilistic weighting vectors (or possibility distributions). This is so because, in general, **u** does not satisfy $\max u_i = 1$ as it is often the case that there is no i such that $A_i(x_0) = 1$. The practical consequences of this fact is that the aggregation operator $\mathbb{C}_{\mathbf{u}}$ that use them does not satisfy unanimity (i.e., it does not hold $\mathbb{C}_{\mathbf{u}}(a, a, \ldots, a) \neq a$). Nevertheless, this property is not a consequence of using the Sugeno integral but it is a property that already held for the weighted minimum and the weighted maximum.

The rewriting of the fuzzy inference system in terms of Sugeno integrals yields to an important consequence. While WMin and WMax assume independence between the values to be aggregated, the Sugeno integral does not require such independence. Therefore, the Sugeno integral is a natural operator to combine the conclusions of several rules in a fuzzy rule based system when such rules are not independent. Fig. 1 illustrates this situation.

Fig. 1 represents (left) the case of a fuzzy rule based system with two input variables x and y. The rules are assumed to follow a grid-like structure, a common structure in real-world fuzzy control applications. This is the case of fuzzy systems represented in a tabular form. In general, such systems are defined in terms of fuzzy partitions [7], one for the domain of each variable. In our case, let $\{A_i^x\}_i$ and $\{A_j^y\}_j$ be the fuzzy partitions of the domains of x and y, then for each pair (A_i^x, A_j^y) a fuzzy rule is defined. Accordingly, for *almost* any pair of input values (x_0, y_0), four rules are applied. This is the case, for example, when firing the rules for $x_0 = 2$ and $y_0 = 2$

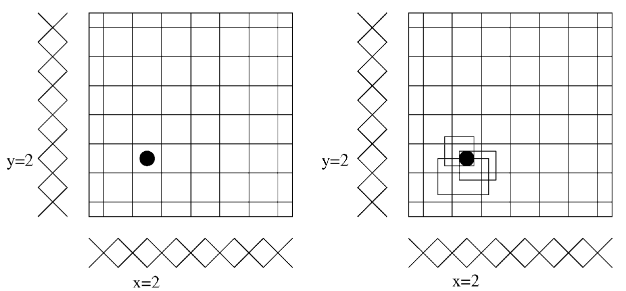

Fig. 1. Graphical representation of two different fuzzy inference systems with two input variables: regions correspond to fuzzy rules (the membership functions for each of the two variables is also given).

indicated with a circle in Fig. 1 (left). Exceptions to this situation correspond to the knots or the lines in the figure, when only one or two fuzzy rules are applied. Note that the regions in the figure represent a rule for which the sets A_i^x and A_j^y are satisfied for a degree of at least 0.5.

Instead, in the case represented on the right hand side, not all rules follow the grid-like structure and there is a region (around $x_0 = 2$, $y_0 = 2$) where the number of rules to be applied do not have an homogeneous structure. In fact, the number of rules to be applied depends on the values (x_0, y_0).

The regularity of the rules in systems following a grid-like structure makes the use of Sugeno integral with fuzzy measures $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$ or $\mu_{\mathbf{u}}^{\mathrm{w\,min}}$ adequate. In this case there are no major interactions between the outcomes of the rules. Each rule has its own *area of influence*. Instead, rules not following this pattern are usually not independent, as there are regions that *accumulate* several rules. In such regions, the outcome of the rules bias the outcome of the whole system. In this case, other fuzzy measures might be used to reduce the influence of such *accumulation* of rules and, thus, take into account the interaction among rules.

The example given below illustrate such situations.

Example 4. Let us consider the following set of four rules, each with one input variable x and one output value y, to model the relation (x, x^2):

R_1: **IF** x is 1 **THEN** y is 1,
R_2: **IF** x is 2 **THEN** y is 3.7,
R_3: **IF** x is 1.95 **THEN** y is 4.04,
R_4: **IF** x is 1.9 **THEN** y is 4.4.

Let us consider the following triangular fuzzy numbers to represent the fuzzy sets in the consequent: $(0.2, 1.00, 2.3)$, $(1.6, 3.7, 6.3)$, $(1.7, 4.04, 6.6)$ and $(1.8, 4.4, 6.8)$. Here, a triangular fuzzy number (a, b, c) stands for the triangular fuzzy set defined with the normal point b and the support (a, c).

It can be easily observed that rules R_2, R_3 and R_4 are redundant as they try to give information about the same region on the domain of x (i.e. the region around the value 2).

Now, let us consider the application of the rules to the input value $x_0 = 1.6$ considering that the rules are disjunctive. Assume that all rules are fired and that the degree of satisfaction of rules R_i ($\alpha_i = A_i(x)$) are $\alpha = (0.25, 0.625, 0.6875, 0.75)$. Then, the fuzzy set $\tilde{B}$ will include the outcomes of the four rules. Fig. 2 shows the result of firing such rules. This output has been computed following the description in Section 4.2 using the fuzzy measure $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$ defined in Table 6. Recall that this procedure corresponds to the Mamdani approach for applying fuzzy rules.

This figure shows that as rules R_2, R_3 and R_4 conclude all about a value near 4, such region has a larger influence in the output (a larger dark region) than the one that would be obtained by a single rule. It should be underlined that such influence can be positive (i.e., increasing the output value) or negative (i.e., decreasing the

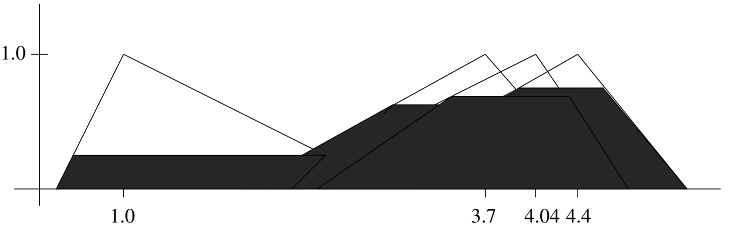

Fig. 2. Outcome of the rules.

Table 6
Fuzzy measure $\mu = \mu_{\mathbf{u}}^{\mathrm{w\,max}}$ for Example 4

$\mu(\{R_1, R_2, R_3, R_4\}) = 0.75$
$\mu(\{R_1, R_2, R_3\}) = 0.6875$
$\mu(\{R_1, R_2, R_4\}) = 0.75$
$\mu(\{R_1, R_3, R_4\}) = 0.75$
$\mu(\{R_2, R_3, R_4\}) = 0.75$
$\mu(\{R_1\}) = 0.25$
$\mu(\{R_2\}) = 0.625$
$\mu(\{R_4\}) = 0.75$
$\mu(\{R_1, R_2\}) = 0.625$
$\mu(\{R_1, R_3\}) = 0.6875$
$\mu(\{R_1, R_4\}) = 0.75$
$\mu(\{R_2, R_3\}) = 0.6875$
$\mu(\{R_2, R_4\}) = 0.75$
$\mu(\{R_3, R_4\}) = 0.75$
$\mu(\{R_3\}) = 0.6875$
$\mu(\emptyset) = 0$

output value). The sign depends on the shape and position of the membership functions.

The influence of the rules in the output can be quantified applying a defuzzification method to the fuzzy set. We have selected the center of gravity as such defuzzification method. The example described above with four rules leads to a defuzzified value equal to 2.8356. Instead, if only rules R_1 and R_3 were applied (with α_1 and α_2), the final output would be 2.5491. Alternatively, if we replace rules R_2, R_3 and R_4 by a rule with an average consequent fired with the average $(\alpha_2 + \alpha_3 + \alpha_4)/3$ we get a defuzzified value of 2.5521. Note that the ideal output for $x = 1.6$ equals to $1.6^2 = 2.56$. Thus, in this example, the redundancy of the rules bias the output towards larger values.

A way to solve the problem illustrated in Example 4, is to use the Sugeno integral but with a fuzzy measure different than $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$. In particular, and as an example, we can consider the fuzzy measure μ_1 defined in Table 7. This measure reduces the effect of the redundant rules. It has been defined so that $\mu_1(Z)$ is lower than μ when Z

Table 7
Fuzzy measure μ_1 to solve the inconveniences of Example 4

$\mu_1(\{R_1, R_2, R_3, R_4\}) = 0.75$
$\mu_1(\{R_1, R_2, R_3\}) = 0.4583$
$\mu_1(\{R_1, R_2, R_4\}) = 0.5$
$\mu_1(\{R_1, R_3, R_4\}) = 0.5$
$\mu_1(\{R_2, R_3, R_4\}) = 0.75$
$\mu_1(\{R_1\}) = 0.25$
$\mu_1(\{R_2\}) = 0.2083$
$\mu_1(\{R_4\}) = 0.25$
$\mu_1(\{R_1, R_2\}) = 0.25$
$\mu_1(\{R_1, R_3\}) = 0.25$
$\mu_1(\{R_1, R_4\}) = 0.25$
$\mu_1(\{R_2, R_3\}) = 0.4583$
$\mu_1(\{R_2, R_4\}) = 0.5$
$\mu_1(\{R_3, R_4\}) = 0.5$
$\mu_1(\{R_3\}) = 0.2292$
$\mu_1(\emptyset) = 0$

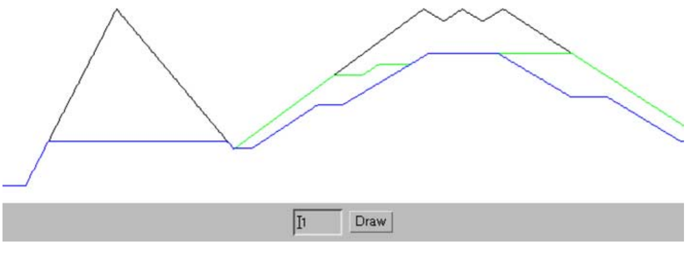

Fig. 3. Using Sugeno integral for combining fuzzy rules.

contains some of the $\{R_2, R_3, R_4\}$ but not all of them (in fact, it is proportional to the number of redundant rules in Z). Note e.g. that $\mu(\{R_1, R_3\}) = 0.6875$ in the original fuzzy measure but that $\mu_1(\{R_1, R_3\}) = 0.25$ in the new one. Similarly, $\mu(\{R_1, R_3, R_4\}) = 0.75$ while for the new measure $\mu_1(\{R_1, R_3, R_4\}) = 0.5$.

With such new fuzzy measure μ_1, the defuzzified value (for the input $x = 1.6$) equals to 2.4827, that is more similar to the goal 2.56 than the original 2.8356. Note that this value is also similar to the outcome when only one of the redundant rules is used. Fig. 3 illustrates the results of the combination using Sugeno integral with both the original measure μ and the alternative measure μ_1.

5. Interpreting the other integrals

Another well-known integral is the Choquet integral that generalizes, among other operators, the weighted mean. The main difference between the Choquet integral and the weighted mean is that the latter uses a weighting vector, that corresponds to a probability distribution, while the Choquet integral uses a fuzzy measure. Under this consideration, the fuzzy measure can be understood as a kind of probability distribution with some uncertainty. On this basis, and taking into account that the weighted mean can be understood as an expected value, the Choquet integral can be interpreted as a kind of expectation for such probability with uncertainty.

Twofold integrals were defined as a generalization of Choquet and Sugeno integrals. They correspond, in fact, to two-step fuzzy integrals: a Choquet integral of Sugeno integrals. Accordingly, such integrals can be interpreted in terms of the Choquet integral and the Sugeno integral. In particular, turning into the example of the rule based system, we can use the twofold to define a fuzzy inference system with randomness on the rules. This is detailed in the following example:

Example 5. Let us consider a rule based fuzzy inference system. Let $B_i(y_0)$ be the certainties that rules R_i assign to a particular value y_0. Then, $\mu_S(A)$ is the certainty assigned to the set of rules A. Naturally, $\mu_S(A)$ is computed from α_i (the degree in which rules x_i have been fired) either using $\mu_{\mathbf{u}}^{\mathrm{w\,max}}$, $\mu_{\mathbf{u}}^{\mathrm{w\,min}}$ or any other composite measure. Additionally, μ_C corresponds to some prior knowledge about the appropriatedness/accuracy of the rules. So, a probability distribution (or a fuzzy measure) is defined over the set of rules.

Then, to combine the values of $B_i(y_0)$ taking into account $\mu_S(A)$ and μ_C the twofold integral of $B_i(y_0)$ with respect to $\mu_S(A)$ and μ_C will be used.

6. Conclusions

In this paper we have considered the interpretation of the Sugeno integral. In particular, we have shown that in the Sugeno integral both the measure and the values being aggregated are in the same domain. Such values can be interpreted as *importances*, *reliabilities* or *certainties*. Several examples are given. In particular, we have shown that the Sugeno integral can naturally be applied to fuzzy inference systems. In fact, we have shown that the Sugeno integral is a natural extension of the operators used in fuzzy inference systems to aggregate the outcomes of the rules. Besides, we have outlined an interpretation for the Choquet and twofold integrals.

References

[1] P. Benvenuti, R. Mesiar, D. Vivona, Monotone set functions-based integrals, in: E. Pap (Ed.), Handbook of Measure Theory, Elsevier, 2002.
[2] T. Calvo, G. Mayor, R. Mesiar, Aggregation Operators, Physica-Verlag, 2002.
[3] T. Calvo, A. Mesiarová, L. Valásková, Composition of aggregation operators—one more new construction method, Kybernetika 39 (2003) 643–650.
[4] G. Choquet, Theory of capacities, Ann. Inst. Fourier 5 (1954) 131–296.
[5] D. Dubois, H. Prade, Weighted minimum and maximum operations, Inform. Sci. 39 (1986) 205–210.

[6] D. Driankov, H. Hellendoorn, M. Reinfrank, An Introduction to Fuzzy Control, Springer-Verlag, USA, 1996.
[7] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, UK, 1995.
[8] L. Magdalena, On the role of context in hierarchical fuzzy controllers, Int. J. Int. Syst. 17 (5) (2002) 471–493.
[9] R. Mesiar, A. Mesiarová, Fuzzy integrals, in: Proceedings of the Modeling Decisions for Artificial Intelligence, LNAI, 3131, 2004, pp. 7–14.
[10] T. Murofushi, M. Sugeno, An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure, Fuzzy Sets Syst. 29 (1989) 201–227.
[11] T. Murofushi, M. Sugeno, Fuzzy measures and fuzzy integrals, in: M. Grabisch, T. Murofushi, M. Sugeno (Eds.), Fuzzy Measures and Integrals: Theory and Applications, Physica-Verlag, 2000, pp. 3–41.
[12] T. Murofushi, M. Sugeno, Fuzzy *t*-conorm integral with respect to fuzzy measures: generalization of Sugeno integral and Choquet integral, Fuzzy Sets Syst 42 (1) (1991) 57–71.
[13] Y. Narukawa, V. Torra, Twofold integral and Multi-step Choquet integral, Kybernetika 40 (2004) 39–50.
[14] M. Sugeno, Theory of fuzzy integrals and its application, Doctoral Thesis, Tokyo Institute of Technology, 1974.
[15] E. Takahagi, On fuzzy integral representation in fuzzy switching functions, fuzzy rules and fuzzy control rules, in: IFSA, Proc. 8th IFSA World Congres, 1999, pp. 289–293.
[16] V. Torra, Twofold integral: a Choquet integral and Sugeno integral generalization, Butlletí de l'Associació Catalana d'Intel.ligència Artificial, vol. 29, pp. 13–19 (in Catalan), Preliminary version: IIIA Research Report TR-2003-08 (in English) 2003.
[17] E. Tunstel, M.A.A. de Oliveira, S. Berman, Fuzzy behavior hierarchies for multi-robot control, Int. J. Int. Syst. 17 (5) (2002) 449–470.
[18] A. Verkeyn, D. Bottledooren, B. De Baets, G. De Tré, Sugeno integrals for the modelling of noise annoyance aggregation, in: Proc. IFSA 2003, LNAI, vol. 2715, 2003, pp. 277–284.
[19] R.R. Yager, On ordered weighted averaging aggregation operators in multi-criteria decision making, IEEE Trans. SMC 18 (1988) 183–190.
[20] R.R. Yager, D.P. Filev, Essentials of Fuzzy Modeling and Control, John Wiley & Sons, New York, 1994.
[21] M. Yoneda, S. Fukami, M. Grabisch, Human factor and fuzzy science, in: K. Asai (Ed.), Fuzzy Science, Kaibundo, 1994, pp. 93–122 (in Japanese).
[22] V. Torra, Y. Narukawa, On the interpretation of some fuzzy integrals, LNAI 3131 (2004) 316–326.