Information Fusion 64 (2020) 92–98

Contents lists available at ScienceDirect

Information Fusion

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

Ordinal scale based uncertainty models for AI

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801

ARTICLE INFO

Keywords:
Linguistically expressed
Fuzzy measure
Ordinal information
Multi-source fusion

ABSTRACT

In human processed AI, HP-AI, we build our AI systems based on knowledge learned by human experts rather then that learned by artificial neural networks such as in the case of deep learning. The information provided by these human experts is typically linguistically expressed. In support of HP-AI we look at the properties of an ordinal scale, **S**, needed to model linguistically expressed quantitative information. Since fuzzy measures provide a very general structure for modeling uncertainty we look at ordinal fuzzy measures. We look at the Sugeno integral based on this ordinal **S** scale. We discuss the modeling of information about an uncertain variable using an ordinal scale. We look at the problem of multi-source in this ordinal environment.

1. Introduction

Artificial Intelligence has become one of the most important technologies in todays technology oriented world. Essentially we can identify two major paradigms for building AI models. The first, and currently the most popular, is based on machine processing of data and experiences. We shall refer to this as MP-AI. Central to this are neural network techniques such as deep learning [1]. In this approach data and other experiential information is feed into some digital device that uses this information to learn decision and other models. The second type of AI is based on the human processing of data and experiential information to learn decision models. We shall refer to this as HP-AI. Often HP-AI is based on human experience. While the models obtained using HP-AI are generally based on less data they are typically more sophisticated. In HP-AI we use logic, fuzzy sets, approximate reasoning and other soft computing technologies to build systems based on this human processed knowledge. In HP-AI we must extract from human experts their learned models and rules of thumb. As humans prefer language to express themselves this kind of human processed observations are generally linguistically expressed. Human language typically uses qualitative quantitative terms such as small, medium and big as well as tall and sort. Fuzzy sets [2–4] and other related technologies [5] can be used to model this kind of imprecise information. However, along another dimension the scales underlying the kinds if qualitative quantitative information provided by human information processes is generally of an ordinal nature. In support of HP-AI in this work we look at techniques for building ordinal scale based uncertainty models.

2. Ordinal scales

Let $\boldsymbol{S} = \{S_1, \ldots, S_q\}$ be a linguistically motivated ordinal scale such that $S_{j+1} > S_j$. Typically such a linguistic scale would be {Smallest, very small, small, medium, big, very big, biggest}. Here S_1 is our minimal element and S_q is our maximal element. At times we shall find it convenient to denote our minimal element S_1 as 0 and our maximal element as 1. Here we see $S_i > S_j$ if $i > j$. We shall say that the power of the scale is q. Here we have the ability to denote q different levels.

We can define a negation operator [6,7] on S. We define the negation operator N: $\mathbf{S} \rightarrow \mathbf{S}$ such that $N(S_j) = S_{q+1-j}$. We note that N has the following properties

1) $N(S_1) = S_q$
2) $N(S_q) = S_1$
3) $N(N(S_j)) = S_j$
4) If $S_i > S_j$ then $N(S_j) \geq N(S_i)$

We can define the binary operation of conjunction (Min), ∧, on **S** such that

$$S_i \wedge S_j = S_{i \wedge j} = S_{Min(i,j)}$$

and the binary operation of disjunction (Max), ∨, on **S** such that

$$S_i \wedge S_j = S_{i \wedge j} = S_{Max(i,j)}$$

We see that for any level S_j we have

$$S_j \wedge S_1 = S_1$$
$$S_j \vee S_1 = S_j$$
$$S_j \wedge S_q = S_j$$
$$S_j \vee S_q = S_q$$

When working with an ordinal scale such as **S** a useful tool is what we refer to as a unitor function, which maps the unit interval into **S**.

E-mail address: yager@panix.com

https://doi.org/10.1016/j.inffus.2020.06.010
Received 12 September 2019; Received in revised form 20 April 2020; Accepted 21 June 2020
Available online 28 June 2020
1566-2535/

Definition. Assume $\boldsymbol{S} = \{S_1, \dots, S_q\}$ is an ordinal scale. A ***unitor function*** U is a mapping
U: $[0, 1] \to \mathbf{S}$ such that

$$U(r) = S_i \quad \text{if } \frac{(i-1)}{q} \le r < \frac{i}{q} \text{ for } i = 1, 2\dots, q$$
$$U(1) = S_q$$

There are many functions and mathematical structures we can model using the ordinal scale **S** to provide associated parameters.

An ***aggregation function*** with respect to **S** is a function of $n > 1$ arguments f: $\mathbf{S}^n \to \mathbf{S}$ having the properties [8,9]

1) $f(S_1, \dots, S_1) = S_1$
2) $f(S_q, \dots, S_q) = S_q$
3) $f(a_1, \dots, a_n) \ge f(b_1, \dots, b_n)$ if $a_i \ge b_i$ for all i.

Two notable examples of aggregation functions are Min and Max. Here for $f =$ Min, we have $f(a_1, \dots, a_n) = \text{Min}_i[a_i]$ and for $f =$ Max, we have $f(a_1, \dots, a_n) = \text{Max}_i[a_i]$. The median is also an aggregation function on **S**.

Another notable example of an aggregation function on **S** is $f(a_1, \dots, a_n) = \underset{j=1\,to\,n}{Max}[w_j \wedge a_j]$ where w_j are a set of weights such 1) $w_j \in \mathbf{S}$ and 2) at least one $w_j = S_q$, that is $\underset{j=1\,to\,n}{Max}[w_j] = S_q$.

Closely related function is ***a basic Monotonic function*** f: $\mathbf{S} \to \mathbf{S}$ such that:

1) $f(S_1) = S_1$
2) $f(S_q) = S_q$
3) $f(a) \ge f(b)$ if $a > b$

One mathematical structure we can model using the ordinal scale **S** is to provide the associated parameters of a fuzzy set. Assume $X = \{x_1, \dots, x_n\}$ is a set of objects we can define a fuzzy subset F on X using the scale **S** to provide the membership grades. At times we refer to this as an **S**-fuzzy set to emphasize that the membership grades are drawn from **S**. Thus here we define the fuzzy set F so its membership grades $F(x_j)$ are such that $F(x_j) \in \mathbf{S}$. We say that F is normal if there exists at least one $x_j \in X$ such that $F(x_j) = S_q$. We say F is the null set $\emptyset$, if $F(x_j) = S_1$ for all x_j and we say F is the whole space if $F(x_j) = S_q$ for all x_j. Assume F is an **S**-fuzzy set we define its negation $\bar{F}$ as an **S**-fuzzy set such that $\bar{F}(x_j) = N(F(x_j))$. Assume F_1 and F_2 are two **S**-fuzzy sets on X we say $F_1 \subseteq F_2$ if $F_1(x_j) \le F_2(x_j)$ for all x_j. Furthermore we can define their intersection, $F_1 \cap F_2$, as an **S**-fuzzy set F_3 such that $F_3(x_j) = F_1(x_j) \wedge F_2(x_j)$ and we can define their union $F_1 \cup F_2$ as the **S**-fuzzy set F_4 such that $F_4(x_j) = F_1(x_j) \vee F_2(x_j)$.

In [10] Zadeh discussed the modeling of proportional linguistic quantifiers using fuzzy sets. Examples of linguistic quantifiers are "most", "about half", "more than α%". As noted by Zadeh these can be modeled as a fuzzy subset of the unit interval [0, 1]. A particularly important class of linguistic quantifiers are the monotonic linguistic quantifiers. For these linguistic quantifiers the membership function Q: $I \to I$ is monotonic. We can generalize this idea to the ordinal domain **S** using an **S**-quantifier where Q: $\mathbf{S} \to \mathbf{S}$ such that

1) $Q(S_1) = S_1$
2) $Q(S_q) = S_q$
3) Q is monotonic, $Q(a) \ge Q(b)$ if $a > b$

A useful concept associated with a classic fuzzy subset A on a finite space $X = \{x_1, \dots, x_n\}$ is the specificity of A, Sp(A) [11-16]. The specificity of A is a quantification of the degree to which A can be considered as a one point set. It is closely related to the concept of entropy and can be seen as an indication of the amount of information contained in the set A. In [17] we discussed a number of properties required of a formulation of the specificity of a fuzzy set:

1) Sp(A) assumes the largest value when A is a crisp set consisting of exactly one element
2) Sp(A) assumes its smallest value when A is the null set or $A = X$
3) If A and B are two normal fuzzy sets such that $A \subset B$ then $Sp(A) \ge Sp(B)$.

While there are many possible formulations for specificity in [17,18] Yager suggested the following strict linear formulation for the specificity of a fuzzy set A

$$Sp(A) = w_1 b_1 - \sum_{j=2}^{n} w_j b_j$$

where b_j is the j^{th} largest membership grade in A and w_j are a set of weights having the following properties:

1) $w_j \in [0, 1]$
2) $w_1 = 1$
3) $w_i \ge w_j$ for $i < j$
4) $\sum_{j=2}^{n} w_j = 1$

In [19] we suggested a formulation for specificity in the ordinal environment, $\mathbf{S} = \{S_1, \dots, S_q\}$. Assume A is an **S**-fuzzy set of $X = \{x_i, i = 1 \text{ to } n\}$. Let a_i for $i = 1$ to n be the membership grades of the elements in A. With b_j being the j^{th} largest membership grade in A let $L = \underset{j=1\,to\,n}{Max}[b_j \wedge U(\frac{j-1}{n-1})]$ using this we define $Sp(A) = N(L) \wedge b_1$. Let us look at the properties of Sp(A).

1) Assume A is one point set. Without loss of generality assume $A(x_1) = S_q$ and $A(x_j) = S_1$ for all $j \ne 1$. In this case $b_1 = S_q$ and $b_j = S_1$ for all other j. We see
$$L = \left(b_1 \wedge U(0) \vee \underset{j=2}{\overset{n}{Max}}\left[\left(b_j \wedge U(j-1/n-1)\right)\right]\right)$$
Since $b_j = S_1$ for $j \ne 1$ then $L = b_1 \wedge U(0)$ and since $U(0) = S_1$ then $L = S_1$. Hence $N(L) = S_q$ and $Sp(A) = S_q$.
2) Assume A is the null set, $b_i = S_1$ for all i including b_1 hence $Sp(A) = N(L) \wedge b_1 = S_1$.
3) Assume A is X. With $L = \underset{j=1\,to\,n}{Max}[b_j \wedge U(\frac{j-1}{n-1})]$ since $b_j = 1$ for all j then $L = \underset{j=1\,to\,n}{Max}[U(\frac{j-1}{n-1})] = U(1) = S_q$ and hence $N(L) = S_1$
4) Assume A and B are two normal ***S***-fuzzy sets such that $A \subseteq B$, that is $A(x_i) \le B(x_i)$ for all x_i. Assume e_j and f_j are the j^{th} largest membership grades in A and B respectively. Since $A \subseteq B$ then $f_j \ge e_j$ for each j. Since A and B are normal then $e_1 = f_1 = S_q$ and $Sp(A) = N(L_A)$ and $Sp(B) = N(L_B)$. Here $L_A = \underset{j}{Max}[e_j \wedge U(j - 1/n - 1)]$ and $L_B = \underset{j}{Max}[f_j \wedge U(j - 1/n - 1)]$, however since $f_j \ge e_j$ then $L_B \ge L_A$ from this we see $N(L_A) \ge N(L_B)$ and hence $Sp(A) \ge Sp(B)$. Thus we see the proposed definition of $Sp(A) = N(L) \wedge b_1$ has all the required properties of Sp(A).

3. S-Fuzzy measures

Assume $X = \{x_1, \dots, x_n\}$ is an arbitrary finite set we now define an ***S-fuzzy measure*** $\boldsymbol{\mu}$ on X. Here μ maps crisp subsets on X in to **S**. Specifically μ: $2^X \to \boldsymbol{S}$ such that

1) $\mu(\emptyset) = S_1$
2) $\mu(X) = S_q$
3) If $A \subset B$ then $\mu(B) \ge \mu(A)$

Here we shall simply refer to μ as a measure or ***S***-measure. If μ_1 and μ_2 are two ***S***-measures such that $\mu_1(A) \ge \mu_2(A)$ for all subsets A, we shall denote this as $\mu_1 \ge \mu_2$.

If μ is an **S**-measure we define the dual of μ, $\hat{\mu}$: $2^X \to \boldsymbol{S}$ such that $\hat{\mu}(A) = N(\mu(\bar{A}))$. We easily see that $\hat{\mu}$ is a ***S***-measure

1) $\hat{\mu}(\emptyset) = N(\mu(\bar{\emptyset})) = N(\mu(X)) = N(S_q) = S_1$
2) $\hat{\mu}(X) = N(\mu(\bar{X})) = N(\mu(\emptyset)) = N(S_1) = S_q$
3) Assume $A \subset B$. In this case $\bar{A} \supseteq \bar{B}$, $\hat{\mu}(A) = N(\mu(\bar{A}))$ and $\hat{\mu}(B) = N(\mu(\bar{B}))$

Since $\bar{A} \supseteq \bar{B}$ then $\mu(\bar{A}) \geq \hat{\mu}(\bar{B})$ and hence $N(\mu(\bar{A})) \leq N(\mu(\bar{B}))$ since $\hat{\mu}(A) = N(\mu(\bar{A}))$ and $\hat{\mu}(B) = N(\mu(\bar{B}))$ then $\hat{\mu}(B) \geq \mu(A)$.

We see that the dual of dual is the original measure $\hat{\hat{\mu}}(A) = \mu(A)$. This is easily shown, $\hat{\hat{\mu}}(A) = N(\hat{\mu}(\bar{A})) = N(N(\mu(A)) = \mu(A)$

We shall now introduce some notable examples of **S**-measures in X. The prototypical example of an ***S*** measure is the ***S***-possibility measure defined as follows. Let $\alpha_j \in \boldsymbol{S}$ for $j = 1$ to n and assume at least one $\alpha_j = S_q$. Here μ is defined so that $\mu(\{x_j\}) = \alpha_j$ and for any subset $A \subseteq X$ we have $\mu(A) = \underset{x_j \in A}{Max}[\alpha_j]$, then μ is an **S**-possibility measure. We easily see that $\mu(X) = S_q$ and if $A \subseteq B$ then $\mu(A) \leq \mu(B)$.

A special case of **S**-possibility measure μ is one focused on x_{j^*}. Here $\alpha_{j^*} = S_q$ and $\alpha_j = S_1$ for $j \neq j^*$. Here we see that for any $A \subseteq X$ we have $\mu(A) = S_q$ if $x_{j^*} \in A$ and $\mu(A) = S_1$ if $x_{j^*} \notin A$. Another special possibility measure is one where $\alpha_j = S_q$ for all j. This is called a uniform possibility measure.

A measure closely related to these possibility ***S***-measures are granular possibility ***S***-measures. Assume A_i for $i = 1$ to K are subsets of X and associated with each A_i is a value $\alpha_i \in \boldsymbol{S}$ such that at least one $\alpha_i = S_q$. Also associated with each A_i is an integer V_i so that $1 \leq V_i \leq Card(A_i)$. Also associated with each A_i is a function G_i: $2^X \to \{S_1, S_q\}$ so that for any subset $B \subseteq X$ we have

$$G_i(B) = S_q \text{ if } Card(B \cap A_i) \geq V_i$$
$$G_i(B) = S_1 \text{ if } Card(B \cap A_i) < V_i$$

Using the preceding we define an **S**-fuzzy measure μ on X such that $\mu(B) = \underset{i=1 \text{ to } K}{Max}[G_i(B) \wedge \alpha_i]$. We now show that μ is an **S**-measure on X:

1) $\mu(\emptyset)$: We have $Card(\emptyset \cap A_i) = 0$ for any A_i. In this case $G_i(\emptyset) = S_1$ and $\mu(\emptyset) = \underset{i=1 \text{ to } K}{Max}[S_1 \wedge \alpha_i] = S_1$
2) $\mu(X)$: We see $X \cap A_i = A_i$ and $Card(X \cap A_i) = Card(A_i)$. In this case $G_i(X) = S_q$ and $\mu(X) = \underset{i=1 \text{ to } K}{Max}[S_q \wedge \alpha_i] = S_q$
3) Assume B and D are two subsets of X so that $D \subseteq B$. Here then for this $Card(B \cap A_i) \geq Card(D \cap A_i)$ for all A_i. From that it follows that $G_i(B) \geq G_i(D)$ for any i. From this we get $\mu(B) = \underset{i=1 \text{ to } K}{Max}[G_i(B) \wedge \alpha_i]$ $\geq \underset{i=1 \text{ to } K}{Max}[G_i(D) \wedge \alpha_i] = \mu(D)$

Some notable examples of G_i and V_i are worth pointing out.

If all $V_i = 1$ then $\mu(B) = \underset{A_i \cap B \neq \emptyset}{Max}[\alpha_i]$

If all $V_i = Card(A_i)$ then $\mu(B) = \underset{i, A_i \subset B}{Max}[\alpha_i]$

If all $V_i = 0.5\ Card(A_i)$ then $\mu(B) = \underset{i, |A_i \subset B| \geq \frac{1}{2}|A_i|}{Max}[\alpha_i]$. Thus here if B contains at least half the elements in A_i then α_i contributes to $\mu(B)$.

More generally we see that each G is an ***S***-measure on X. This inspires us to consider the following general formulation. Assume μ_i for $i = 1$ to K are collection of ***S*** measures on X and associated with each μ is a value $\alpha_i \in \boldsymbol{S}$ such that at least one $\alpha_i = S_q$. We can define an S-measure μ on X, μ: $2^X \to \boldsymbol{S}$, such that $\mu(B) = \underset{i=1 \text{ to } K}{Max}[\alpha_i \wedge \mu_i(B)]$.

Another interesting example of an ***S***-measure on X is what we shall refer to as a prioritized measure [20-22]. Assume the elements in $X = \{x_1, \ldots, x_n\}$ are prioritized so that $x_1 \geq x_2 \geq x_3 \ldots \geq x_n$. We now define a type of measure to capture this prioritization. First we define

$$L_j = \{x_k / k = 1 \text{ to } j\} \text{ for } j = 1 \text{ to } n$$
$$L_0 = \emptyset$$

Here we see that $L_4 = \{x_1, x_2, x_3, x_4\}$ and $L_n = X$. We now associate with each subset L_j a value $\beta_j \in \boldsymbol{S}$ such that $\beta_0 = S_1$, $\beta_n = S_q$ and $\beta_j \geq \beta_{j-1}$. We now define the measure μ so that $\mu(A) = Max_j[\beta_j \wedge G_j(A)]$ where

$$G_j(A) = S_q \text{ if } L_j \subseteq A$$
$$G_j(A) = S_1 \text{ if } L_j \not\subset A$$

We see $\mu(A) = \beta_j$ where L_j is the largest L_j that is contained in A. We observe that $\mu(L_j) = \beta_j$ and $\mu(X) = S_q$ and $\mu(\emptyset) = S_1$.

Another notable class of **S**-fuzzy measures are the cardinality based ***S***-measures. Assume for $j = 0$ to n that $a_j \in \boldsymbol{S}$ where $a_{j+1} \geq a_j$ and $a_0 = S_1$ and $a_n = S_q$ then the **S**-measure μ defined such that $\mu(A) = a_{|A|}$ is called a cardinality based measure. We easily see that $\mu(\emptyset) = a_0 = S_1$, $\mu(X) = a_n = S_q$ and if $A \subset B$ then $\mu(B) = a_{|A|} \leq a_{|B|} = \mu(A)$. Further we observe that all sets A and B such that they have the same number of elements have the same measure independent of elements in the set.

We can point out some special cases of cardinality based ***S***-measure. One case μ_* is such that $\mu_*(A) = S_1$ for $A \neq X$ and $\mu_*(X) = S_q$. Another cardinality based measure is μ^* define such that $\mu^*(A) = S_q$ for $A \neq \emptyset$ and $\mu^*(\emptyset) = S_1$. Actually these two measures are the smallest and largest **S**-measures on X, that is $\mu_* \leq \mu \leq \mu^*$.

Another special case of cardinality-based measure is a generalization of these two cases called a tipping measure. Here $a_j = S_1$ for $j < K$ and $a_j = S_q$ for $j \geq K$. Here we note $K \leq n$.

4. Sugeno integral

An important and useful operation which we can define using the ordered space ***S*** is the Sugeno integral [23–25]. Assume $X = \{x_1, \ldots, x_n\}$ is a set of elements and μ is an ***S***-measure on X, μ: $2^X \to \boldsymbol{S}$. Let f: $X \to \boldsymbol{S}$ be a function that maps elements of X into **S**. The Sugeno integral of f with respect to μ is defined as

$$Sug_\mu(f) = \underset{j=1 \text{ to } n}{Max}\left[\mu(H_j) \wedge f(x_{\rho(j)})\right]$$

where ρ is an index function so that $\rho(j)$ is the index the element x_i having the j^{th} largest value of $f(x_i)$ and $H_j = \{x_{\rho(1)}, \ldots, x_{\rho(j)}\}$, it is the subset of X with j largest values for $f(x_i)$.

Note: If multiple elements in X have the same value for $f(x_i)$ the Sugeno integral is indifferent as how we order these tied value elements, we see this as follows. Assume m elements are tied and assume these are $x_{\rho(K+1)}, \ldots, x_{\rho(K+m)}$. Here

$$Sug_\mu(f) = \underset{j=1 \text{ to } K}{Max}\left[\mu(H_j) \wedge f(x_{\rho(j)})\right] \vee \underset{j=K+1 \text{ to } K+m}{Max}\left[\mu(H_j) \wedge f(x_{\rho(j)})\right]$$
$$\vee \underset{j=K+m+1 \text{ to } n}{Max}\left[\mu(H_j) \wedge f(x_{\rho(j)})\right]$$

We see that $f(x_{\rho(j)})$ for $j = K+1$ to $K+m$ are the same, let us denote this S_T. We now see that

$$\underset{j=K+1 \text{ to } K+m}{Max}\left[\mu(H_j) \wedge S_T\right] = \underset{j=K+1 \text{ to } K+m}{Max}\left[\mu(H_j) \wedge S_T = (H_{K+m}) \wedge S_T\right]$$

Here $H_{K+m} = \{x_{\rho(j)}, \ldots, x_{\rho(K+m)}\}$.

We can easily see that the Sugeno integral has the following properties

1) If $f(x_i) = S_k$ for all i, then $Sug_\mu(f) = S_k$, it is idempotent
2) As special cases of 1 we have
   a) All $f(x_i) = S_1$ then $Sug_\mu(f) = S_1$
   b) All $f(x_i) = S_q$ then $Sug_\mu(f) = S_q$
3) It is monotonic, if $f \geq \hat{f}$, $f(x_i) \geq \hat{f}(x_i)$ for x_i then $Sug_\mu(f) \geq Sug_\mu(\hat{f})$
4) The Sugeno integral can be expressed as a median function [8]

   $$Sug_\mu(f) = Med(f(x_1), \ldots, f(x_n), \mu(H_1), \ldots, \mu(H_n - 1))$$

5) The Sugeno integral is bounded: $\underset{i=1 \text{ to } j}{Min}[f(x_i)] \leq Sug_\mu(f) \leq \underset{i=1 \text{ to } j}{Max}[f(x_i)]$

From these properties we see that for any measure μ the Sugeno integral is an aggregation function. Also we see that the Sugeno integral is a mean function [8].

We can use the Sugeno integral to extend a measure to act on ***S***-fuzzy subsets. Assume μ is an ***S***-measure on $X = \{x_1, \ldots, x_n\}$. Thus for any crisp subset F on X we have $\mu(F)$. Let $\tilde{F}$ be a ***S***-fuzzy subset of X. Here then for each $x_i \in X$ we have $\tilde{F}(x_i) \in \boldsymbol{S}$. We now define $\mu(\tilde{F}) = Sug_\mu(f)$ where $f(x_i) = \tilde{F}(x_i)$. We see using this definition for the case where $\tilde{F}$ is a crisp subset of X, $\tilde{F}(x_i) \in \{S_1, S_q\}$ then $Sug_\mu(f) = \underset{j=1 \text{ to } n}{Max}[\mu(H_j) \wedge f(x_j)] = \mu(H_K 0$ where $H_K = \{x_{\rho(1))}, \ldots, x_{\rho(K)}\}$. Here $x_{\rho(j)}$, for $j = 1$ to K are the subset of elements in X for which $f(x_{\rho(j)}) = 1$ thus $H_K = F$ and hence $Sug_\mu(f) = \mu(F)$.

5. Combining measures for new measures

Theorem: Assume μ is an ***S***-measure on $X = \{x_1, \ldots, x_n\}$ and assume f is a basic monotonic function of **S**, f: $\mathbf{S} \to \mathbf{S}$. Then the **S** set function μ_1: $2^X \to S$ defined such that for all $A \in 2^X$ $\mu_1(A) = f(\mu(A)$ is an **S**-measure.

Proof. 1) If $A = \emptyset$ then $\mu_1(\emptyset) = f(\mu(\emptyset)) = f(S_1) = S_1$

2) If $A = X$ then $\mu_1(X) = f(\mu(X)) = f(S_q) = S_q$

3) Let $\mu_1(A) = f(\mu(A))$ and $\mu_2(B) = f(\mu(B)$ and if $A \subset B$ then $\mu(A) \leq \mu(B)$ and $\mu_1(A) = f(\mu(A) \leq f(\mu(B)) = \mu_1(B)$

Thus here we see we can create new ***S***-measures from μ by transforming $\mu(A)$ using a basic monotonic function.

A very important and useful property is noted in the following theorem.

Theorem: Assume for $j = 1$ to m that μ_j are a collection of ***S***-measures on X. If G is an ***S***-aggregation function of dimension m then the set function μ defined such that for all $A \subseteq X$

$\mu(A) = G(\mu_1(A), \mu_2(A), \ldots, \mu_m(A))$ is an ***S***-measure on X.

Proof.

1) $\mu(\emptyset) = G(\mu_1(\emptyset), \ldots, \mu_m(\emptyset)) = G(S_1, \ldots, S_1) = S_1$
2) $\mu(X) = G(\mu_1(X), \ldots, \mu_m(X)) = G(S_q, \ldots, S_q) = S_q$
3) Consider $\mu(A) = G(\mu_1(A), \ldots, \mu_m(A))$ and $\mu(B) = G(\mu_1(B), \ldots, \mu_m(B))$, if $B \subseteq A$ then $\mu_j(A) \geq \mu_j(B)$ for all j and hence $\mu(A) \geq \mu(B)$.

There are a large number of **S** aggregation functions. In applications the choice of aggregation function depends on the user's imperative for combining the measures. One class of aggregation functions are conjunctive functions [8], these are defined by having the additional property $G(a_1, \ldots, a_m) \leq \text{Min}(a_1, \ldots, a_m)$. These conjunctive functions are often used for implementation of the logical "anding" type aggregation. One property of these operations is that if $a_j = S_1$ for some argument then $G(a_1, \ldots, a_m) = S_1$. Some notable examples of these conjunctive aggregation functions are the Min and Drastic conjunction aggregator. For the Min we have $G(a_1, \ldots, a_m) = \text{Min}_i[a_i]$ and for the drastic operator we have

$$G_1(a_1, \ldots, a_m) = S_1 \text{ if } \text{Max}_j [a_j] < S_q$$
$$G_1(a_1, \ldots, a_m) = \text{Min}_i [a_i] \text{ if } \text{Max}_j [a_j] < S_q$$

Actually, the Min and the drastic conjunction are the bounding conjunction type aggregation operators, with the Min the biggest and the Drastic the smallest.

Another class of **S** aggregation operators are the disjunctive aggregation operators. These operator have the additional property that $G(a_1, \ldots, a_m) \geq \text{Max}_j[a_j]$. These are often used to implement an "oring" type aggregation. One property of these disjunctive type aggregation operators is that if $a_j = S_q$ for at least one argument then $G(a_1, \ldots, a_m) = S_q$. Two notable examples of these operators are the Max and the drastic disjunction, $G(a_1, \ldots, a_m) = \text{Max}_j[a_j]$ and

$$G_1(a_1, \ldots, a_m) = S_1 \text{ if } \text{Min}_j [a_i] < S_1$$
$$G_1(a_1, \ldots, a_m) = \text{Max}_j [a_j] \text{ if } \text{Min}_j (a_j] = S_1$$

Another class of **S** aggregation operators are mean operators, these are defined as having the additional of property $\text{Min}_j[a_j] \leq G(a_1, \ldots, a_m) \leq \text{Max}_j[a_j]$. Mean operator are idempotent, if G is a mean operator and all $a_j = a$ then $G(a_1, \ldots, a_m) = a$. A notable example of mean operators is the weighed max, $G(a_1, \ldots, a_m) = \underset{i=1\,to\,m}{Max}[C_i \wedge a_i]$, where $C_i \in \mathbf{S}$ and Max $[C_i] = S_q$. The median is an example of a mean operator. The Sugeno integral is another example of aggregation operator that is a mean.

The Sugeno integral provides a very general approach for the aggregation of a collection of **S**-measures on X to obtain another ***S***-measure on X. Assume μ_i for $i = 1$ to m are a collection of ***S***-measures on X. Let us denote $\boldsymbol{Z} = \{\mu_i$ for $i = 1$ to m$\}$ and let λ be an ***S***-measure on Z. Here we are interested in the set function μ^* on X that results from the Sugeno aggregation of the μ_i with respect to the ***S***-measure λ, we denote $\mu^* = \text{Sug}_\lambda(\mu_1, \ldots, \mu_m)$. If μ^* is defined such that for each subset $A \subseteq X$ we have $\mu^*(A) = \text{Sug}_\lambda(\mu_1(A), \ldots, \mu_m(A))$ then μ^* is a measure on X. Here

$$\mu*(A) = \text{Sug}_\lambda(\mu_1(A), \ldots, \mu_m(A)) = [\lambda(H_j) \wedge \mu_{\rho_A(j)}(A)]$$

where ρ_A is an index function so that $\rho_A(j)$ is the index of j^{th} largest $\mu_i(A)$ and $H_j = \{\mu_{\rho_A(j)}, \ldots, \mu_{\rho_A(j)}\}$ is the subset of **Z** consisting of the μ_i with the j largest values for $\mu_i(A)$. We emphasize that $\mu^*(A)$ must be calculated for all subsets $A \subseteq X$.

Some notable examples of λ are worth pointing out. If λ is the Min measures then $\mu^*(A) = \underset{i}{Min}[\mu_i(A)]$ and if λ is Max measure then $\mu^*(A) = \underset{i}{Max}[\mu_i(A)]$. If λ is a possibility measure on **Z** with $\alpha_i \in S$ the possibility of μ_i then $\lambda(H_j) = \underset{k=1\,to\,j}{Max}[\alpha_{\rho_A(k)}]$ and hence

$$\mu*(A) = \left[\underset{j=1\,to\,m}{Max}\ \underset{K=1\,to\,j}{Max}\left[\alpha_{\rho_A(K)}\right] \wedge \mu_{\rho_A(j)}(A)\right]$$

Finally if λ is a cardinality based measure with parameters $w_j \in \mathbf{S}$ where $w_0 = S_1$, $w_m = S_q$ and $w_{j+1} \geq w_j$ then $\lambda(H_j) = w_j$ and hence $\mu^*(A) = \underset{j=1\,to\,m}{Max}[w_j \wedge \mu_{\rho_A(j)}(A)]$

6. Modeling information about an uncertain variable using the S scale

Assume V is an uncertain variable that takes its value in the space $X = \{x_1, \ldots, x_n\}$. A general structure for representing information about a variable V that takes its value in a space X is a measure μ on the space X. In this representation for any subset A of X, $\mu(A)$ indicates the anticipation of finding the value of V in A [26, 27]. Since we only have available the **S** scale we must use an S-measure on X. Here for any subset $A \subseteq X$ we have $\mu(A) \in \mathbf{S}$ as the anticipation that the value of V is in A. For this measure $\mu(\emptyset) = S_1$, $\mu(X) = S_q$ and for any two subsets $A \subset B \subseteq X$ we have $\mu(A) \leq \mu(B)$. The anticipation of finding V in B is at least as much as finding V in A. The knowledge that $V = x_K$ is represented as a Dirac **S**-measure. Here for x_K in A we have $\mu(A) = S_q$ and for $x_K \notin A$ we have $\mu(A) = S_1$.

Another notable measure for representing information about V is a possibility measure μ. Here for each $x_i \in X$, we have $\mu(\{x_i\}) = \pi_i \in \mathbf{S}$ denoted the possibility of x_i. For this measure for any subset $A \subseteq X$, $\mu(A) = \underset{x_i \in A}{Max}[\pi_i]$. Since $\mu(X) = 1$ we require that at least one $\pi_i = 1$.

We note that a probability measure is not available to us as this requires a scale that allows addition. We can associate with V a cardinality-based measure μ. Here we assume a_i is a set of parameters for $i = 0$ to n with $a_i \in \mathbf{S}$ and $a_0 = S_1$, $a_n = S_q$ and $a_i \leq a_{i+1}$. For this measure $\mu(A) = a_{|A|}$. In this case the anticipation of a set A depends on how many elements are in A.

Assume V is an uncertain variable that takes its value in the space $X = \{x_i, i = 1 \text{ to } n\}$. Assume our knowledge about V is expressed via an **S**-measure μ on X. In addition let f be a function of V such that f: $X \to \mathbf{S}$. A natural question is to find expected value of V, EV(V). Here we can use the Sugeno integral

$$EV(V) = \text{Sug}_\mu(f) = \underset{j=1\,to\,n}{Max}\left[\mu(H_j) \wedge f(x_\rho(j))\right]$$

where ρ is an index function such that $\rho(j)$ is the index of the element x_i with the j^{th} largest value for $f(x_i)$ and $H = \{x_{\rho(k)}$, for $k = 1$ to j$\}$, the subset of X with the j largest values of $f(x_i)$.

Assume μ is an S-Dirac measure focused at x_L in this case

$$EV(V) = \text{Sug}_\mu(f) = \underset{j=1\,to\,n}{Max}\left[\mu(H_j) \wedge f(x_\rho(j))\right],$$

however here $\mu(H_j) = S_q$ if $x_L \in H_j$ and $\mu(H_j) = S_1$ if $x_L \notin H_j$. Assume $f(x_L)$ is the M^{th} largest value of $f(x_i)$. In this case $\mu(H_j) = S_q$ for $j \geq M$ and $\mu(H_j) = S_1$ for $j < M$. In this case

$$EV(V) = \underset{j=M\,to\,n}{Max}\left[S_q \wedge f(x_{\rho(j)}\right] = f(x_{\rho(M)}) = f(x_L).$$

Assume now μ is a cardinality based measure with parameters a_i, in the case

$EV(V) = Sug_\mu(f) = \underset{j=1\,to\,n}{Max}\,[a_j \wedge f(x_{\rho(j)})]$

This is what Yager [28] referred to as the ordered OWA operator.

Assume μ is a possibility measure on X where $\mu(\{x_i\}) = \pi_i$. In this case $\mu(H_j) = \underset{k=1\,to\,j}{Max}\,[\pi_{\rho(k)}]$ and we have

$EV(V) = Sug_\mu(f) = \underset{j=1\,to\,n}{Max}\,[(\mu(H_j) \wedge f(x_{\rho(j)}))] = \underset{j=1\,to\,n}{Max}\,[(\underset{k=1\,to\,j}{Max}\,[\pi_{\rho(k)}] \wedge f(x_{\rho(j)}))]$

$Sug_\mu(f) = (\pi_\rho(_1) \wedge f(x_{\rho(1)})) \vee ((\pi_{\rho(1)} \vee \pi_{\rho(2)}) \wedge f(x_{\rho(2)})) \vee \underset{j=3\,to\,n}{Max}\,[(\underset{k=1\,to\,j}{Max}\,[\pi_{\rho(k)}] \wedge f(x_{\rho(j)}))]$.

We see that $\pi_{\rho(1)} \wedge f(x_{\rho(2)}) \le \pi_{\rho(1)} \wedge f(x_{\rho(1)})$ hence

$Sug_\mu(f) = \underset{j=1\,to\,2}{Max}\,(\pi_{\rho(j)} \wedge f(x_{\rho(j)})) \vee \underset{j=3\,to\,n}{Max}\,[\underset{k=1\,to\,j}{Max}\,(\pi_{\ \rho(k)} \wedge f(x_{\rho(j)}))]$

We now see that $(\pi_{\rho(1)} \vee \pi_{\rho(2)}) \wedge f(x_{\rho(3)}) \le \underset{j=1\,to\,2}{Max}\,(\pi_{\rho(j)} \wedge f(x_{\rho(j)}))$ hence

$Sug_\mu(f) = \underset{j=1\,to\,3}{Max}\,(\pi_{\rho(j)} \wedge f(x_{\rho(j)})) \vee \underset{j=4\,to\,n}{Max}\,[\underset{k=1\,to\,j}{Max}\,(\pi_{\ \rho(k)} \wedge f(x_{\rho(j)}))]$

Continuing in this manner we get that

$EV(V) = Sug_\mu(f) = \underset{j=1\,to\,n}{Max}\,[(\pi_{\rho(j)} \wedge f(x_{\rho(j)}))]$.

7. Multi-Source fusion

In the preceding we considered the situation where V is an uncertain variable that takes its value in the space $X = \{x_1, \ldots, x_n\}$ and our knowledge about V is expressed via an ***S***-measure μ on X. In many cases we have multiple pieces of information about the variable V. More particularly, assume we have r pieces of information V *is* μ_k for $k = 1$ to r where each μ_k is an **S**-measure. If we have some prescribed aggregation formulation G for combining these r pieces of information where G can be expressed as an aggregation function, then our fused measure is $\mu^* = G(\mu_1, \mu_2, \ldots\ldots, \mu_r)$. As we have previously shown our fused measure μ^* in this case is such that for each $A \subset X$, $\mu^*(A) = G(\mu_1(A), \ldots, \mu_r(A))$.

Here we shall consider the case where we have no prescribed formulation for combining these r pieces of information. We observe that one objective in fusing these r sources of information is to obtain a fused **S**-measure that is informative. The most informative measure μ is one that tells us that one value, x_K, is completely possible and all other x_j are impossible. This can be captured by a Dirac measure μ_K such that $\mu_K(A) = S_q$ for any A such that $x_K \in A$ and $\mu_K(A) = S_1$ for any subset A such that $x_K \notin A$. Thus measure tells us the x_K is the value of V.

Given some proposed fused measure μ^*, to determine how informative it is, we must find its closeness to some μ_K. In general the problem of finding the closeness of two measures is not easy, however, the special structure of a Dirac measure μ_K allows us to attain a formula to accomplish this. Consider the measure μ_K on 2^X. We can partition 2^X into two collections of subsets of X, E_K^+ and E_K^-. Here E_K^+ consists of the subsets of X that contain x_K and E_K^- consists of all the subsets that do not contain x_K. Furthermore, for any $A \in E_K^+$ we have $\mu_K(A) = S_q$ and for any $A \in E_K^-$ we have $\mu_K(A) = S_1$. Here for our proposed fused value μ^* to be close to μ_K we desire for any $A \in E_K^+$ that $\mu^*(A) = S_q$ and for any $A \in E_K^-$ we desire $\mu^*(A) = S_1$.

We also observe that the each of the subsets E_K^+ and E_K^- can be partially ordered. In particular, $\{x_K\}$ is the smallest set in E_K^+ and hence $\mu^*(\{x_K\}) = \underset{A \in E_K^+}{Min}\,[\mu^*(A)]$ and X - $\{x_K\}$ is the largest set in E_K^- and hence $\mu^*(X - \{x_K\}) = \underset{A \in E_K^-}{Max}[\mu^*(A)]$. We see that for μ^* to be close to μ_K we desire all $A \in E_K^+$ to be close to S_q and all $A \in E_K^-$ to be close to S_1, as small as possible. Using these observations we see

$$Close(\mu^*, \mu_K) = \underset{A \in E_K^+}{Min}\,[\mu^*(A)] \wedge N(\underset{A \in E_K^-}{Max}\,[\mu^*(A)])$$

$$Close(\mu^*, \mu_K) = \mu^*(\{x_K\}) \wedge N(\mu^*\ (X - \{x_K\}))$$

To find the *informativeness* of a proposed fused measure μ^*, $Inf(\mu^*)$, we calculate

$$Inf(\mu^*) = \underset{x_K \in X}{Max}[Close(\mu^*, \mu_K)]$$

$$Inf(\mu^*) = \underset{x_K \in X}{Max}\ [(\mu^*(\{x_K\}) \wedge N(\mu^*(X - \{x_K\}))]$$

Table 1

$K(Ricness)$	Inf	$Fused\ Value$
r	$Inf(\mathfrak{I}_r)$	$\mu_{rj_r^*}$
$r-1$	$Inf(\mathfrak{I}_{r-1})$	$\mu_{r-1j_{r-1}^*}$
$r-2$		
.		
.		
.		
1	$Inf(\mathfrak{I}_1)$	$\mu_{1j_1^*}$

Here we shall suggest a **C**rude **Q**uantification of the **I**nformativeness of a measure μ, $CQI(\mu)$. Assume μ is an **S**-measure on $X = \{x_1, \ldots, x_n\}$ and let $a_i = \mu(\{x_i\})$. Let g be an index function so that g(j) is the index of the x_i with the j^{th} largest value of a_i, thus $a_{g(j)}$ is the j^{th} largest value $\mu(\{x_i\})$. In particular $a_{g(1)}$ is the largest value of a_i and $a_{g(2)}$ is the second largest value of a_i. Here we see $a_{g(1)} = \underset{j=1\,to\,n}{Max}\,[a_i]$. We now define our crude quantification of the informativeness of measure μ as

$CQI(\mu) = a_{g(1)} \wedge N(a_{g(2)})$.

In the preceding we provided a method for determining the information associated with a proposed fused value. We now turn to the issue of generating a proposed fused value.

Assume $R = \{\mu_i, i = 1$ to r$\}$ are sources of value for V. One approach to fusing these is to take the conjunction of some subset of R. Assume $F \subseteq R$ is a subset of source values for V, the fused value μ corresponding to the conjunction of these sources from R is

$$\mu_F = \underset{\mu_i \in F}{Min}\,[\mu_i]$$

where $\mu_F(A) = \underset{\mu_i \in F}{Min}[\mu_i(A)]$. Here our fused value is trying to satisfy all the sources in F.

In addition to the informativeness of the fused value, $Inf(\mu_F)$, a second feature of a good fused value is that it is based upon many of elements in R we shall refer to this as the ***Richness*** of the fusion, $Rich(\mu_F)$. Here we note $Rich(\mu_F) \approx Card(F)$. The larger the cardinality of F, the richer the fusion. A good fused value μ_F is one that has a large $Inf(\mu_F)$ and large Card(F). At times we shall find it convenient to use $Rich(\mu_F)$ or Rich(F) synonymously.

Let $\mathfrak{I}_K$ denote the set of all subsets of R having cardinality K

Example: Assume $R = \{\mu_1, \mu_2, \mu_3\}$ then

$\mathfrak{I}_1 = \{\{\mu_1\}, \{\mu_2\}, \{\mu_3\}\}$
$\mathfrak{I}_2 = \{\{\mu_1, \mu_2\}, \{\mu_1, \mu_3\}, \{\mu_2, \mu_3\}\}$
$\mathfrak{I}_3 = \{\{\mu_1, \mu_2, \mu_3\}\}$

We see the following properties of Rich are desired

1) If F and $F^* \in \mathfrak{I}_K$ then $Rich(\mu_F) = Rich(\mu_{F^*})$, all fused values based on sets of the same cardinality have the same richness. We t refer to this as $Rich\{\mathfrak{I}_K\}$
2) If $K_2 > K_1$ then if $F_1 \in \mathfrak{I}_{K_1}$ and $F_2 \in \mathfrak{I}_{K_2}$ then $Rich(F_2) > Rich(F_1)$. Thus $Rich(\mathfrak{I}_{K_2}) \ge Rich\ (\mathfrak{I}_{K_1})$.

Let $F_{Kj} \in \mathfrak{I}_K$, it is a subset of K sources from R. Let us denote μ_{Kj} as the fused value based on the conjunction of sources in F_{Kj}. Our objective is to select the fused value having the largest informativeness and richness. Assume $Inf(\mu_{Kj_K^*}) = \underset{F_{Kj} \in \mathfrak{I}_K}{Max}\ [Inf(\mu_{Kj})]$. If we had to select a fused value from $\mathfrak{I}_K$ it would be $\mu_{Kj_K^*}$ since all $F_{Kj} \in \mathfrak{I}_K$ have the same Richness and this fused value has the largest informativeness. Let us denote $Inf(\mathfrak{I}_K) = Inf(\mu_{Kj_K^*})$.

Construct the following the table

Construct from Table 1 a Table 2 that is a subset of the rows in Table 1 using the following rule. A row j in Table 1 appears in Table 2 if there is no row $K > j$ such that $Inf(\mathfrak{I}_K) > Inf(\mathfrak{I}_j)$.

Operational we form Table 2 as follows. We place a row in Table 2 if no row above it in Table 1 has a larger informativeness. We note for sure the top row in table 1 appears in Table 2.

Table 2

$K(Richnesss)$	Inf	$Fused\ Value$
r	$Inf(\mathfrak{I}_r)$	$\mu_{rj_r^*}$

The fused values in Table 2 are the possible optimal fused values, best Richness and best Informativeness. At this point it is up to the responsible decision maker to select a fused value from Table 2 by looking at its Richness and Informativeness.

In order to simplify the burden on the decision maker we can suggest the following procedure to associate with each row in Table 2 a unique value which can be used to select the best fused value.

We associate with each K a unique value in $\boldsymbol{S}$. Thus let H(K) $\in \boldsymbol{S}$ indicate the Richness value associated with a fused based on K sources. We note that H should have the following properties:

1) $H(r) = S_q$
2) If $K_2 > K_1$ the $H(K_2) \geq H(K_1)$

In the following we shall denote a row in Table 2 by its K value. Here we shall associate with each row in Table 2 a unique value, D(K) based on its Richness and Informativeness

$D(K) = Min[Inf(\mathfrak{I}_K), H(K)]$

We then select as our fused value the $\mu_{Kj_K^*}$ corresponding to the K with the maximal value for D(k)

The one open question is having to obtain the function H(K). The choice of H(K) is very subjective other than the requirements that 1) $H(r) = S_q$ and 2) If $K_2 > K_1$ then $H(K_2) \geq H(K_1)$. A reasonable choice is a quasi-linear formula for H. Here

$$H(K) = S_i \text{ if } \frac{(i-1)}{q} \leq \frac{K}{r} \leq \frac{i}{q}.$$

Thus given the value of K associated with a row in Table 2 then we calculate H(K) using the above. Here emphasize we are saying the above form for H(K) is a reasonable choice for H(K) not the only choice.

8. Granular information

Assume V is a variable with domain $X = \{x_1, \ldots, x_n\}$. Our interest is in the value of variable V. Let $\boldsymbol{F} = \{F_1, \ldots, F_m\}$ be a granulation of the space X, each $F_k \subseteq X$. Here we assume $\bigcup_{k=1}^{m} F_k = X$ but the F_k are not necessary disjoint. Let W be a related variable that takes its value in the space **F**. Assume the information available to us concerns the value of the variable W. In particular we have an **S**-measure μ on the space **F** so for any subset E of **F**, μ(E) is the anticipation that the value of W lies in E. Our interest is in using this available information about W to tell us something about the value of V.

Assume A and all of the F_j be crisp subsets of X. What is clear in this case is that if $W = F_K$, where F_K is also a crisp subset of X then

$Poss(V \text{ is in } A/W = F_K) = 1, S_q, \text{ if } A \cap F_K \neq \emptyset$
$Poss(V \text{ is in } A/W = F_K) = 0, S_q, \text{ if } A \cap F_K = \emptyset$

Furthermore

$Cert(V \text{ is in } A/W = F_K) = 1, S_q, \text{ if } F_K \subseteq A$
$Cert(V \text{ is in } A/W = F_K]) = 0, S_q, \text{ if } F_K \not\subset A.$

In the more general setting where A and the sets F_j are $\boldsymbol{S}$-fuzzy sets of X then using Zadeh definitions [29,30]

$$Poss(V \text{ is in } A/W = F_K) = \underset{x_i}{Max}[A(x_i) \wedge F_K(x_i)]$$
$$Cert(V \text{ is in } A/W = F_K) = \underset{x_i}{Min}[A(x_i) \vee N(F_K(x_i))]$$

Here we can say that Truth(V is in $A/W = F_K$) lies in an interval bounded below by the certainty and above by the possibility,

Truth(V is in $A/W = F_K$) $\in$ [Cert(V is in $A/W = F_K$), Poss(V is in $A/W = F_K$)].

We now more generally consider the situation we do not know the exact value of W, all we have is a measure μ on **F** as our knowledge about the value of W. Here we must provide an expected like value for the possibility and certainty, EV(Poss(V is in A/W is μ) and EV(Cert(V is in A/W is μ). Since our knowledge involves a measure μ and is expressed involving parameters from **S** we must use the Sugeno integral to express the expected value. In particular

EV(Poss(V is in A/W is μ) = $Sug_\mu[Poss(A/F_K] \text{ for } K = 1 \text{ to m}]$
EV(Cert(V is in A/W is μ) = $Sug_\mu[Cert(A/F_K] \text{ for } K = 1 \text{ to m}]$

Using our preceding discussion about the Sugeno integral we have $Sug_\mu[Poss(A/F_K) \text{ for } K = 1 \text{ to m}] = \underset{j=1\,to\,m}{Max}[(\mu(H_j) \wedge Poss(A/F_{\rho(j)}))]$ where ρ is an index function so that ρ(j) is the index of the F_k with the j^{th} largest value of Poss[A/F_K] and $H_j = \{F_{\rho(1)}, \ldots, F_{\rho(j)}\}$, the subset of elements in **F** with the j largest values of Poss(A/F_K). Similarly $Sug_\mu[Cert(A/F_K), K = 1 \text{ to m}] = \underset{j=1\,to\,m}{Max}[(\mu(G_j) \wedge Cert(A/F_{\tau(j)}))]$ where τ is an index function so that τ(j) is the index of the F_k with the j^{th} largest value of Cert[A/F_K] and $G_j = \{F_{\tau(1)}, \ldots, F_{\tau(j)}\}$. Using the terminology of Shafer [31] we refer to these as $\boldsymbol{S}$-plausibility(A) and $\boldsymbol{S}$-belief(A). Thus we have

EV(Truth V is A /W is μ) $\in$[***S**-belief*(A), ***S**-plausibility*(A)]

We shall look at the preceding for some special cases. Consider the case where A and all the granular sets, the F_K, are crisp sets. For this situation Poss(A/F_K) and Cert(A/F_K) are all either S_q or S_1. Thus in this case

$$Pl_\mu(A) = \underset{j=1\,to\,m}{Max}[(\mu(H_j) \wedge Poss(A/F_{\rho(j)}))] = \mu(H^P)$$

where H^P is the subset of **F** consisting of all the granular sets F_k so that $F_k \cap A \neq \emptyset$, thus $H^P = \{F_k/F_k \cap A \neq \emptyset\}$. Similarly $Bel_\mu = \underset{j=1\,to\,m}{Max}[(\mu(G_j) \wedge Cert(A/F_{\tau(j)}))] = \mu(H^B)$ where $H^B = \{F_k/F_k \subseteq A\}$.

Another special case is where A is one element in X, $A = \{x^*\}$. In case $Poss(A/F_k) = F_k(x^*)$ and $Cert(A/F_k) = \underset{x_i \neq x^*}{Max}[N(F_k(x_i))] = N(\underset{x_i \neq x^*}{Min}[F_k(x_i)])$

9. Conclusion

We first looked at the properties of an ordinal scale, **S**, needed to model linguistically expressed quantitative information. Since fuzzy measures provide a very general structure for modeling uncertainty we looked at **S** fuzzy measures. We looked at the Sugeno integral based on this ordinal **S** scale. We discussed the modeling of information about an uncertain variable using an ordinal scale. We turned to the problem of multi-source in this ordinal environment.

Author statement

I am sole author and I submit this revision.

Declaration of Competing Interest

There is no conflict of interest.

References

[1] T.J. Sejnowski, The Deep Learning Revolution, MIT Press, Cambridge, MA, 2018.
[2] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[3] L.A. Zadeh, Fuzzy logic and approximate reasoning, Synthese 30 (1975) 407–428.
[4] L.A. Zadeh, Approximate reasoning based on fuzzy logic, in: Proceedings of Sixth International Conference on Artificial Intelligence, Tokyo, 1979, pp. 1004–1010.
[5] J. Kacprzyk, D.P. Filev, G. Beliakov, R. Yager (Ed.), Springer, Heidelberg, 2017.
[6] R.R. Yager, On the measure of fuzziness and negation part I: membership in the unit interval, Int. J. of General Systems 5 (1979) 221–229.
[7] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, Upper Saddle River, NJ, 1995.
[8] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide For Practitioners, Springer, Heidelberg, 2007.

[9] R. Mesiar, A. Kolesárová, T. Calvo, M. Komornikova, A review of aggregation functions, in: H. Bustince, F. Herrera, J. Montero (Eds.), Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, Eds., Springer, Heidelberg, 2008, pp. 121–144. ed.
[10] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Computing and Mathematics with Applications 9 (1983) 149–184.
[11] M. Higashi, G.J. Klir, Measures of uncertainty and information based on possibility distributions, Int. J. of General Systems 9 (1983) 43–58.
[12] R.R. Yager, Entropy and specificity in a mathematical theory of evidence, Int. J. of General Systems 9 (1983) 249–260.
[13] D. Dubois, H. Prade, A note on measures of specificity for fuzzy sets, Int J Gen Syst 10 (1985) 279–283.
[14] G.J. Klir, Generalized information theory, Fuzzy Sets and Systems 40 (1991) 127–142.
[15] R. Gonzalez del Campo, L. Garmendia, R.R. Yager, Specificity for interval-valued fuzzy sets, Int J of Computational Intelligence Systems 5 (2012) 452–459.
[16] N. Marin, G. Rivas-Gervilla, D. Sanchez, R.R. Yager, On families of bounded specificity measures, in: Proceedings of FUZZ-IEEE 2017 Naples, Italy, 2017.
[17] R.R. Yager, Default knowledge and measures of specificity, Inf Sci (Ny) 61 (1992) 1–44.
[18] R.R. Yager, On the specificity of a possibility distribution, Fuzzy Sets and Systems 50 (1992) 279–292.
[19] R.R. Yager, Ordinal measures of specificity, Int. J. of General Systems 17 (1990) 57–72.
[20] R.R. Yager, Prioritized aggregation operators, Int J of Approximate Reasoning 48 (2008) 263–274.
[21] R.R. Yager, E. Walker, C. Walker, A prioritized measure for multi-criteria aggregation and its Shapley index, in: Proceedings of the North American Fuzzy Information Processing Society Annual Meeting, El Paso, TX, 2011, pp. 158–161.
[22] R.R. Yager, On prioritized multiple criteria aggregation, IEEE Transactions on Systems, Man and Cybernetics: Part B 42 (2012) 1297–1305.
[23] M. Sugeno, Fuzzy measures and fuzzy integrals: a survey, in: M.M. Gupta, G.N. Saridis, B.R. Gaines (Eds.), Fuzzy Automata and Decision Process, eds., North-Holland Pub, Amsterdam, 1977, pp. 89–102.
[24] T. Murofushi, M. Sugeno, Fuzzy measures and fuzzy integrals, in: M. Grabisch, T. Murofushi, M. Sugeno (Eds.), Fuzzy Measures and Integrals, Eds., Physica-Verlag, Heidelberg, 2000, pp. 3–41. ed.
[25] M. Grabisch, C. Labreuche, A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, Ann Oper Res 175 (2010) 247–290.
[26] R.R. Yager, A measure based approach to the fusion of possibilistic and probabilistic uncertainty, Fuzzy Optimization and Decision Making 10 (2011) 91–113.
[27] R.R. Yager, Uncertainty modeling using fuzzy measures, Knowl Based Syst 92 (2016) 1–8.
[28] R.R. Yager, Applications and extensions of OWA aggregations, Int J Man Mach Stud 37 (1992) 103–132.
[29] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.
[30] L.A. Zadeh, Eds., Advances in Fuzzy Set Theory and Applications, Amsterdam: North-Holland, 1979, pp. 3–18. ed.
[31] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, N.J., 1976.