Imprecise Sugeno integral based decision

Agnes Rico[2], Florentin Kucharczak[1], and Olivier Strauss[1]

[1]LIRMM, Univ. Montpellier, Montpellier, France, strauss@lirmm.fr
[2]ERIC, Univ. Lyon 1, Lyon, France, agnes.rico@univ-lyon1.fr

Abstract. One of the main difficulties when it comes to aggregating partial evaluations using tools such as Sugeno integral, is having access to precise knowledge of the capacity to perform this aggregation. In this article, we propose an intervallist extension of Sugeno integral that allows partial knowledge of this ability to be taken into account, and to measure the impact of this lack of knowledge on the aggregation result. We propose an example in the field of surgical gesture evaluation.

Keywords: Sugeno integral · imprecision · limited knowledge · qualitative aggregation.

1 Introduction

For several decades, many researchers have been interested in replacing precise aggregations with imprecise ones. This type of approach can be found in system identification [1], control [2], economics [3], statistics [4], etc. One of the main objectives of imprecise-valued approach is to take into account either the inconsistency between input data and the underlying model of the aggregation process under study, or the limited knowledge we have of the aggregative model. Several methods can be considered including Monte Carlo-type techniques, ensemble methods [5], bootstrapping, uncertainty propagation [6], interval analysis [7] and so on. In most approaches, convex models are used, i.e. intervallist models. Most of the methods referenced above are relevant to the aggregation of quantitative data.

When it comes to aggregating qualitative data, there are few dedicated methods and to our knowledge only [8] proposed to consider an additive-based intervallist approach. As shown by the work of Grabisch [9], Dubois and Prade [10, 11], the Sugeno integral is a tool dedicated to qualitative aggregation, which can be used to model a wide range of aggregation approaches.

The question we address in this paper is the following: assuming that we have an aggregation model based on a Sugeno integral, can we create an estimate of the final decision in the form of an interval whose width would reflect the dispersion of the aggregated judgements towards that decision? Would such a model also make it possible to represent poor knowledge of how to aggregate these partial judgements?

In a recent paper [12], Loquin et al. proposed an interval-valued extension of Choquet integral for imprecise aggregation, known as maxitive aggregation.

This extension is based on an interpretation of possibility measures as defining a set of probability measures. A maxitive aggregation then defines a convex set of classical weighted aggregations.

We propose to build on recent results [13] to define an analogous extension in the qualitative domain using the Sugeno integral. As in the case of maxitive aggregation, we restrict ourselves to maxitive capacities, i.e. possibility measures which replace probability measures in qualitative setting [14].

2 Background and notations

2.1 Capacities

Let Ω be the finite reference set of possible states of nature: $\Omega = \{1, \ldots, n\}$. In the context of decision making, Ω can be seen as a list of experts providing different assessments based on the same information, or as a set of criteria enabling the same expert to make an assessment.

We consider a totally ordered scale of $t+1$ elements $L = \{0 = \lambda_0 < \ldots < \lambda_t = 1\}$ with a top denoted 1 and a bottom denoted 0. We assume that L is equipped with an order-reversing map $\eta : L \to L$ with $\eta(1) = 0$ and $\eta(0) = 1$. η is unique and such that $\eta(\lambda_i) = \lambda_{t-i}$.

To simplify notations, we will assume that the scale L is such that $\lambda_{t-i} = 1 - \lambda_i$. This case can always be handled by adding elements to the scale. Thus $\forall \lambda \in L$, $\eta(\lambda) = (1 - \lambda)$.

A (qualitative) vector of Ω is a function $\boldsymbol{x} : \Omega \to L$ denoted $\boldsymbol{x} = (x_1, \ldots x_n) \in L^n$.

A (qualitative) L-valued capacity (or L-valued fuzzy measure) is a set function $\mu : 2^\Omega \to L$ such that: $\forall A, B \subseteq \Omega$, $A \subseteq B \implies \mu(A) \leq \mu(B)$, $\mu(\Omega) = 1$ and $\mu(\varnothing) = 0$, $\varnothing$ being the empty set of Ω.

The set of all L-valued capacities of Ω is denoted $\mathcal{K}_L(\Omega)$. Let μ and ν be two capacities of Ω, then $\max(\mu, \nu)$ is a capacity and $\min(\mu, \nu)$ is a capacity.

The **conjugate** capacity of μ is denoted μ^c. It is defined by $\forall A \subseteq \Omega$, $\mu^c(A) = 1 - \mu(A^c)$ where A^c is the complement of A in Ω.

Let μ and ν be two capacities of Ω, we say that μ **dominates** ν if $\forall A \subseteq \Omega$, $\mu(A) \geq \nu(A)$. In the remainder of this article we will denote $\mu \succcurlyeq \nu$ the fact that μ dominates ν and $\mu \preccurlyeq \nu$ the fact that μ is dominated by ν.

In the qualitative context a capacity μ is said to be **optimistic** if dominates its conjugate capacity: $\mu \succcurlyeq \mu^c$ and **pessimistic** if it is dominated by its conjugate capacity: $\mu \preccurlyeq \mu^c$. If μ is optimistic then μ^c is pessimistic and vice versa.

In the context of qualitative aggregation for decision, μ can be interpreted as a degree of relevance of a set of possible state of nature ($A \subseteq \Omega$) to lead to a (qualitative) decision.

As mentioned in [15], in a qualitative context an additive capacity (a probability) makes little sense. On a practical level, we need to ensure that any additive combination of scale elements belongs to the scale. On a theoretical level, a qualitative scale only makes sense because of the order of the terms.

The distance between two terms λ_i and λ_j, on which probability measures rely, is meaningless. In this qualitative context, the maxitive and minitive capacities play a predominant role.

A **maxitive** capacity is a capacity μ satisfying the maxitive axiom:

$$\forall A, B \subseteq \Omega,\ \mu(A \cup B) = \max(\mu(A), \mu(B)).$$

Such a capacity is called **possibility measure**. A possibility measure is completely defined by the finite vector of its values on singletons of Ω [16]. Such a set is called a possibility distribution and denoted $\boldsymbol{\pi} : \Omega \to L$. π_i can be interpreted as the degree of relevance of state i to lead to the decision. A possibility measure defined by the possibility distribution $\boldsymbol{\pi}$ is generally denoted $\Pi_{\boldsymbol{\pi}}$. Due to the maxitivity axiom, $\forall A \subseteq \Omega$, $\Pi_{\boldsymbol{\pi}}(A) = \max_{i \in A} \pi_i$, and $\forall i \in \Omega$, $\Pi_{\boldsymbol{\pi}}(\{i\}) = \pi_i$.

A **minitive** capacity is a capacity μ satisfying the minitive axiom:

$$\forall A, B \subseteq \Omega,\ \mu(A \cap B) = \min(\mu(A), \mu(B)).$$

Such a capacity is called **necessity measure**. It is also completely defined by a possibility distribution $\boldsymbol{\pi}$ and denoted $N_{\boldsymbol{\pi}}$: $\forall A \subseteq \Omega$, $N_{\boldsymbol{\pi}}(A) = \min_{i \in A}(1 - \pi_i)$.

In that case we have, $\forall i \in \Omega$, $N_{\boldsymbol{\pi}}(\{i\}) = 1 - \pi_i$. By definition, $N_{\boldsymbol{\pi}}$ is the conjugate of $\Pi_{\boldsymbol{\pi}}$, i.e. $\forall A \subseteq \Omega$, $N_{\boldsymbol{\pi}}(A) = 1 - \Pi_{\boldsymbol{\pi}}(A^c)$, A^c being the complementary set of A in Ω. A necessity measure can also be defined by its values on singletons: $\iota_i = N_\pi(\{i\}) = 1 - \pi_i$. $\boldsymbol{\iota}$ is then referred as the *impossibility distribution* associated to the necessity measure.

By construction, a possibility measure is an optimistic capacity while a necessity measure is a pessimistic capacity.

When it comes to possibility measures, domination relationships are much simpler: let $\boldsymbol{\pi}$ and $\boldsymbol{\delta}$ be two possibility distributions, $\forall i \in \Omega, \pi_i \geq \delta_i \iff \Pi_{\boldsymbol{\pi}} \succcurlyeq \Pi_{\boldsymbol{\delta}}$ (and also $N_{\boldsymbol{\pi}} \preccurlyeq N_{\boldsymbol{\delta}}$).

2.2 Sugeno integral

The Sugeno integral [17] is a qualitative aggregation method commonly used in multi-criteria decision making.

Let us consider the vector $\boldsymbol{x} = (x_1, \cdots, x_n) \in L^n$, where x_i is the qualitative evaluation of state i (e.g. the evaluation of the i^{th} expert, or the evaluation of a single expert for the i^{th} criterion). Let us consider that to each subset of Ω is associated a measure of its relevance for a global evaluation. In such a context, Sugeno integral is designed to calculate a global valuation based on the partial valuations represented by the vector $\boldsymbol{x}$.

Definition 1 *The Sugeno integral of $\boldsymbol{x} \in L^n$ with respect to a capacity $\mu \in \mathcal{K}_L(\Omega)$ is defined by*

$$S_\mu(\boldsymbol{x}) = \max_{A \subseteq \Omega} \min(\mu(A), \min_{i \in A} x_i). \tag{1}$$

Expression (1) has a computational complexity of 2^n since it has to be valuated on each subset of Ω. Sugeno integral computation can be simplified by considering a permutation σ on Ω such that $x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)}$ and the sets – also called coalitions – $A_{\sigma(i)} = \{\sigma(i), \cdots, \sigma(n)\}$. In that case the Sugeno integral of $\boldsymbol{x} \in L^n$ with respect to μ reduces to:

$$S_\mu(\boldsymbol{x}) = \max_{i=1}^{n} \min(x_{\sigma(i)}, \mu(A_{\sigma(i)})), \tag{2}$$

whose computational complexity is n.

When considering expression (2), Sugeno integral appears as a median of $2n-1$ terms:

$$S_\mu(\boldsymbol{x}) = median(x_1, \cdots, x_n, \mu(A_{\sigma(1)}), \cdots \mu(A_{\sigma(n-1)})).$$

In the particular case where the capacity is a possibility measure defined by a distribution $\boldsymbol{\pi} \in L^n$, this integral reduces to a weighted maximum [18]:

$$S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x}) = \max_{i \in \Omega} \min(\pi_i, x_i).$$

In the particular case where the capacity is a necessity measure defined by a possibility distribution $\boldsymbol{\pi} \in L^n$, this integral reduces to a weighted minimum [18]:

$$S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}) = \min_{i \in \Omega} \max(1 - \pi_i, x_i).$$

Proposition 1 *[19] Let μ, ν be two capacities of Ω, we have the following equivallence: μ dominates ν $\iff$ $\forall \boldsymbol{x} \in L^n$, $S_\mu(\boldsymbol{x}) \geq S_\nu(\boldsymbol{x})$.*

2.3 Core of a (qualitative) capacity

In the quantitative context the core of a capacity is the set of probabilities dominated by the considered capacity. Since probability measures are unsound in the qualitative context, the notion of core should be replaced by those of maxitive and minitive cores [14, 20].

Let μ be a capacity of Ω,

- the maxitive core of μ, denoted $\overline{\mathcal{C}}(\mu)$, is the set of possibility distributions defining a possibility measure that dominates μ:
  $\overline{\mathcal{C}}(\mu) = \{\boldsymbol{\pi} \in L^n / \ \forall A \subseteq \Omega, \Pi_{\boldsymbol{\pi}}(A) \geq \mu(A)\}$.
- the minitive core of μ, denoted $\underline{\mathcal{C}}(\mu)$, is the set of possibility distributions defining a necessity measure dominated by μ:
  $\underline{\mathcal{C}}(\mu) = \left\{\boldsymbol{\pi} \in L^n / \ \forall A \subseteq \Omega, N_{\boldsymbol{\pi}(}A) \leq \mu(A)\right\}$.

It is obvious that the vacuous distribution $\boldsymbol{v} = (1, \ldots, 1)$ belongs to any maxitive or minitive core [15].

Proposition 2 *[21] For any capacity μ of Ω, if $\boldsymbol{\pi} \in L^n$ belongs to $\overline{\mathcal{C}}(\mu)$ then it also belongs to $\underline{\mathcal{C}}(\mu^c)$.*

Moreover, let $\boldsymbol{\pi}$ and $\boldsymbol{\delta}$ be two possibility distributions belonging to $\overline{\mathcal{C}}(\mu)$, then $\boldsymbol{\beta}$ defined by $\forall i \in \Omega$, $\beta_i = \max(\pi_i, \delta_i)$ also belongs to $\overline{\mathcal{C}}(\mu)$.

As proved in [15], any capacity μ can be represented by its maxitive core, or the minitive core of its conjugate since $\forall A \subseteq \Omega$:

$$\mu(A) = \min_{\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu)} \Pi_{\boldsymbol{\pi}}(A) = \max_{\boldsymbol{\pi} \in \underline{\mathcal{C}}(\mu^c)} N_{\boldsymbol{\pi}}(A). \tag{3}$$

This representation induces that Sugeno integral of $\boldsymbol{x} \in L^n$ with respect to capacity μ can be computed by using its maxitive and minitive cores [15]:

$$S_\mu(\boldsymbol{x}) = \min_{\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu)} S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x}) = \max_{\boldsymbol{\pi} \in \underline{\mathcal{C}}(\mu^c)} S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}).$$

2.4 Capacity and Sugeno integral value

This section is a brief reminder of interesting results presented in [22].

Let $\boldsymbol{x} \in L^n$ and $\alpha \in [\min_{i=1}^n x_i, \max_{i=1}^n x_i]$.

– A capacity μ is solution of $S_\mu(\boldsymbol{x}) \geq \alpha$, if and only if it dominates the capacity $\check{\mu}_{\boldsymbol{x},\alpha}$ defined by

$$\check{\mu}_{\boldsymbol{x},\alpha}(A) = \begin{cases} 1 \text{ if } A = \Omega \\ \alpha \text{ if } \{i \mid x_i \geq \alpha\} \subseteq A \\ 0 \text{ otherwise} \end{cases}$$

In other words:

$\{\mu \in \mathcal{K}_L(\Omega) \mid S_\mu(\boldsymbol{x}) \geq \alpha\} = \{\mu \in \mathcal{K}_L(\Omega) \mid \check{\mu}_{\boldsymbol{x},\alpha} \preccurlyeq \mu\}$

By construction, $\check{\mu}_{\boldsymbol{x},\alpha}$ is a necessity measure associated to the distribution $\delta^{\boldsymbol{x},\alpha}$ defined by: $\forall i \in \Omega$, $\delta_i^{\boldsymbol{x},\alpha} = 1 - \alpha$ if $x_i < \alpha$ and 1 otherwise.

– A capacity μ is solution of $S_\mu(\boldsymbol{x}) \leq \alpha$, if and only if it is dominated by the capacity $\hat{\mu}_{\boldsymbol{x},\alpha}$ defined by:

$$\hat{\mu}_{\boldsymbol{x},\alpha}(A) = \begin{cases} 0 \text{ if } A = \varnothing \\ \alpha \text{ if } A \subseteq \{i | x_i > \alpha\} \\ 1 \text{ otherwise.} \end{cases}$$

In other words: $\{\mu \in \mathcal{K}_L(\Omega) \mid S_\mu(\boldsymbol{x}) \leq \alpha\} = \{\mu \in \mathcal{K}_L(\Omega) \mid \mu \preccurlyeq \hat{\mu}_{\boldsymbol{x},\alpha}\}$.

By construction, $\hat{\mu}_{\boldsymbol{x},\alpha}$ is a possibility measure associated to the distribution $\boldsymbol{\pi}^{\boldsymbol{x},\alpha}$ defined by: $\forall i \in \Omega$, $\pi_i^{\boldsymbol{x},\alpha} = \alpha$ if $x_i > \alpha$ and 1 otherwise. Moreover, we always have $\check{\mu}_{\boldsymbol{x},\alpha} \preccurlyeq \hat{\mu}_{\boldsymbol{x},\alpha}$.

Now, if a capacity μ is such that $S_\mu(\boldsymbol{x}) = \alpha$, then $S_\mu(\boldsymbol{x}) \geq \alpha$ and $S_\mu(\boldsymbol{x}) \leq \alpha$, therefore $\check{\mu}_{\boldsymbol{x},\alpha} \preccurlyeq \mu \preccurlyeq \hat{\mu}_{\boldsymbol{x},\alpha}$. In other words: $\{\mu \in \mathcal{K}_L(\Omega) \mid S_\mu(x) = \alpha\}$ is identical to $\{\mu \in \mathcal{K}_L(\Omega) \mid \check{\mu}_{\boldsymbol{x},\alpha} \preccurlyeq \mu \preccurlyeq \hat{\mu}_{\boldsymbol{x},\alpha}\}$.

3 Imprecise aggregation via interval-valued Sugeno integral

3.1 Dominating an unknown capacity

When using Sugeno's integral to aggregate expert opinions for a decision, the choice of capacity to perform the aggregation is critical. As mentioned above, a capacity gives weight to each group of criteria. However, such information is rarely available. At best, we can prioritize or value the partial assessments to be aggregated. The information available for aggregating data is therefore usually available in the form of weight on each criterion, i.e. in the form of a possibility distribution.

It is sometimes possible to obtain from an expert a partial evaluation of the (qualitative) weight of some particular sets. In this case, however, it is important to check that the function obtained is indeed increasing with inclusion, and to bear in mind that this information only gives an imprecise assessment of the capacity to use.

Here, we hypothesize that the true capacity μ to accurately aggregate partial evaluations exists, but is unknown. We assume that the possibility distribution $\boldsymbol{\pi}$ provided by expert interrogation is consistent with this unknown capacity, i.e. that it belongs to the maxitive cores of both μ and μ^c:

$$\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu) \text{ and } \boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu^c).$$

Proposition 3 *$\forall \boldsymbol{\pi} \in L^n, \forall \mu \in \mathcal{K}_L(\Omega)$, if $\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu)$, and $\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu^c)$ then $\forall A \in \Omega$, $N_{\boldsymbol{\pi}}(A) \leq \mu(A) \leq \Pi_{\boldsymbol{\pi}}(A)$.*

Proof. $\forall A \subseteq \Omega$,
$\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu) \implies \mu(A) \leq \Pi_{\boldsymbol{\pi}}(A)$and $\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu^c) \implies \mu^c(A^c) \leq \Pi_{\boldsymbol{\pi}}(A^c)$, thus $1 - \mu(A) \leq \Pi_{\boldsymbol{\pi}}(A^c)$ thus $1 - \Pi_{\boldsymbol{\pi}}(A^c) \leq \mu(A)$, therefore $N_{\boldsymbol{\pi}}(A) \leq \mu(A)$.

Remark that if μ is optimistic then μ dominates μ^c and thus if $\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu)$ then $\boldsymbol{\pi} \in \overline{\mathcal{C}}(\mu^c)$.

Let $\boldsymbol{\pi}$ and $\boldsymbol{\delta}$ be two possibility distributions such that $\Pi_{\boldsymbol{\pi}}$ dominates $\Pi_{\boldsymbol{\delta}}$ (i.e. $\forall i \in \Omega$, $\delta_i \leq \pi_i$). Thus it is easy to check that $N_{\boldsymbol{\pi}} \preccurlyeq N_{\boldsymbol{\delta}} \preccurlyeq \Pi_{\boldsymbol{\delta}} \preccurlyeq \Pi_{\boldsymbol{\pi}}$.

The Sugeno integral being increasing according to a capacity (Proposition 1), then $\forall \boldsymbol{x} \in L^n$:

$$S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}) \leq S_{\mu}(\boldsymbol{x}) \leq S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x}).$$

Moreover, $S_{N_{\boldsymbol{\pi}}} = \min_{\boldsymbol{\delta} \in \overline{\mathcal{C}}(N_{\boldsymbol{\pi}})} S_{\Pi_\delta}$ and $S_{\Pi_{\boldsymbol{\pi}}} = \max_{\boldsymbol{\delta} \in \underline{\mathcal{C}}(N_{\boldsymbol{\pi}})} S_{N_\delta}$.

3.2 Imprecise Sugeno integral

Now let's assume that information about the weight of each criterion in the final decision is known in the form of a possibility distribution $\boldsymbol{\pi}$.

The imprecise Sugeno integral is defined as:

$$\forall \boldsymbol{x} \in L^n,\ IS_{\boldsymbol{\pi}}(\boldsymbol{x}) = [S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}), S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x})]\,. \tag{4}$$

Proposition 4 *$IS_{\boldsymbol{\pi}}(\boldsymbol{x})$ is the bounded set of all Sugeno integral based aggregations of $\boldsymbol{x}$ with regard to a capacity μ of Ω such that both μ and μ^c are dominated by $\Pi_{\boldsymbol{\pi}}$.*

Proof. This property comes directly from the definition of maxitive and minitive cores and the link between capacity domination and Sugeno integral (Prop. 1).

If μ is dominated by $\Pi_{\boldsymbol{\pi}}$ then $\forall \boldsymbol{x} \in L^n$, $S_\mu(\boldsymbol{x}) \leq S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x})$ and thus $S_{\mu^c}(\boldsymbol{x}) \geq S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x})$. If μ^c is dominated by $\Pi_{\boldsymbol{\pi}}$ then $\forall \boldsymbol{x} \in L^n$, $S_{\mu^c}(\boldsymbol{x}) \leq S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x})$ and thus $S_\mu(\boldsymbol{x}) \geq S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x})$

Due to Proposition 1 we have:

Corollary 5 *let $\boldsymbol{\pi}$ and $\boldsymbol{\delta}$ be two possibility distributions of L^n such that $\forall i \in \Omega$, $\delta_i \leq \pi_i$ then $\forall A \subseteq \Omega$, $[S_{N_{\boldsymbol{\delta}}}(A), S_{\Pi_{\boldsymbol{\delta}}}(A)] \subseteq [S_{N_{\boldsymbol{\pi}}}(A), S_{\Pi_{\boldsymbol{\pi}}}(A)]$*
(i.e. $IS_{\boldsymbol{\delta}}(\boldsymbol{x}) \subseteq IS_{\boldsymbol{\pi}}(\boldsymbol{x})$).

It would now be interesting to know whether the interval thus constructed is dense in a way, i.e.

$\forall \boldsymbol{\pi} \in L^n$, $\forall \boldsymbol{x} \in L^n$ and $\forall \alpha \in IS_{\boldsymbol{\pi}}(\boldsymbol{x})$, $\exists \mu \in \mathcal{K}_L(\Omega)$ such that $S_\mu(\boldsymbol{x}) = \alpha$.

We construct this as a theorem in several successive steps.

First, remark that $\forall \boldsymbol{x} \in L^n$ and $\forall \mu \in \mathcal{K}(\Omega)$ we have:

$$\min_{i=1}^n x_i \leq S_\mu(\boldsymbol{x}) \leq \max_{i=1}^n x_i.$$

Proposition 6 *Let $\boldsymbol{x} \in L^n$ be a vector,*

$$\forall \alpha \in \left[\min_{i=1}^n x_i, \max_{i=1}^n x_i\right], \exists \boldsymbol{\pi}^{\boldsymbol{x},\alpha}, \text{ such that } S_{\Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}}(x) = \alpha.$$

and

$$\forall \alpha \in \left[\min_{i=1}^n x_i, \max_{i=1}^n x_i\right], \exists \boldsymbol{\delta}^{\boldsymbol{x},\alpha}, \text{ such that } S_{N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}}(x) = \alpha.$$

Proof. The proof of this proposition is trivial when referring to Section 2.4.

For the rest of the construction, it is important to remember that $S_{N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}}(\boldsymbol{x}) = \alpha = S_{\Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}}(\boldsymbol{x})$ and $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}} \preccurlyeq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}$.

Proposition 7 *Let $\boldsymbol{x} \in L^n$ be a vector, $\boldsymbol{\pi} \in L^n$ be a possibility distribution, $\alpha \in IS_{\boldsymbol{\pi}}(\boldsymbol{x})$, and $\boldsymbol{\pi}^{\boldsymbol{x},\alpha}$, $\boldsymbol{\delta}^{\boldsymbol{x},\alpha}$ being the possibility distributions defined above, then $\forall A \subseteq \Omega$, $[N_{\boldsymbol{\pi}}(A), \Pi_{\boldsymbol{\pi}}(A)] \cap [N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A), \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}(A)] \neq \varnothing$.*

Proof. To prove this proposition, we show that there is always at least a capacity $\tilde{\mu} \in \mathcal{K}_L(\Omega)$ such that $N_{\boldsymbol{\pi}} \preccurlyeq \tilde{\mu} \preccurlyeq \Pi_{\boldsymbol{\pi}}$ and $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}} \preccurlyeq \tilde{\mu} \preccurlyeq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}$.

Since $S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}) \leq \alpha \leq S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x})$, according to [22], $N_{\boldsymbol{\pi}} \preccurlyeq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}$ and $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}} \preccurlyeq \Pi_{\boldsymbol{\pi}}$. There are thus four possible cases for each subset $A \subseteq \Omega$:

1. $N_{\boldsymbol{\pi}}(A) \leq N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A) \leq \Pi_{\boldsymbol{\pi}}(A) \leq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}(A)$,
2. $N_{\boldsymbol{\pi}}(A) \leq N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A) \leq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}(A) \leq \Pi_{\boldsymbol{\pi}}(A)$.
3. $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A) \leq N_{\boldsymbol{\pi}}(A) \leq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}(A) \leq \Pi_{\boldsymbol{\pi}}(A)$,
4. $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A) \leq N_{\boldsymbol{\pi}}(A) \leq \Pi_{\boldsymbol{\pi}}(A) \leq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}(A)$.

Let us define the following capacities:
$\mu = \max(N_{\boldsymbol{\pi}}, N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}})$ and $\mu' = \min(\Pi_{\boldsymbol{\pi}}, \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}})$. First remark that: $\forall A \subseteq \Omega$,

$$\mu^c(A) = 1 - \mu(A^c) = 1 - \max\left(N_{\boldsymbol{\pi}}(A^c), N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A^c)\right) = \min\left(\Pi_{\boldsymbol{\pi}}(A), \Pi_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}(A)\right).$$

Thus, $\mu^c = \mu'$ and, by construction, $\mu \preccurlyeq \mu'$. Thus μ satisfies $N_{\boldsymbol{\pi}} \preccurlyeq \mu \preccurlyeq \mu^c \preccurlyeq \Pi_{\boldsymbol{\pi}}$ and $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}} \preccurlyeq \mu \preccurlyeq \mu^c \preccurlyeq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}$.

Theorem 8 *Let $\boldsymbol{x} \in L^n$ be a vector, $\boldsymbol{\pi} \in L^n$ be a possibility distribution, $\forall \alpha \in IS_{\boldsymbol{\pi}}(\boldsymbol{x})$, $\exists \mu \in \mathcal{K}_L(\Omega)$ such that $N_{\boldsymbol{\pi}} \preccurlyeq \mu \preccurlyeq \Pi_{\boldsymbol{\pi}}$ and $S_\mu(\boldsymbol{x}) = S_{\mu^c}(\boldsymbol{x}) = \alpha$.*

Proof. Theorem 8 can be regarded as a corollary of Proposition 7. Let $\boldsymbol{\pi}^{\boldsymbol{x},\alpha}$, $\boldsymbol{\delta}^{\boldsymbol{x},\alpha}$ be the two possibility distributions defined ins Section 2.4 and let μ be the capacity defined by $\mu = \max(N_{\boldsymbol{\pi}}, N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}})$.
Since $S_{N_{\boldsymbol{\pi}}}(\boldsymbol{x}) \leq \alpha \leq S_{\Pi_{\boldsymbol{\pi}}}(\boldsymbol{x})$, $N_{\boldsymbol{\pi}} \preccurlyeq \Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}$ and $N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}} \preccurlyeq \Pi_{\boldsymbol{\pi}}$.
$S_{N_{\boldsymbol{\delta}^{\boldsymbol{x},\alpha}}}(\boldsymbol{x}) = \alpha \leq S_\mu(\boldsymbol{x}) \leq S_{\mu^c}(\boldsymbol{x}) \leq S_{\Pi_{\boldsymbol{\pi}^{\boldsymbol{x},\alpha}}}(\boldsymbol{x}) = \alpha$. Thus $S_\mu(\boldsymbol{x}) = S_{\mu^c}(\boldsymbol{x}) = \alpha$.

In addition, information on how to aggregate partial evaluations may be imprecisely available, i.e. a family of p possibility distributions $\{\boldsymbol{\pi}^i\}_{i=1\ldots p}$ is assumed to be known. This is the case, for example, when several experts are asked to each give a weighting to each criterion, and these weights are different.

By supposing that this information is coherent in a way, i.e. $\exists \mu \in \mathcal{K}_L(\Omega)$ such that $\forall i \in \{1 \ldots p\}$, $\boldsymbol{\pi}^i \in \overline{\mathcal{C}}(\mu)$ and $\boldsymbol{\pi}^i \in \overline{\mathcal{C}}(\mu^c)$. In that case, due to Expressions (3) $\forall \boldsymbol{x} \in L^n$ we have:

$$S_\mu(\boldsymbol{x}) \in \bigcap_{i=1}^{p} IS_{\boldsymbol{\pi}}(\boldsymbol{x}) = \left[\max_{i=1}^{p} S_{N_{\boldsymbol{\pi}^i}}(\boldsymbol{x}), \min_{i=1}^{p} S_{\Pi_{\boldsymbol{\pi}^i}}(\boldsymbol{x})\right].$$

This technique can reveal the fact that the weight distributions given by each expert are conflicting, and in this case $\exists \boldsymbol{x} \in L^n$, such that $\bigcap_{i=1}^{p} IS_{\boldsymbol{\pi}}(\boldsymbol{x}) = \varnothing$.

Example: Let $\boldsymbol{x} = \{0.1, 0.5, 0.8\} \in L^3$ be the partial evaluations on $n = 3$ criteria. Let us suppose that interviewing two experts leads to two possibility distributions $\boldsymbol{\pi}^1 = \{0.0, 0.4, 0.8\}$ and $\boldsymbol{\pi}^2 = \{0.8, 0.4, 0.0\}$. Thus $IS_{\boldsymbol{\pi}^1}(\boldsymbol{x}) = [0.7, 0.8]$ while $IS_{\boldsymbol{\pi}^2}(\boldsymbol{x}) = [0.2, 0.4]$. Obviously $IS_{\boldsymbol{\pi}^1}(\boldsymbol{x}) \cap IS_{\boldsymbol{\pi}^2}(\boldsymbol{x}) = \varnothing$.

4 Experiment

In this section, we propose an application of the proposed interval-valued Sugeno integral based on the technical assessment of ENT (ear, nose, and throat) surgeons.

Training to surgery can be difficult for the more critical procedures during residency. In this context, researchers from the Montpellier University Hospital in France investigated the dead porcine model for training to total laryngectomy (TL). Eighteen surgeons, decomposed into three level groups – young residents (post graduate year (YR – 1 to 3)), experienced residents (ER – 4 to 6) and senior surgeons (SS) – were asked to perform the full surgical operation in dead swines.

Seven main steps ($O_{1...7}$) of the surgery were video-recorded and rated from 5 (poor) to 25 (excellent) by 3 surgeon's experience blinded experts of the procedure using modified Objective Structured Assessment of Technical Skills (OSATS). OSATS is a validated assessment tool that evaluates the technical competency in a particular technique, in order to grade the overall technical proficiency for open surgery. It consists of a procedure specific checklist, a pass/fail judgment and a global rating scale. For each task and each surgeon, the mean OSATS score of the three experts was considered. The results of this evaluation are shown in Table 1.

Surgeon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Group	ER	YR	SS	SS	SS	ER	YR	SS	ER	ER	YR	YR	YR	YR	ER	YR	ER	SS	Diff.
O_1	12	11	23	14	20	15	13	22	17	19	16	14	16	14	17	13	20	25	1
O_2	12	13	23	20	22	17	8	25	16	19	17	13	13	14	12	10	18	22	5
O_3	11	8	24	21	21	12	11	23	15	16	12	14	12	13	11	5	15	25	4
O_4	16	9	22	21	21	16	11	23	17	17	13	11	13	14	13	9	15	24	6
O_5	13	8	25	19	17	15	9	24	15	15	14	9	12	10	11	12	18	25	4
O_6	8	8	24	17	20	10	5	24	13	9	15	13	13	8	12	11	18	25	7
O_7	16	14	20	18	18	18	12	24	16	14	15	15	15	13	9	7	20	25	8

Table 1: Mean OSATS for the 7 steps of a laryngectomy with the level of technicality of each step.

A weight is attached to each stage, defining its level of technicality (last column of Table 1). We use this weight to define the possibility distribution associated with this OSATS. As the two scales (OSATS and weights) are not commensurable, we reduce them to 0 and 1 in two different ways, because they don't represent the same thing.

OSATS are scores and thus can be reduced e.g. by dividing each value by 25.

In contrast, the weights associated to the stages were given by the evaluators as a level of technicality. They see it more as an additive scale: for a surgery to be successful, it is not sufficient for a single step to be performed at its best, but rather to have a maximum score on the technicality scale. We therefore propose to transform this 1 to 8 scale into a possibility distribution using a probability / possibility transformation.

There are two main transformations. One is more relevant in a subjective context while the other gives the most specific possibility distribution dominating the probability distribution of interest [16]. We propose to investigate both approaches.

Let w_i be the weights associated to the i^{th} step of the surgery. Let $\boldsymbol{\rho} \in \mathbb{R}^{+n}$ be the values obtain by normalizing the weights ($\rho_i = \frac{w_i}{35}$, since $\sum_{i=1}^{7} w_i = 35$).

The so-called **optimal** transformation of the probability distribution $\boldsymbol{\rho}$ into the possibility distribution $\boldsymbol{\pi}^{\mapsto\rho}$ is given by:

$$\forall i \in \{1, \dots, 7\},\ \pi_i^{\mapsto\rho} = \sum_{j=1}^{7} \rho_j.\chi(\rho_j \leq \rho_i),$$

where, $\forall a, b, \in [0,1]$ $\chi(a \leq b) = 1$ if $a \leq b$ and 0 else.

The **subjective** transformation of the probability distribution $\boldsymbol{\rho}$ into the possibility distribution $\boldsymbol{\pi}^{(\rho)}$ is given by:

$$\forall i \in \{1, \ldots, 7\},\ \pi_i^{(\rho)} = \sum_{j=1}^{7} \min(\rho_i, \rho_j).$$

By construction we have $\forall i \in \Omega$, $\pi_i^{\mapsto\rho} \leq \pi_i^{(\rho)}$, thus $\Pi_{\boldsymbol{\pi}^{(\rho)}}$ dominates $\Pi_{\boldsymbol{\pi}^{\mapsto\rho}}$.

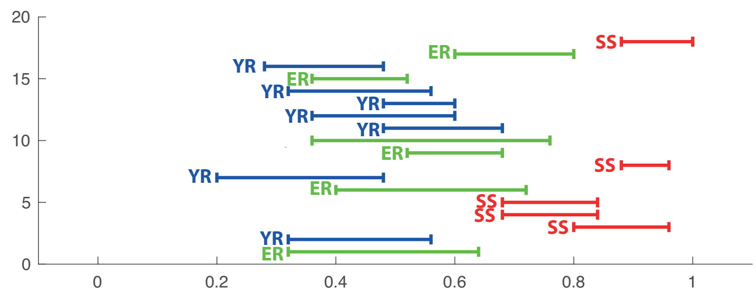

Fig. 1: Imprecise aggregation by using the subjective transformation.

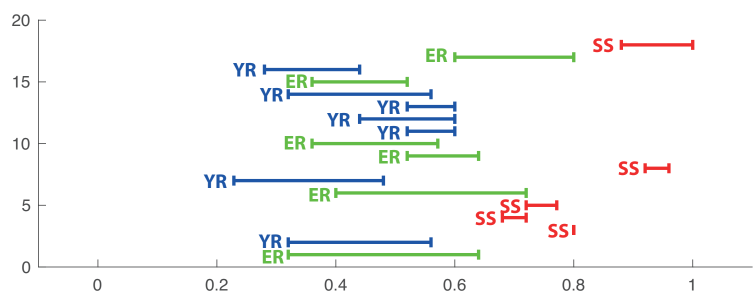

Fig. 2: Imprecise aggregation by using the optimal transformation.

Figures (1) and (2) plot the intervals obtained by aggregating the scores obtained by each surgeon for each step of the surgery. The intervals associated to the young residents are plotted in blue, experience residents in green and senior surgeons in red.

Naturally, due to Corollary 5, since $\Pi_{\boldsymbol{\pi}^{(\rho)}}$ dominates $\Pi_{\boldsymbol{\pi}^{\mapsto\rho}}$, the interval obtained with the optimal transformation are tighter than those obtained by the subjective transformation. For example, for the 3^{rd} surgeon (senior surgeon) the interval-valued aggregation is reduced to a single value (0.8) when using the optimal transformation while it stays as an interval ($[0.8, 0.96]$) when using the subjective transformation. On the other hand, for the second surgeon, the interval-valued aggregation is unchanged ($[0.32, 0.56]$) whatever the possibility distribution.

On a more pragmatic level, what emerges in these two graphs is that there is a very strong difference between the evaluations of senior surgeons and those of residents, while this distinction is not clear between young and experienced residents. This distinction is much less clear when we simply look at the min-max intervals of the partial scores, as can be seen in Figure (3).

The fact that steps $4, 6$ and 7 are more significant according to the evaluator tends to highlight the difference between senior surgeons and residents. Moreover, we can see that, according to this aggregation, these evaluations show that senior surgeons results completely dominate young residents and most of the experienced residents.

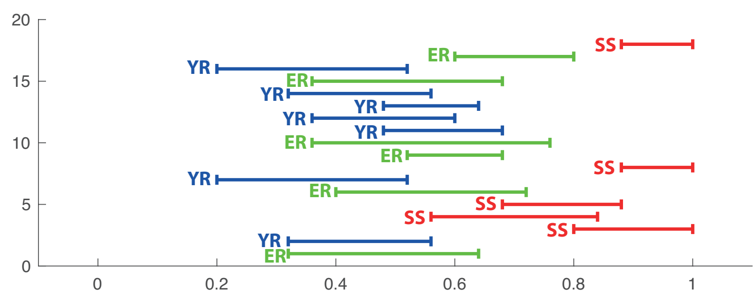

Fig. 3: Range of partial score values.

For comparison, Figure (4) shows the projection of the (unweighted) principal component analysis data onto the first two principal axes. As in previous figures, blue is for young residents, green for experience residents and red for senior surgeons. We can see that the senior surgeons stand out quite well from the residents, while the residents form a group that is difficult to distinguish. The weighted interpretation given by imprecise aggregation gives richer information in the sense that we find this result but for a whole set of possible aggregations, which is a stronger conclusion.

5 Conclusion and discussion

In this article, we proposed the construction of an intervallist extension of the Sugeno integral. The aim of this construction is to be able to handle imprecise

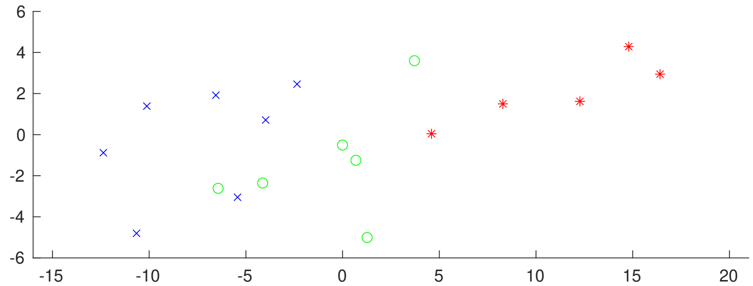

Fig. 4: Projections of the individual evaluations of the surgeons on the two first principal axes (×: YR, ○: ER, ⋆: SS)

knowledge of the appropriate capacity for a qualitative aggregation. We assume that imprecise knowledge is available in the form of one or several coherent possibility distributions. The main result of this article is Theorem 8 showing that the interval produced by this extension is dense in L. A medical example shows the benefits of using such an approximation.

Among the many future works we're considering, we'd like to find explicit conditions for knowing whether a set of possibility distributions is consistent, i.e. there exists a capacity μ that can be imprecisely reconstructed by using this set of possibility. Similarly, on a practical level, it would be interesting to explore a little more how to exploit expert information on sets that are not singletons.

Finally, as you may have noticed, the scores associated with each surgical step were in fact averages of assessments given by three different experts. It might be interesting to carry on the intervallist approach by calculating an aggregated interval for each expert. The final aggregation could consist in aggregating the intervals obtained, or in developing a method taking into account multiple evaluations for several criteria.

References

1. L. Bako, S. Ndiaye, and E. Blanco, "An interval-valued recursive estimation framework for linearly parameterized systems," *Systems & Control Letters*, vol. 168, p. 105345, 2022.
2. L. Jaulin, M. Kieffer, O. Didrit, and W. E, *Applied Interval Analysis*, ., Ed. Springer Verlag, 2001.
3. D. Butler and G. Loomes, "Imprecision as an account of violations of independence and betweenness," *Journal of Economic Behavior & Organization*, vol. 80, no. 3, pp. 511–522, 2011.
4. J. Perolat, I. Couso, K. Loquin, and O. Strauss, "Generalizing the wilcoxon rank-sum test for interval data," *International Journal of Approximate Reasoning*, vol. 56, pp. 108–121, 2015.

5. P. C. Pendharkar, "Ensemble based point and confidence interval forecasting in software engineering," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9441–9448, 2015.
6. J. Liu, Y. Shi, C. Ding, and M. Beer, "Hybrid uncertainty propagation based on multi-fidelity surrogate model," *Computers & Structures*, vol. 293, p. 107267, 2024.
7. R. Calm, M. García-Jaramillo, J. Bondia, M. Sainz, and J. Vehí, "Comparison of interval and monte carlo simulation for the prediction of postprandial glucose under uncertainty in type 1 diabetes mellitus," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 325–332, 2011.
8. A. Mármol, J. Puerto, and F. Fernández, "The use of partial information on weights in multicriteria decision problems," *Journal of Multi-criteria devision analysis*, vol. 7, pp. 322–329, 1998.
9. M. Grabisch and C. Labreuche, "A decade of application of the choquet and sugeno integrals in multi-criteria decision aid," *Annals of Operations Research*, vol. 175, pp. 247–286, 2010.
10. D. Dubois, H. Prade, and R. Sabbadin, "Qualitative decision theory with Sugeno integrals," in *14th Conference on Uncertainty in Artificial Intelligence (UAI 1998)*, G. F. Cooper and S. Moral, Eds. Madison, Wisconsin, United States: Morgan Kaufmann, 1998, pp. 121–128.
11. D. Dubois, J.-L. Marichal, H. Prade, M. Roubens, and R. Sabbadin, "The use of the discrete Sugeno integral in decision-making: a survey," in *A-MCD-A - Aide Multi Critère à la Décision (Multiple Criteria Decision Aiding)*, A. Colorni, M. Paruccini, and B. Roy, Eds. European Commission Joint Research Centre, 2001, pp. 81–98.
12. K. Loquin and O. Strauss, "On the granularity of summative kernels," *Fuzzy Sets and Systems*, vol. 159, no. 15, pp. 1952–1972, 2008, from Knowledge Representation to Information Processing and Management.
13. D. Dubois, H. Prade, and A. Rico, "Qualitative integrals and desintegrals as lower and upper possibilistic expectations," in *Proceedings of the 8th EUSFLAT conference*, J. Montero, G. Pasi, and D. Ciucci, Eds. Atlantis Press, 2013.
14. D. Dubois, F. Faux, H. Prade, and A. Rico, "Qualitative capacities: Basic notions and potential applications," *Int. J. Approx. Reason.*, vol. 148, pp. 253–290, 2022.
15. D. Dubois, H. Prade, and A. Rico, "Representing qualitative capacities as families of possibility measures," *International Journal of Approximate Reasoning*, vol. 58, pp. 3–24, 2015.
16. D. Dubois, "Possibility theory and statistical reasoning," *Computational Statistics & Data Analysis*, vol. 51, no. 1, pp. 47–69, 2006, the Fuzzy Approach to Statistical Analysis.
17. M. Sugeno, "Theory of fuzzy integrals and its applications," Ph.D. dissertation, Tokyo Institute of Technology, 1974.
18. D. Dubois and H. Prade, "Weighted minimum and maximum operations in fuzzy set theory," *Inf. Sci.*, vol. 39, no. 2, pp. 205–210, 1986.
19. M. Sugeno, "Fuzzy measures and fuzzy integrals—a survey," in *Readings in Fuzzy Sets for Intelligent Systems*, D. Dubois, H. Prade, and R. R. Yager, Eds. Morgan Kaufmann, 1993, pp. 251–257.
20. M. Grabisch, *Set functions, Games and Capacities in Decision-Making*, Springer-Verlag, Ed., 2016.
21. D. Dubois, F. Faux, H. Prade, and A. Rico, "Qualitative capacities: Basic notions and potential applications," *Int. J. Approx. Reason.*, vol. 148, pp. 253–290, 2022.
22. A. Rico, M. Grabisch, C. Labreuche, and A. Chateauneuf, "Preference modelling on totally ordered sets by the Sugeno integral," *Discrete Applied Mathematics*, vol. 147, no. 1, pp. 113–124, 2005.