A Weighted Matrix Visualization for Fuzzy Measures and Integrals

Andrew R. Buck, Derek T. Anderson, James M. Keller, Timothy Wilkin, and Muhammad Aminul Islam

***Abstract*—Fuzzy integrals are useful general purpose aggregation operators, but they can be difficult to understand and visualize in practice. The interaction between an exponentially increasing number of variables–2^n fuzzy measure variables for n inputs–makes it hard to understand what exactly is going on in a high dimensional space. We propose a new visualization scheme based on a weighted indicator matrix to better understand the inner workings of an arbitrary fuzzy measure. We provide ways of viewing the Shapley and interaction indices, as well as an optional data coverage histogram. This approach can give insight into which sources are the most relevant in the overall aggregation and decision making process, and it provides a way to visually compare fuzzy measures and subsequently integrals.**

***Index Terms*—fuzzy measure, fuzzy integral, visualization**

I. Introduction

Fuzzy measures and fuzzy integrals are powerful tools for performing generalized non-linear aggregation. Their expressiveness, however, has led to some mystification regarding the actual operation of their inner workings in practice. At the heart of the matter is the issue of understanding the interaction and relationships between the 2^n variables that arise from n inputs. On a purely computational level, the fuzzy integral and measure is well understood, but it remains difficult to intuitively grasp what occurs in high dimensional spaces.

There have been some approaches for visualizing fuzzy measures[1] [1] and high dimensional set interactions [2, 3], but these have fallen short of offering a holistic understanding of the complete aggregation process, importance of individual sources and relationships between sources. For instance, in [1], a fuzzy measure is drawn as a lattice of all subset elements, arranged by cardinality in a Hasse diagram. Nodes are scaled to be proportional in size to the value of each subset, and a path is drawn for the walk taken by the permutation of each training data sample used in evaluating the fuzzy integral. This approach is useful for understanding the data coverage problem [1, 4, 5] at a high-level, but it does little to indicate which subsets or data sources are over or underutilized.

We seek to address this problem by presenting a weighted matrix visualization of an arbitrary fuzzy measure. Our approach utilizes basic data visualization guidelines established in [6] and [7], such as minimizing "chartjunk" and maximizing the "data to ink" ratio. The main idea of our method involves constructing a weighted indicator matrix of all possible subsets and scaling the widths and heights of the rows and columns to be proportional to useful values. We then include the incremental contribution of each source, the Shapley and interaction indices, and an optional data visitation histogram. The resulting graphic serves as both an exploratory and an explanatory visualization of a fuzzy measure, providing an overview of the general nature of the measure while also allowing one to inspect the numerous interactions at play. The diagram is compact enough to be used as part of a small multiple, comparing several different measures at once, and arguably serves as a form of "modern" art [8].

The remainder of this paper is structured as follows. Section II provides the background notation for fuzzy measures and fuzzy integrals. Section III describes our approach in detail using an illustrative example. Section IV shows several examples of the proposed visualization technique on various types of fuzzy measures, and Section V gives our conclusions.

II. Background

A. Fuzzy Measures

A fuzzy measure g defined on a finite set $X = \{x_1, ..., x_n\}$ is a function $g : \mathcal{P}(X) \to \Re^+$ satisfying the following[2]:

(i) $g(\emptyset) = 0$

(ii) $A \subseteq B \subseteq X$ implies $g(A) \leq g(B)$.

Here we note that $\mathcal{P}(X)$ is the power set of X, which includes all possible combinations of the elements of X. For instance, if $X = \{x_1, x_2, x_3\}$, then $\mathcal{P}(X) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_2\}, \{x_1, x_3\}, \{x_2, x_3\}, \{x_1, x_2, x_3\}\}$. The usefulness of a fuzzy measure comes from its ability to model the worth of any subset of X. In general, $g(A)$ represents the value or utility of the subset $A \subseteq X$.

There are several properties of a fuzzy measure that can be computed to give insight into the inner workings of the measure. One is the Shapley value [9], which can be used to assess the relative importance of each individual source element in X. The Shapley value of a fuzzy measure g is defined as the vector $[s_1, ..., s_n]$, where

$$s_i = \sum_{K \subseteq X \setminus i} \frac{(n - |K| - 1)!|K|!}{n!} \left[g(K \cup i) - g(K) \right], \quad (1)$$

and n is the number of elements in X (i.e. $n = |X|$). The Shapley index s_i of an element $x_i \in X$ represents the average contribution that x_i makes when added to an existing subset.

Andrew, Derek, James, and Muhammad are with the Electrical Engineering and Computer Science (EECS) department, University of Missouri (MU), MO, USA, e-mail: buckar@missouri.edu. Tim Wilkin is with the School of Information Technology, Deakin University, Victoria, AU.

[1] Hereafter, for the sake of brevity, we only make reference to the fuzzy measure versus the measure and integral, unless there is a specific reason to discuss the measure relative to the integral.

[2] Typically g is defined such that $g(X) = 1$, e.g., in decision level fusion, however this is not strictly necessary.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

The Shapley value gives a normalized weight to each element such that $\sum_{i=1}^{n} s_i = g(X)$.

While the Shapley value is defined only for individual elements of X, the process has been generalized [10] to extend to arbitrary groups of elements. The interaction index of a subset $A \subseteq X$ for a fuzzy measure g is defined as

$$I(A) = \sum_{B \subseteq X \setminus A} \frac{(n - |B| - |A|)!|B|!}{(n - |A| + 1)!} \sum_{C \subseteq A} (-1)^{|A \setminus C|} g(C \cup B). \quad (2)$$

Like the Shapley value, the interaction index of a set A gives a sense of the worth of the set in the context of the fuzzy measure. When $I(A)$ is positive, the set is said to have positive synergy, indicating that the elements are complementary and that there is value in their combined usage. In contrast, when $I(A)$ is negative, the set is said to have negative synergy, indicating that the elements are redundant and the set as a whole brings no added value [11].

B. Fuzzy Integrals

The fuzzy integral is a validated tool with wide reaching applications from information fusion to multicriteria decision-making [12]. The fuzzy integral is defined with respect to a fuzzy measure. Let $h : X \rightarrow [0, 1]$ be a function that specifies the value of a single element $x \in X = \{x_1, ..., x_n\}$. Given h and a fuzzy measure g defined on X, the discrete Choquet integral is defined as

$$C_g(h) = \int_C h \circ g = \sum_{i=1}^{n} h(x_{\pi(i)}) \left[g(A_i) - g(A_{i-1}) \right], \quad (3)$$

where π is a permutation of X such that $h(x_{\pi(1)}) \geq h(x_{\pi(2)}) \geq ... \geq h(x_{\pi(n)})$ and $A_i = \{x_{\pi(1)}, ..., x_{\pi(i)}\}$ with $g(A_0) = 0$ [13].

We can consider an individual data sample as an instance of h that produces an output $C_g(h)$. In the evaluation of the Choquet integral, the elements of X are ordered according to h, resulting in a sequence of n subsets $A_1, ..., A_n$ that are used by g. Note that while the fuzzy measure is defined over 2^n possible subsets, only n of these are visited for a single data sample. We call the sequence of subsets visited by a data sample h a walk notated as W_h. The distribution of visited subsets over all possible subsets of X can have serious implications in the quality of any data-driven learning method for the fuzzy measure.

III. Method

A. General Approach

We now present the method for constructing the weighted matrix visualization of a fuzzy measure. The process is most clearly explained by working through an example step by step. Consider the fuzzy measure g defined in Table I. This is a measure defined on a set of three elements, $X = \{x_1, x_2, x_3\}$. We start by ordering the elements of the power set $\mathcal{P}(X)$ first by cardinality and then lexicographically within each subset of equal size as in Table I. A binary indicator matrix M is defined such that each row i of the matrix corresponds to an

TABLE I: Fuzzy Measure Example

A	$g(A)$
$\emptyset$	0
$\{x_1\}$	0.3
$\{x_2\}$	0.2
$\{x_3\}$	0.4
$\{x_1, x_2\}$	0.7
$\{x_1, x_3\}$	0.8
$\{x_2, x_3\}$	0.4
$\{x_1, x_2, x_3\}$	1

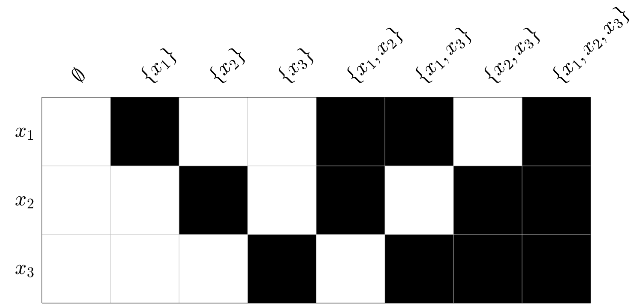


Fig. 1: Binary indicator matrix for the power set of $\{x_1, x_2, x_3\}$.

element $x_i \in X$ and each column j corresponds to a subset $A \subseteq X$, where

$$M_{ij} = \begin{cases} 1 & x_i \in A_j \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The matrix for a fuzzy measure defined on a set with three elements is shown in Fig. 1.

On its own, the binary indicator matrix says nothing about the fuzzy measure itself. There are several ways to encode the values of the fuzzy measure as part of the visualization. We choose to adjust the width of each matrix column corresponding to a subset $A \subseteq X$ to be proportional to $g(A)$, as shown in Fig 2. Here, the overall width of the matrix remains the same (taken to be $disp_w$), and the width of each column j is calculated as $disp_w \cdot g(A_j) / \sum_{B \subseteq X} g(B)$. Note that since $g(\emptyset) = 0$, the empty set is not shown. Measures that include several small values of $g(A)$ may end up with columns that are too narrow to interpret, but this is an indication that these subsets do not contribute greatly in the overall evaluation of the fuzzy integral. In small examples, we can label each column j with $g(A_j)$ on the top of the diagram. However, as the number of columns increases with larger measures, the labels are removed to improve legibility.

To clarify the visual presentation, particularly as we begin to work with larger and more complex fuzzy measures, we separate the diagram into separate parts for each cardinality set. Within each set, the columns are sorted in order of increasing values of $g(A)$ as shown in Fig. 3. This has the effect of moving the strongest subsets to the right and will make it easier to identify trends and substructure within the fuzzy measure.

We now introduce what is perhaps the most important feature of the weighted matrix visualization. The incremental

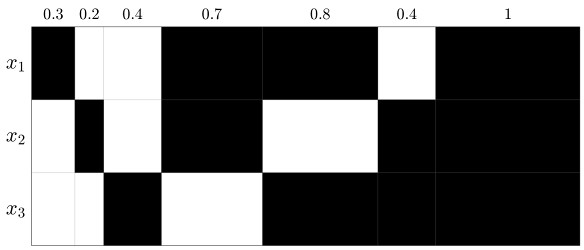


Fig. 2: Column widths are scaled to be proportional to the fuzzy measure value.

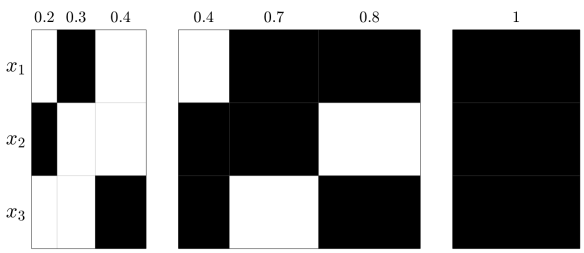


Fig. 3: Sets are separated by cardinality and columns are reordered to have increasing values.

contribution of each source i in every subset $A_j \subseteq X$ is defined as

$$\Delta g_{ij} = g(A_j) - g(A_j \backslash i). \quad (5)$$

The incremental contribution values form the basis of the Shapley value calculated in Eq. 1. In Fig. 4, the incremental contributions are shown as the black shaded parts of each grid cell in the matrix visualization. The binary indicator matrix of Fig. 3 is now shown with a lighter gray color to allow the darkened parts to stand out. The Shapley index of a source (row) is proportional to the sum of the widths of the black regions in that row.

Clearly, for subsets with a single element ($A_j = \{x_i\}$), $\Delta g_{ij} = g(x_i) = g(A_j)$, so the three left-most columns are drawn with completely black elements. In the center part of the diagram corresponding to subsets with a pair of elements $\{x_u, x_v\}$, the black bars indicate the amount that each source x_u contributes to the fuzzy measure value of the pair of sources. For instance, $g(\{x_2, x_3\}) = 0.4$, $g(\{x_2\}) = 0.2$, and $g(\{x_3\}) = 0.4$. The incremental contribution of x_2 in the set $\{x_2, x_3\}$ is $g(\{x_2, x_3\}) - g(\{x_3\}) = 0$, so there is no darkened black region in the grid cell for x_2 in the column for the set $\{x_2, x_3\}$. Likewise, the incremental contribution of x_3 in the same set is $g(\{x_2, x_3\}) - g(\{x_2\}) = 0.2$, so the grid cell for x_3 is half shaded. We can see in this example that x_1 has the greatest incremental contribution to the full set X, which shows that the subset without x_1 was the lowest valued subset of size two, and so the addition of x_1 has the largest impact on the fuzzy measure value. In other words, the two two-element subsets that do not include x_2 and x_3 respectively already have relatively high values assigned by the fuzzy measure, and so do not see as much benefit from including those elements.

An alternative interpretation of the shaded area is to consider the gray portion of each grid cell. The width of this region is

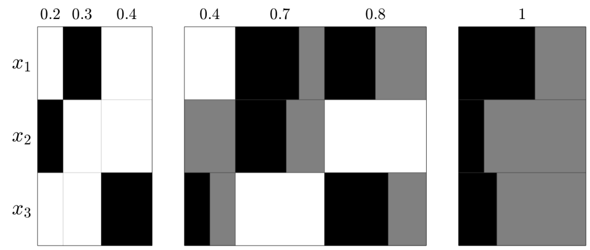


Fig. 4: Each grid cell is shaded with the incremental contribution of that element.

equivalent to the value of the fuzzy measure with the specified element removed. For instance, in the right-most column of the example in Fig. 4 corresponding to the full set X, the grid cell for source x_2 is 80% gray with only 20% shaded black. This shows that the value of the set $\{x_1, x_3\}$ is already 80% of the value of the full set, so adding x_2 has only a relatively small impact.

B. Shapley and Interaction Indices

To further enhance the information presented in the visualization, we can include both the Shapley value of the fuzzy measure and the interaction indices. The Shapley value is obtained with Eq. 1 and to more clearly illustrate it, the row heights of the diagram are scaled to be proportional to the Shapley index of each source. Fig. 5 shows that the Shapley value for the example is $[0.45, 0.2, 0.35]$, suggesting that source x_1 is the most important and source x_2 is the least important. Modifying the row heights makes the Shapley value proportional to the widths of the black regions in each row as opposed to the total black area, but altering the row heights makes any differences in the Shapley value more obvious.

The Shapley indices alone do not tell the whole story, as they apply only to single source elements. The interaction indices are an extension of the Shapley indices for arbitrary sets. For each subset $A \subseteq X$ we can compute the interaction index $I(A)$ and plot it as a bar graph below the matrix diagram. This "row" of the diagram is labeled with an "I" and the scale[3] is fixed to the range [-1, 1]. Positive values are colored red and negative values are colored blue to emphasize positive and negative interaction. The interaction indices of the three left-most columns are equivalent to the Shapley values, although note that they may be presented in a different order due to the ordering of the $g(A)$ values. The example shows that $I(\{x_1, x_2\})$ and $I(\{x_1, x_3\})$ are both positive, whereas $I(\{x_2, x_3\})$ is negative. This suggests that the subset $\{x_2, x_3\}$ is redundant and adds no value, which can be gathered from the fact that $g(\{x_2, x_3\}) = g(\{x_3\})$. The interaction of the full set is zero in this example, which indicates that this is a 2-additive measure [14].

Measures with small values of $g(A)$ will have reduced column widths that can make it difficult to view the graph of interaction indices, or may prevent it from being drawn at

[3] Although the interaction index of a subset A with $|A| > 2$ can exceed ± 1, keeping a fixed range provides consistency across multiple diagrams.

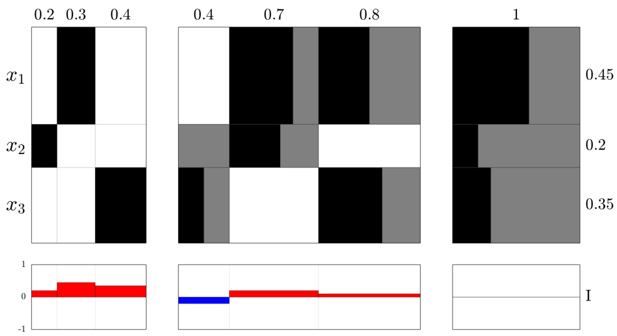


Fig. 5: A bar graph of the interaction indices is included for each subset.

all. In these cases, it is useful to rely on the row heights to show the Shapley indices, although the interaction indices of subsets with two or more elements may still be difficult to see. Often, the interaction indices for subsets with very low values of $g(A)$ are disregarded as being unimportant. Though not explored in this paper, an alternative plotting method using fixed column widths, or widths proportional to the interaction index values may be considered.

C. Data Coverage

A common use case for visualizing a fuzzy measure is to assess the quality of a measure learned from data. In general, a fuzzy measure has $2^n - 1$ variables that can be assigned. During training and evaluation, a single instance uses only n of these variables. As n grows large, it becomes increasingly likely that some variables will never have been encountered in the training process and are assigned based on boundary conditions only. For details, see [4].

The data coverage can be shown in the visualization by including an additional row above the main matrix diagram, labeled "D". As with the bar graph for the interaction indices, the height of the bar in each column shows the relative visitation frequency v for the corresponding subset. Let $h_1, ..., h_m$ be the data set used to construct the fuzzy measure, and let W_{h_i} be the set of subsets $A \subseteq X$ visited in the walk of h_i (See Section II-B). The visitation frequency of a subset A is defined as

$$v(A) = \frac{1}{m} \sum_{i=1}^{m} \mathbb{1}(A \in W_{h_i}). \quad (6)$$

where $\mathbb{1}$ is an indicator function that equals 1 if the condition is true and 0 otherwise.

The visitation frequencies are plotted in yellow and scaled such that the maximum value within each cardinality set is set to 1, or the full height of the row. Each column is scaled such that

$$v'(A) = \frac{v(A)}{\max_{|B|=|A|} v(B)}. \quad (7)$$

Within each cardinality set, the mean visitation frequency is shown as a horizontal line. Columns with greater than average visitation frequencies have their bars darkened above the mean to emphasize the degree to which they may be considered

TABLE II: Example Data

$h(x_1)$	$h(x_2)$	$h(x_3)$	$\pi_{(1)}$	$\pi_{(2)}$	$\pi_{(3)}$
0.74	0.13	0.14	1	3	2
0.94	0.09	0.74	1	3	2
0.97	0.13	0.75	1	3	2
0.92	0.96	0.74	2	1	3
0.91	0.20	0.92	3	1	2

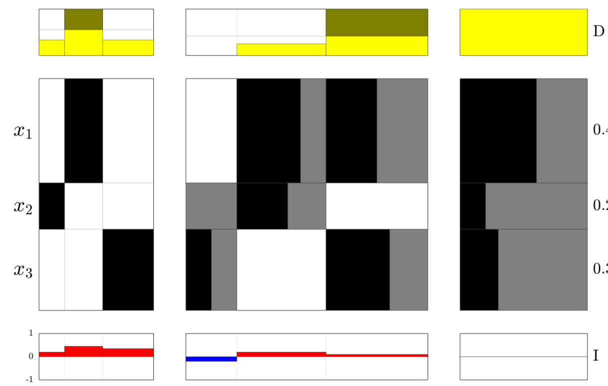


Fig. 6: A data coverage histogram is added above the diagram.

outliers, having been visited by a disproportionately large number of walks from the data set. As with the graph of interaction indices, the visitation frequency graph may be difficult to interpret for subsets with small values of $g(A)$ due to narrow column widths. If these values are considered important, it may be helpful to plot the diagram with equal column widths.

An example data set is shown in Table II and the corresponding visualization is shown in Fig. 6. There are five samples with three sources in this data set and three unique walks. The first thee data samples share the same sort order for the elements and have the walk 1–3–2. The remaining two data samples have the walks 2–1–3 and 3–1–2 respectively. These walks can be more clearly observed in Fig. 7, which shows the FM lattice visualization for this example. While this diagram shows the three walks and the visited subsets, it does not indicate which sources were included in these subsets.[4] Conversely, while the specific walks are not shown

[4]Although the subsets can be identified based on lexicographic sort order, this becomes impractical for large measures.

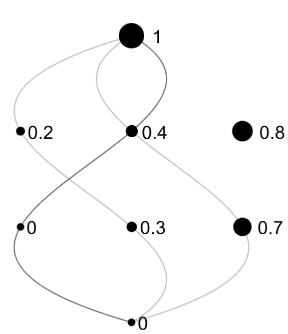


Fig. 7: The FM lattice visualization for the example.

TABLE III: Fuzzy Measure Examples for $X = \{x_1, x_2, x_3\}$

A	$g_1(A)$	$g_2(A)$	$g_3(A)$	$g_4(A)$	$g_5(A)$	$g_6(A)$
$\emptyset$	0	0	0	0	0	0
$\{x_1\}$	0	0.33	0	1	0.2	0.86
$\{x_2\}$	0	0.33	0	1	0.3	0.03
$\{x_3\}$	0	0.33	0	1	0.5	0.05
$\{x_1, x_2\}$	0	0.67	1	1	0.5	0.98
$\{x_1, x_3\}$	0	0.67	1	1	0.7	0.91
$\{x_2, x_3\}$	0	0.67	1	1	0.8	0.42
X	1	1	1	1	1	1

in the weighted matrix visualization of Fig. 6, the visitation frequency of each subset is clearly seen. In this example, we note that the subset $\{x_2, x_3\}$ is never visited by this data set. Furthermore, the darkened bars on the subsets $\{x_1\}$ and $\{x_1, x_3\}$ indicate their dominance, as they were each visited by 60% of the walks.

IV. Examples

A. Three-Source Measures

We now show several examples of the weighed matrix fuzzy measure visualization for different fuzzy integrals. Table III shows six different fuzzy measures defined for three sources. The first four integrals (measures g_1–g_4) correspond to different OWA operators [15] that demonstrate how the visualization method appears in edge cases.

The first measure (g_1) turns the Choquet integral into the min operator, shown in Fig. 8a. Since the only element to have a non-zero value is the full set X, the diagram shows only one large black region. The two smaller cardinality sets are empty and are shown as thin lines on the left side of the diagram.

The second measure (g_2) corresponds to a mean operator, defined such that $g(A) = |A|/|X|$. The diagram in Fig. 8b shows equal row sizes and uniform black bar sizes, indicating identical Δg values. We note that these vertical "stripes" are characteristic of averaging operators in which all sources are treated equally. Note also the lack of interaction index bars on cardinality sets greater than 1.

The third measure (g_3) is a median operator shown in Fig. 8c, in which $g(A) = 1$ for all subsets where $|A| \geq 2$ and 0 otherwise. Like the min operator, the cardinality one set is shown only as a thin line on the left side of the diagram. The full value of the measure is assigned in the cardinality two set, showing all black boxes. Since there is no more room for improvement, the full cardinality three set is all gray.

The fourth measure (g_4) is a max operator shown in Fig. 8d, where all subsets except the empty set are assigned a value of one. Similar to the median operator, all value is assigned in the first cardinality set (shown with black boxes) and the remaining sets are drawn in gray.

The fifth measure (g_5) is an example of an additive fuzzy measure in which $g(A \cup B) = g(A) + g(B)$. The diagram in Fig. 8e shows differing row heights indicating the Shapley indices and a lack of interaction index bars.

The sixth measure (g_6) shows a fuzzy measure with a dominant source. Fig. 8f shows a much wider row for source x_1 indicating its larger Shapley index. Also, the shaded regions in the x_1 row are mostly black, which means that x_1 contributes the most to the subsets that it is a part of.

B. Multicriteria Decision-Making

Besides aggregation, fuzzy integrals can be used for multi-criteria decision-making (MCDM) as a way of comparing the importance of various subsets of criteria. Table IV shows a fuzzy measure used in a MCDM example from [11]. This measure is a learned representation of the evaluation of five different individuals against four judging criteria, x_1 to x_4. The weighted matrix visualization of the measure is shown in Fig. 9. Each row represents a different scoring criteria. The integral of this measure with each candidate's performance scores gives a combined score that can be used to rank the individuals.

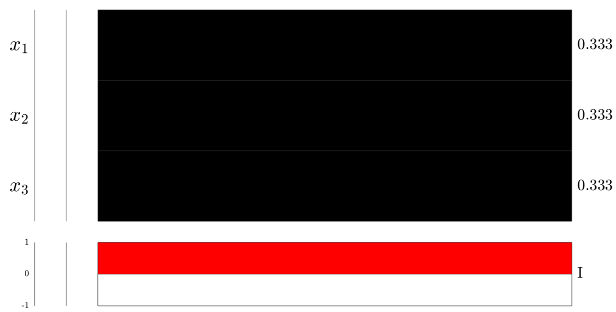
The visualization highlights the interactions between the scoring criteria. We notice first that the cardinality one set is shown only as a thin line, indicating that at least two criteria are needed to perform an assessment. The fourth criteria is dominant, with the largest Shapley index and widest row. It also remains black throughout the width of the diagram showing that it is the main contributor to the value of each subset for which it is a part. The third criteria is shown with a low Shapley index and a narrow row that has no black markings until the rightmost, complete set. This confirms that it plays a small role in the aggregation, only contributing once all other criteria have been considered. We can reason that the first two criteria behave somewhere between these two extremes, with the first criteria being somewhat more important due to its larger Shapley index and row width.

C. Embedded OWA Operators

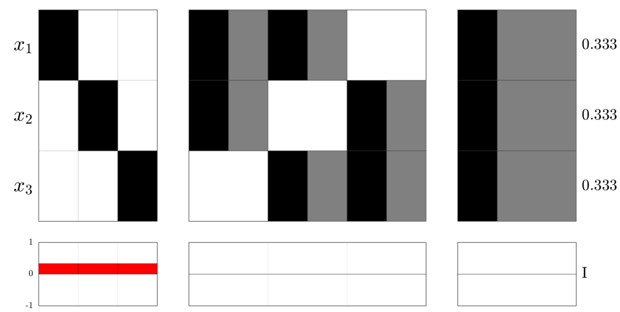
The expressiveness of the fuzzy integral allows it to represent both simple and complex operators simultaneously. Fig. 10 shows the weighted matrix visualization for a fuzzy integral with an *embedded OWA operator*. In this example, $g(A) = 0.4$ when $|A| = 2$ and $g(A) = 0.7$ when $|A| = 3$. The remaining values are randomly selected based on the constraints. We note that this diagram shares several of the characteristics of OWA operators from Section IV-A. In particular, we notice the vertical stripes formed by the black bars in the cardinality three and four sets. These show that all sources contribute equally when considering sets of three or four elements, which is a property of an OWA operator. Observing trends such as these stripes in the visualization can help identify substructure such as embedded OWA operators within the fuzzy measure.

D. Measures Learned From Data

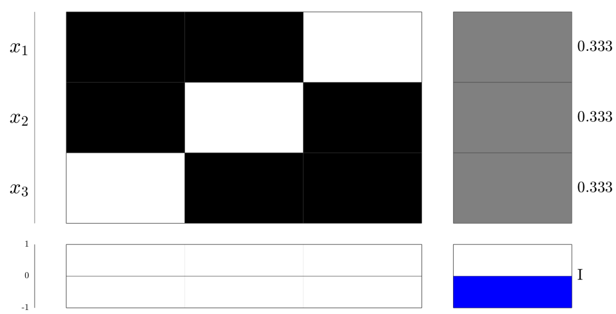
A common use case for the fuzzy measure is to aggregate multiple sources of information using a fuzzy integral. There are several methods for learning a fuzzy measure from data [4, 16]. In [17], a fuzzy measure is learned from the output of seven heterogeneous neural network architectures on a classification problem. Fig. 11 shows one of the fuzzy measures learned from this data set. From this diagram, we can deduce that the learned measure is acting mainly as a minimum operator, based on the narrow columns for all cardinality sets except the full set. The vertical striping pattern suggests that the measure could be represented as an OWA operator without



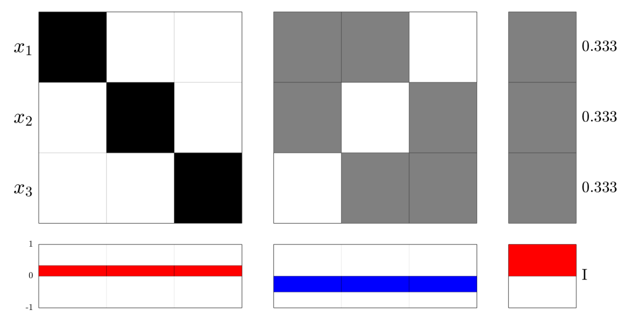
(a) A min operator (g_1).



(b) A mean operator (g_2).

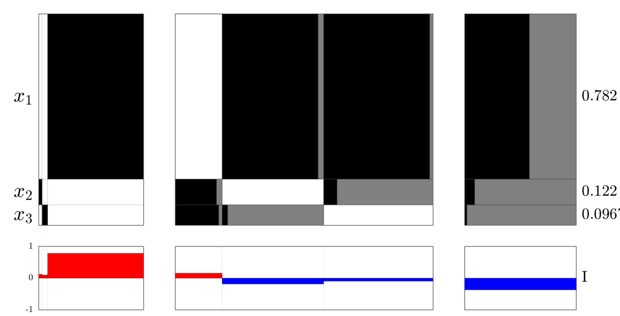


(c) A median operator (g_3).



(d) A max operator (g_4).

(e) An additive fuzzy measure (g_5).



(f) A fuzzy measure with a single dominant source (g_6).

Fig. 8: Weighted matrix fuzzy measure visualizations for three inputs.

TABLE IV: Multicriteria Decision-Making Example

A	$g(A)$
$\emptyset$	0
$\{x_1\}$	10^{-6}
$\{x_2\}$	10^{-6}
$\{x_3\}$	10^{-6}
$\{x_4\}$	10^{-6}
$\{x_1, x_2\}$	10^{-6}
$\{x_1, x_3\}$	10^{-6}
$\{x_1, x_4\}$	0.666667
$\{x_2, x_3\}$	10^{-6}
$\{x_2, x_4\}$	0.389743
$\{x_3, x_4\}$	10^{-6}
$\{x_1, x_2, x_3\}$	10^{-6}
$\{x_1, x_2, x_4\}$	0.666667
$\{x_1, x_3, x_4\}$	0.666667
$\{x_2, x_3, x_4\}$	0.389743
$\{x_1, x_2, x_3, x_4\}$	1

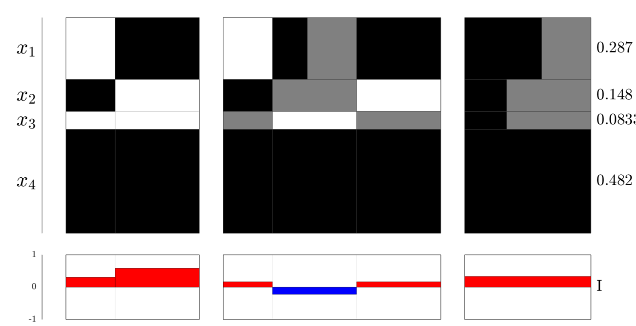


Fig. 9: The fuzzy measure from the MCDM example.

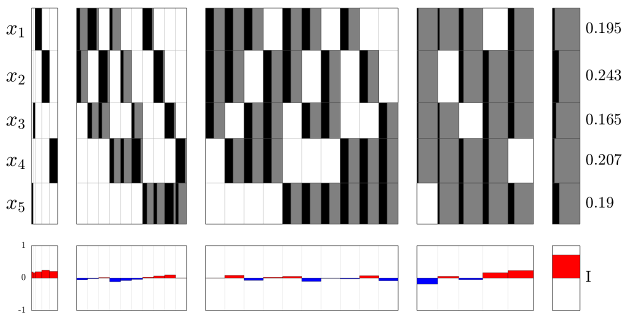


Fig. 10: A fuzzy measure with an embedded OWA operator.

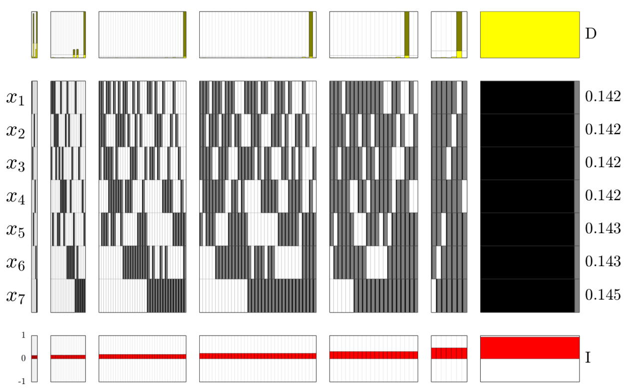


Fig. 11: A fuzzy measure learned from data.

significant loss of accuracy. The uniform interaction indices across cardinality sets, rising to the largest value for the full set, and the near uniform row widths and Shapley indices imply that all sources have roughly equivalent utility.

The data visitation is shown by the yellow histogram at the top of the diagram. Within each cardinality set, the visitation frequency is concentrated mainly on a single subset. This shows that despite the large amount of training data used to learn the fuzzy measure, almost all the data utilized a single walk. The FM lattice visualization of this measure in Fig. 12 shows that most subsets were in fact visited by at least one data sample, but only a few walks are dominant.

V. Conclusion

The weighted matrix visualization is a useful tool for understanding the properties of a fuzzy measure. As an explanatory graphic, it provides a way to quickly see which sources are being utilized and in what combinations. It also provides enough detail into the inner workings of the measure to allow for exploration into specific interactions. When used in conjunction with other visualization approaches, such as the FM lattice, a practitioner can identify if a fuzzy measure is suitable for a particular application, or if another approach may be more appropriate. For instance, when learning the fuzzy measure parameters from a data set, one may wish to know if the problem requires the full expressive power of the fuzzy integral, or if a simpler operator such as an OWA would suffice.

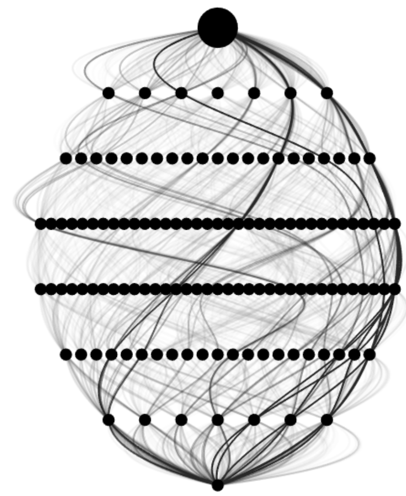


Fig. 12: The FM lattice visualization for the data-driven example, as described in [1].

Although this visualization approach is possible for an arbitrary number of sources, it becomes less feasible and harder to interpret as the number of sources grows large. While our method is designed for static display in print, it may be possible to utilize an interactive version of the diagram that can better handle larger problem sizes. Other variations on this approach may be helpful for exploring specific aspects of a problem, such as mapping column widths to the interaction index or data visitation frequency. Since the true value of any visualization technique comes from real world use, we have made the code available on Code Ocean (https://codeocean.com/capsule/6663959).

References

[1] A. J. Pinar, T. C. Havens, M. A. Islam, and D. T. Anderson, "Visualization and learning of the Choquet integral with limited training data," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Naples, Italy, July 2017.

[2] A. Lex and N. Gehlenborg, "Sets and intersections," *Nature Methods*, vol. 11, no. 8, pp. 779–779, Aug. 2014.

[3] B. Alsallakh and L. Ren, "PowerSet: a comprehensive visualization of set intersections," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 361–370, Jan. 2017.

[4] M. A. Islam, D. T. Anderson, A. J. Pinar, and T. C. Havens, "Data-driven compression and efficient learning of the Choquet integral," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 4, pp. 1908–1922, Aug. 2018.

[5] B. Murray, M. A. Islam, A. J. Pinar, T. C. Havens, D. T. Anderson, and G. Scott, "Explainable AI for understanding decisions and data-driven optimization of the choquet integral," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Rio de Janeiro, July 2018.

[6] E. R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, USA, 1986.

[7] J. Bertin, *Graphics and Graphic Information Processing*, Walter de Gruyter, Berlin, 1981.
[8] R. Krauss, “Grids,” *October*, vol. 9, pp. 51–64, 1979.
[9] L. S. Shapley, “A value for n-person games,” in *Contributions to the Theory of Games (AM-28)*, H. W. Kuhn and A. W. Tucker, Eds., vol. 2, pp. 307–318. Princeton University Press, Princeton, 1953.
[10] G. Owen, “Multilinear extensions of games,” *Management Science*, vol. 18, no. 5-part-2, pp. 64–79, Jan. 1972.
[11] M. Grabisch and M. Roubens, “Application of the Choquet integral in multicriteria decision making,” *Fuzzy Measures and Integrals-Theory and Applications*, pp. 348–374, 2000.
[12] M. Grabisch, T. Murofushi, and M. Sugeno, Eds., *Fuzzy Measures and Integrals: Theory and Applications*, Studies in Fuzziness and Soft Computing. Physica-Verlag Heidelberg, 2000.
[13] G. Choquet, “Theory of capacities,” *Annales de l’Institut Fourier*, vol. 5, pp. 131–295, 1953.
[14] M. Grabisch, “k-order additive discrete fuzzy measures and their representation,” *Fuzzy Sets and Systems*, vol. 92, no. 2, pp. 167–189, Dec. 1997.
[15] R. Yager, “On ordered weighted averaging aggregation operators in multicriteria decision making,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988.
[16] D. T. Anderson, S. R. Price, and T. C. Havens, “Regularization-based learning of the Choquet integral,” in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, July 2014, pp. 2519–2526.
[17] M. A. Islam, D. T. Anderson, A. J. Pinar, T. C. Havens, G. Scott, and J. M. Keller, “Enabling explainable fusion in deep learning with fuzzy integral neural networks,” *IEEE Transactions on Fuzzy Systems (accepted)*, 2019, arXiv: 1905.04394.