Human-Oriented Fuzzy Set Based Explanations of Spatial Concepts

Brendan Young
Electrical Engineering and
Computer Science Department
University of Missouri
Columbia, MO, USA
bmywzx@umsystem.edu

Derek T. Anderson
Electrical Engineering and
Computer Science Department
University of Missouri
Columbia, MO, USA
andersondt@missouri.edu

James M. Keller
Electrical Engineering and
Computer Science Department
University of Missouri
Columbia, MO, USA
kellerj@missouri.edu

Fred Petry
U.S. Naval Research Lab
Stennis Space Center, MS

Chris Michael
U.S. Naval Research Lab
Stennis Space Center, MS

Blake Ruprecht
Electrical Engineering and
Computer Science Department
University of Missouri
Columbia, MO, USA
bcrf53@mail.missouri.edu

***Abstract*—This article explores generating fuzzy explanations for applications involving spatial intelligence. Two approaches are presented that produce an explanation of a spatial concept from a fuzzy attributed relation graph acquired via a human-in-the-loop algorithm. The first method automatically yields a total spatial concept with explicit graphical and linguistic explanations by identifying the medoid, i.e., an explanation is derived directly from an observed example of the concept. This has the benefit of being both consistent and correct across all relations. The second method uses Zadeh's extension principle to produce a parts-based explanation, however this solution is not guaranteed to be a previously known member of the concept. However, it has some advantages over the medoid that will be shown. To illustrate these methods, two examples are presented. Example one shows the generation of a single explanation for a two-object concept, which highlights how each method creates an explanation for a single relation. The second example generates an explanation for a multi-object concept, illustrating how an explanation is created for more complex concepts.**

Keywords—Spatial Attribute Relation Graph, Spatial Relations, Fuzzy Explanation, Explainable Artificial Intelligence

I. Introduction

To classify data into preexisting classes, it is first necessary to have a *concept* of the features and attributes exhibited by those classes. One attribute that is extremely useful in fields like computer vision [1], robot navigation [2], signal-to-text [3, 4], scene understanding [5-7], and human-robot interaction [8] is relative spatial position. This should come as no surprise as this is one of Gardner's seven core human intelligences [9]. When the concept features and attributes are explicitly known, it is possible to determine if an input is a member of a class and provide an explanation for why it is or is not a member. A simple approach that could be attempted is to define a class by a single prototypical example. This is desirable because a set of features can be determined to be in a class by comparing its similarity to the class's prototype. However, this simple approach is ill-suited and insufficient to define most real-world concepts, such as objects in computer vision, even something as simple as a chair, by a single example.

While a single example (e.g., prototype) cannot sufficiently encapsulate most concepts, an approximation can be achieved using multiple examples to develop a flexible concept that allows for variability and uncertainty. In this form, the concept may be stored more implicitly as an aggregate of examples. This, however, comes at the cost of both human readability and understandability. This can be shown by realizing that after a concept has been defined using a significantly large number of examples, it is often difficult to see past individual examples and understand the concept as a whole. This is important when a human must interact with a machine over an extended period, such as in the case of online learning or human-in-the-loop (HITL) recommender algorithms. To this end, new methods are needed to summarize concepts in order to communicate concepts clearly and concisely. This is important for human-machine trust and ensuring that a machine learns valid concepts and parts-based reasoning.

This paper examines two methods for simplifying an implicit concept into an explicit explanation. Method one is a holistic approach in which prototypes are picked by examining the concept as a whole. This approach is based on aggregating similarities of user typicalities in an attributed relation graph (ARG) [10], where spatial relations are modeled by utilizing histograms of forces (HoF) [11]. The second approach uses an atomic, parts-based method to create a prototype where each relation is examined individually to find a sufficiently useful prototypical relationship. An algorithm is proposed that solves this task based on a shape-preserving extension [12] of Zadeh's extension principle [13]. As we show in a later section, these are different explanations, and it is not universally clear if there is a correct method across all cases needing an appropriate explanation. It is more likely that each approach is better suited for different contexts and applications.

The remainder of the paper is organized as follows. Section II explores background knowledge from previous work. Section III discusses the implementation of the two methods of developing a prototype. Section IV shows how the two methods

2023 IEEE International Conference on Fuzzy Systems (FUZZ) | 979-8-3503-3228-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/FUZZ52849.2023.10309800

interact on a single relationship, while Section V uses an example to show how the two methods act on a multi-object concept. Section VI concludes the paper and discusses future work.

II. Related Work

The last section aims to identify a clear objective and discuss our contributions. However, it should be noted that our work is motivated by applications like aided target recognition (AiTR) and cases in which a user (e.g., analyst) is performing complex and repetitive operations across large spatiotemporal geospatial data (e.g., site recognition, land cover classification, shoreline detection, etc.). As such, there is related work in concept learning, HITL recommender systems, active learning, online learning, stream classification, and parts-based learning and reasoning. The reality is that there is simply too much related work to cover for this article. This is further compounded if the reader includes explainable AI (XAI) [14] and related fuzzy explanation research, e.g., fuzzy description logics [15], linguistic summarization [16], fuzzy quantification [17], fuzzy referring expression generation [18], etc. The following sub-sections review the spatial and data representation concepts required to understand the current article. On a final note, while other work exists, we are unaware of any directly related work that can be experimentally compared apples-to-apples to the current article.

Specifically, we focus on parts-based concept modeling via fuzzy ARGs where a human is in the loop providing a small and very limited set of labeled data. This users provides weights that are a measure of the degree to which an example belongs to the target concept. This problem formulation differs from contexts where a large and diverse set of labeled data can be collected and used for training, e.g., supervised deep learning. The reader can refer to [2, 5-7] for spatial intelligence in modern deep learning, from scene graphs to learning spatial relations and subsequent decision-making. These works are not considered herein as they do not have a human-in-the-loop, and they learn distributed and implicit spatial concepts. It is not clear at this moment how to extract the necessary spatial knowledge from these concepts. As such, we focus on ARGs, an explicit representation structure.

A. Spatial Relationships as Histogram of Forces (HoF)

Our aim is to build and explain spatial concepts from instances presented to a human. To effectively achieve this goal, the spatial relationships between objects need to be clearly represented. There have been several approaches for describing spatial orientation between objects in an image [19–23]. The histogram of forces (HoF), proposed by Matsakis in [11], provides a fuzzy set-based approach that models the relative position between objects that can incorporate some aspects of both size and shape. The HoF has been used to evaluate the directional spatial relationships between two objects. This occurs by evaluating the amount that a second object is at angle θ in relation to a first object. This is then used to create a histogram with the angles represented on the x-axis and the amount of force acting between the objects for the specified angle represented on the y-axis. The HoF has been shown to be affine invariant [24], allowing for a robust similarity measure between two scenes which can handle arbitrary changes in rotation, scale, and translation. The HoF has been implemented in different ways. Two main ways are separated by the method used to calculate the effect of distance on the forces. The histogram of gravitational forces (HoGF) mimics equations that describe gravity by multiplying the forces by $1/d^2$, where d is the distance between pixels in each object. This increases the forces of objects that are closer together and decreases the forces for objects that are further apart (see Fig. 1a). The histogram of constant forces (HoCF) results when the distance between objects is disregarded, as is shown in Fig. 1b. This form of the HoF is useful for calculating the relative angle and size between objects in situations where the distance should not be impactful for the desired output, such as cases where a distance robust classifier is desired. Since, in many cases, the relative distance between objects is a necessary component for defining a concept, the HoGF is used for the remainder of the paper. These histograms can be viewed as non-normalized generalized fuzzy subsets with a domain of [0, 360]. Normalization is not desirable since the heights of these sets provide metric information about the relative position of the two objects [24].

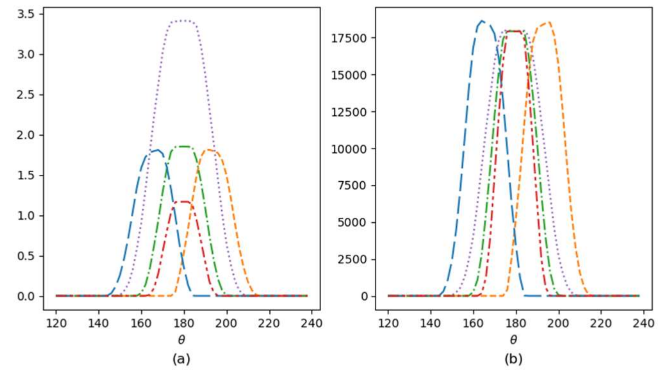

Fig. 1. Histogram of forces (HoF) showing pairs of spatial relationships between the black and colored objects in Fig. 3. For example, the blue HoF is the spatial relation between the black (referent) and blue (argument) object. Fig. 1a is the Histogram of Gravitational Forces (HoGF) and Fig. 1b is the Histogram of Constant Forces (HoCF).

B. Spatial Attributed Relation Graph (SARG)

To represent a spatial concept, a graph-based structure is used [8], where $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, vertices $(\mathbf{V})$ are objects (concept *parts*) and edges $(\mathbf{E})$ are object/part relationships. Example vertex attributes, $V_i \in \{V_1, V_2, \dots\}$, in computer vision include features such as color (e.g., channel statistics, histograms, etc.), texture (e.g., fractals, Gabor, etc.), shape (e.g., moments, morphology), size, etc. An example edge attribute between V_i and V_j, $E_{ij} = r(V_i, V_j)$, is an HoF. Since we are working with graphs in a spatial context, these attributed data structures are often referred to as Spatial ARGs (SARGs). Different example uses of SARGs for applications like classification and scene understanding in domains such as computer vision and remote sensing are illustrated in Fig 2.

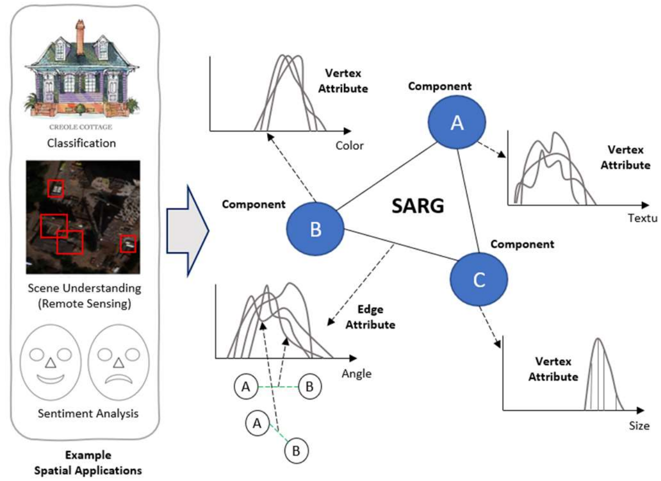

Fig 2. Example computer vision applications represented as a SARG. While our article is focused on spatial attributes, our work naturally extends to other vertex and edge features.

III. Methods

A. Concept Modeling as a Collection of SARGs

Herein, multiple potentially suboptimal examples are used to approximate a target concept and its associated uncertainty. Each example is a SARG and is associated with a user-defined confidence value (typicality) that informs us about the degree to which it supports a concept. Let SARG i be denoted by $\mathbf{G}_i$ and its typicality is μ_i. Therefore, a concept is a collection, $\overline{\mathbf{G}} = \{(\mathbf{G}_1, \mu_1), (\mathbf{G}_2, \mu_2), \ldots, (\mathbf{G}_N, \mu_N)\}$, of N user-attributed SARGs. Collection $\overline{\mathbf{G}}$ can be used to determine the degree to which a new SARG belongs to a concept via an aggregated similarity of its attributes. The next two subsections outline this process and, ultimately, a method to produce an explanation that can be concisely conveyed to a user.

B. Method 1: Holistic Explanation (Medoids)

The first method we explore to summarize and explain $\overline{\mathbf{G}}$ is based on medoids. This can be desirable because it selects an example from our data, which assures that the resulting prototype is guaranteed to be an example of the concept. We select a prototype by picking the example that has both a high relative confidence and is most like the other high confidence examples (see Algorithm 1). Let N be the number of examples in $\overline{\mathbf{G}}$, M is the number of edge attributes (HoFs) in our concept, μ_i is the user-provided typicality for $\mathbf{G}_i$, $C_{i,k}$ is the k-th edge attribute of $\mathbf{G}_i$, and $S(C_{i,k}, C_{j,k})$ is the similarity between the two edge attributes (HoFs). The medoid is defined herein as

$$M(C) = \arg\max_{i \in N} \sum_{j=1}^{N} \sum_{k=1}^{M} \mu_i \mu_j S(C_{i,k}, C_{j,k}). \qquad \text{Eq. 1}$$

The similarity measure used in this paper (originally used on HoFs [25]) is the cross-correlation, defined as

$$S(h_1, h_2) = \frac{\sum_\theta h_1(\theta) h_2(\theta)}{\sqrt{\sum_\theta h_1^2(\theta)}\sqrt{\sum_\theta h_2^2(\theta)}}, \qquad \text{Eq. 2}$$

Algorithm 1: Concept Reduction via Medoid

```
INPUT: Concept Ḡ
INIT: r = -Inf and p = 1
for i = 1 to N do                         # iterate over examples
    s = 0
    for j = 1 to N (i ≠ j) do  # consider all example pairs
        for k = 1 to M do        # loop over edge attributes
            s += μ_i μ_j S(C_ik, C_jk)    # weighted similarity
        end for
    end for
    if s > r
        s = r    and    p = i        # top score and its index
end for
RETURN: p.                       # index of our medoid in N
```

where h_1 *and* h_2 are HoFs and $h_1(\theta)$ is the value of the force angle θ in h_1. Since this solution uses an example from our data (directly encountered concepts), we can use the original example as a visual representation of the concept, and a corresponding linguistic explanation can be produced.

C. Method 2: Atomic Explanation (EP/SPWA)

A second approach is to aggregate the individual attributes in $\overline{\mathbf{G}}$. Since a single HoF is an unnormalized fuzzy set, Zadeh's extension principle (EP) can be used to combine HoFs. Herein, we calculate the user typicality weighted average (WA),

$$F(C_k) = \frac{\sum_{i=1}^{N} \mu_i C_{i,k}}{\sum_{i=1}^{N} \mu_i}, \qquad \text{Eq. 3}$$

which is an accurate calculation of the directional relationship between two objects. This process generates a new relation, fuzzy set $F(C_k)$, that has likely not been directly encountered in our data. However, Zadeh's EP results may not be intuitive. Namely, his formulation of the EP is

$$[f(A_1, \ldots, A_N)](y) = sup_{\mathbf{x}|y=f(\mathbf{x})}(A_1(x_1) \wedge \cdots \wedge A_N(x_N)),$$

where f is our function (e.g., weighted average), $A_i(x_i)$ is the i-th fuzzy set evaluated at x_i, and $\wedge$ is a t-norm (Zadeh let it be the min). The problem is the min of the membership degrees, which restricts our resultant HoF, $F(C_k)$, according to the furthest object. This choice leads to HoFs that are seemingly "cutoff," resulting in inaccurate estimates for the size and distance between objects. This issue is illustrated in Fig 4c.

In [12], this EP limitation is remedied in the context of the fuzzy integral, a solution we coined the Shape Preserving Fuzzy Integral (SPFI). Herein, we are not modeling weights for combinations of inputs, i.e., there is no fuzzy measure. Instead, we are using a single weight from a human. As such, we refer to the reduced operator as the Shape Preserving Weighted Average (SPWA), see Algorithm 2.

In Algorithm 2, $H(C_{ik})$ is the maximum height of the HoF C_{ik}, and D_{ik} is the height normalized C_{ik}. Step one is to normalize the heights of all the HoFs, which makes the modified HoFs normalized general fuzzy sets. This allows the EP to be applied while mitigating the above-mentioned interpretation issues. The result is multiplied by the weighted average of all the previously normalized heights. This solution allows for a collection's reduction (via aggregation) into a single example on a per relationship (parts) basis. Under this method, the linguistic solution is easy to obtain using the directional information of the aggregated spatial relation. The graphical explanation, however, is harder to obtain and compute. Both details are expanded on in the next two sections.

IV. Example 1: Single Parts-Based Explanation

In this section, a simple yet revealing example is shown to highlight differences between concept explanation techniques. This example is focused on a single spatial relation between two objects in an overall concept. Specifically, in Figure 3, the first object, notated as A, is represented by the black box, and the colored circles are five instances of object B, which has been moved around to represent variability in the concept.

Fig 3. Simple relative spatial relations task where a referent (square) is shown in black and different arguments (circles) are color coded (red, blue, etc.). The corresponding HoF for each square-circle combo is shown color coded in Figure 4.

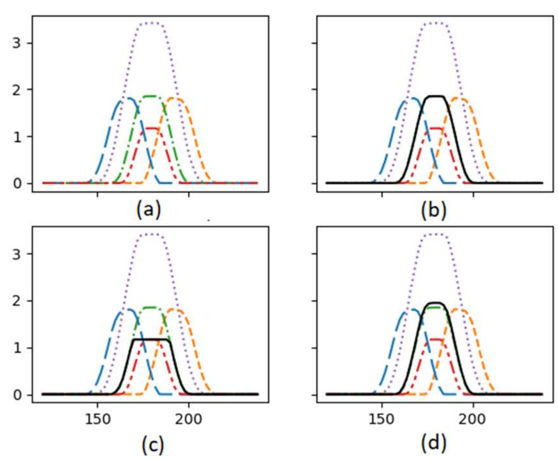

Fig 4. (a) Spatial relations from objects in Fig. 3, (b) result (in black) for medoid, (c) result for Zadeh EP, and (d) SPWA. The results were computed for a typicality of 1 on each relation.

Based on Figure 3, it would be semantically expected that the center configuration of object B would be the best prototype for the examples due to it being at a location equidistant from the other examples. The HoF is computed for each configuration of object B and the resulting HoFs are shown in Fig 4(a). The methods shown in this paper were then ran based on the previously mentioned methods for the example and the results are shown in Fig 4(b-d).

Algorithm 2: Parts-Based Concept Reduction via SPWA

```
INPUT: Attribute C_k
D_ik = C_ik / H(C_ik)   # normalize each w.r.t. height, ∀i ∈ N
D̃_k = EP(D_1k, …, D_Nk)   # calculate WA via α-cut based EP
r = WA of H(C_ik)   # calculate WA of heights
R_k = r D̃_k   # scale EP result by r
RETURN: R_k
```

Each of the methods for creating a summary of the concept is successful in generating a seemingly reasonable explanation of a relationship. The medoid selected the example of the object in the center as the ideal prototype. The EP can correctly approximate the relative angle between the two objects; however, the magnitude indicates that the object is farther out than expected based on Fig 3 from the properties developed in [24]. The SPWA summarizes the concept by selecting a shape that is very similar to the center variation of object B.

To show the robustness of the two proposed algorithms, the complexity of Example 1 can be slightly increased by removing the green center object from Fig 3. This makes the representation of the concept sparser and removes the most ideal example of the concept. The remaining HoFs are shown in Fig 5(a), with the aggregated HoFs in Fig 5(b-d). Similar to the previous example, the ideal prototypical example should remain in the center of the four variations.

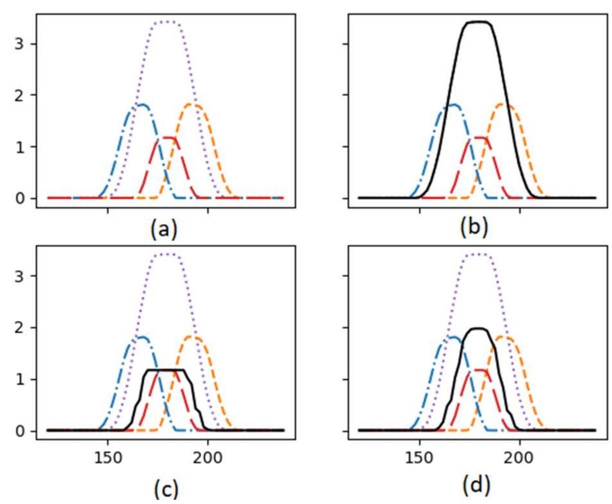

Fig 5. (a) Spatial relations for the red, purple, yellow, and red arguments in Fig. 3; i.e., no green object., (b) result (in black) for medoid, (c) result for the EP, and (d) for SPWA.

The summarization in this slightly more complex example is different from before. The medoid selects the variation closest to the reference object. This is expected because the HoF of the closest object has the most overlap over the other three objects. This shows that, while the medoid finds the best existing example of the set to use as a prototype, it will not be able to find an ideal prototype in a sparse set of variations where the ideal prototype is not provided. The EP has very little change to its result from the previous example. It shows the correct relative angle between the two objects but puts the prototype further out than expected from Figure 3. Last, the

SPWA method manages to select an example that is located at the center between the four variations. This shows that the SPWA method is more robust in cases where some variations are missing, and it can find a prototype that has not been shown to the system.

V. Example 2: Total Concept Explanation

Next, a more complicated yet analytically tractable multi-object example is provided to show the differences between the two concept explanation techniques. This example is focused on three spatial relations between three objects that define the overall concept. Specifically, in Figure 6, the box is object A, the three colored circles to the right of object A are where object B has been moved, and the three colored triangles below object A are where the position of object C has been varied.

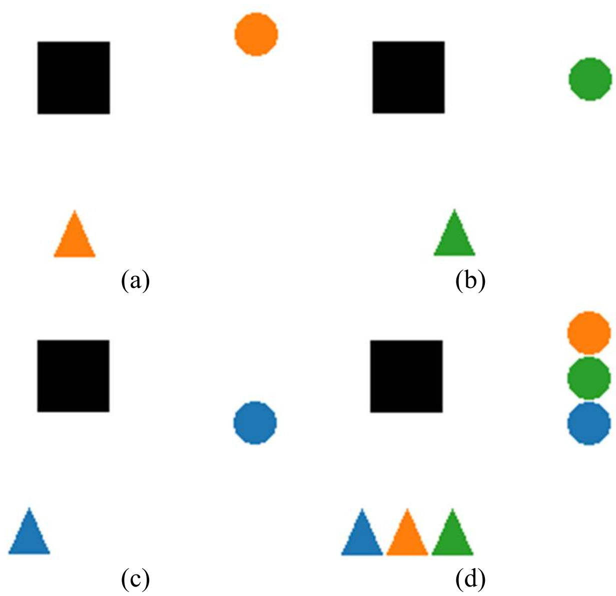

Fig 6. A slightly more complex spatial relationship showing a set of spatial relations between three objects. The three objects are object A (square), object B (circle), and object C (triangle). object B and object C are varied in position. Figures a-c are three provided examples (i.e., SARGs), and figure d shows how the examples relate to one another. The corresponding HoF for each is shown in Figure 7.

In this example, unlike in Example 1, the best prototype is not clear-cut. Based on Fig. 6, the best prototype for object B, by itself, is the green case. However, the best prototype for object C is examined by itself is the orange case. These two relationships show how the best prototype is not always clear cut and can vary from user to user and relationship to relationship.

This ambiguity in selecting a best prototype is exemplified by how the different methods handle this example. Both the medoid and the SPWA show different ways of handling a concept where the best prototype is not immediately clear. The differences are shown in Fig 7. The medoid looks at all the objects as a whole and picks the case that is most like the other cases. For this example, the blue case is not selected due to the relationship from A-B being at a completely different angle than in the orange or green case. The orange and green cases were almost the same score (0.4692 vs. 0.4686) when selecting the medoid. This method can be useful when dealing with cases where outliers are present within the concept, as the resulting prototype is less likely to be affected by an outlier. However, this will never choose a prototype that is an intermediary between two cases.

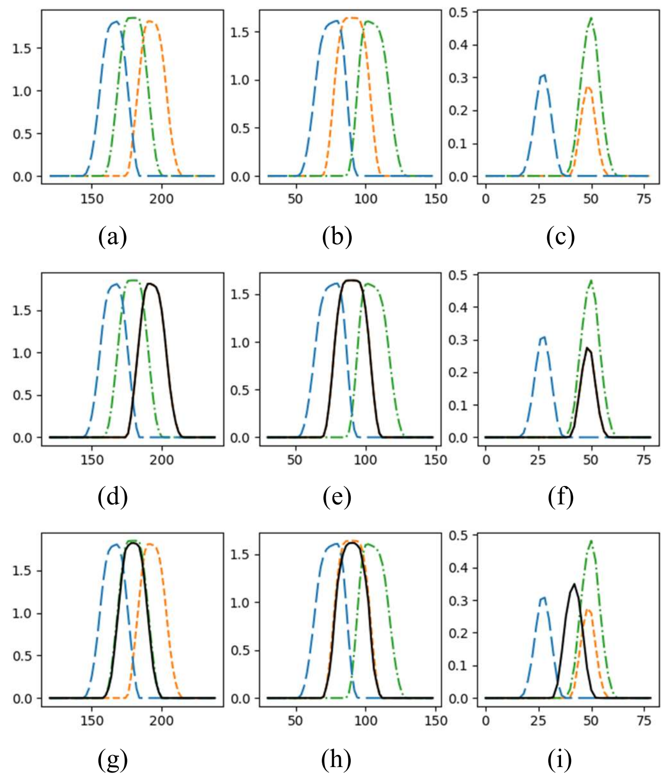

Fig 7. *(a) Spatial relations of the referent to object A in Fig. 6, (b) spatial relations of the referent to Object B, (c) spatial relations of A to B, (d-f) medoid result for spatial relations in (a-c), (g-i) SPWA for relations in (a-c). Results were computed for a typicality of 1 on each relation.*

The SPWA, rather than examining the entire concept to find a prototype, examines each relationship individually and creates a set of prototypical relationships. When examining the object A-object B and the object A-object C relationships, a version that closely resembles the green case in the object A-object B relationship and the orange case in the object A-object C relationship is generated as the prototype. In the case of the object B-object C relationship, it generates a new HoF that gives the same angle as the green version of object B and the orange version of object C. The maximum of the generated HoF also gives an approximate value for what an HoF would be for the green version of object B and the orange version of object C. This result generates a prototype that is different from any of the cases of the concept given and tries to find a version that blends the three cases together.

The XAI community has shown that there are many ways to provide an explanation to users. Common examples include graphical, statistical, local, and linguistic methods. In the current article, an explanation can be provided to a user in at least two ways. The first method is to produce a linguistic explanation of a concept. Using the resultant fuzzy sets from the medoid or SPWA, along with the fuzzy rule-based approach outlined in [4], linguistic spatial relations can be generated. With this method, the linguistic explanation for both the medoid and the SPWA would be that object B is "to the right" of object A, that object C is "below" object A, and that object B is "above and to the right of" object C. This allows for the three rules to provide an explanation that is somewhat easy to visualize; however, it leaves a lot of variability to the user's imagination. Note that for the three instances of the spatial concept in Fig.6, the linguistic descriptions would be as follows:

Object B is "to the right" of Object A,
Object C is "below" Object A,
Object B is "above and to the right of" Object C.

As shown in [4], linguistic relations like these are useful in HITL applications like sketch-to-text, robotic navigation, and computer vision (e.g., AiTR). A linguistic description is a summarization process, which will result in *information loss*. However, in many HITL applications, time can be of the essence. For example, consider human-robot teaming task in a high stress environment. If the human and robot are paired and are working to achieve a common goal, a human might not have time to stop and look at a graphical explanation. A linguistic statement such as the above, stated verbally, could be ideal. The point being, in this article we are showing how to generate different explanations, but the harder task of when and where to best use them is not always clear.

Fig 8. Graphical explanation of SPWA for the HoFs found in Fig. 7g-i overlayed on Fig. 6d.

The second type of explanation that our current article can produce is a graphical explanation. The graphical explanation of the medoid is the simplest operation, as it is merely one of our already observed data points (see Fig. 6a). The graphical explanation for the EP or SPWA requires computation. In [26], we presented a way to effectively back project and combine multiple HoFs, a method we coined "fuzzy silhouettes." An optimization procedure is then run to search for an optimal parts placement configuration. In Fig. 8, we show the result of this operation for the SPWA result in example 2. Black lines are the answer. The result is one part of our "green example" and one part of our "orange example". This is simply due to the way that we set up example 2. For example 1, if we did not place the ideal solution in our data, the EP and SPWA could still identify the solution or another location that does not exist in our input data, and the result can be explained to a user in the form of a graphical result.

VI. Summary

In this article, we investigated ways to simplify a complex implicit spatial concept into an explicit concept that can be communicated to a user for applications requiring a human in the loop (HITL). To this end, we explored two algorithms. The first method is medoid based, while the second involves a modified extension principle (EP). Overall, the medoid yields a graphical and linguistic explanation, and the solution is satisfying in the regard that it is one of our data points and, therefore, is globally consistent across the parts. The second process yields a reduced parts-based explanation that is an aggregated combination of our data. As we demonstrated, this solution can be better than that of the medoid. We also showed that linguistic statements can be produced, and a method for back projecting and discovering an ideal configuration exists if a graphical explanation is required. However, while an improved parts-based explanation might be useful for some applications, this could impact some problems as the solution is not guaranteed to be globally consistent. Lastly, two numeric and analytically tracible examples were provided to show the reader each step in the proposed work. Specifically, Example 1 focused on a single two-object and single-part answer. Example 2 focused on a multi-part solution.

In this article, we established a groundwork for how to make different types of explanations for spatially motivated problems requiring parts-based modeling and reasoning. In future work, we will take these ideas and apply them to one of our geospatially driven HITL tasks, such as site recognition for remote sensing [27]. While we have already used SARGs for classification, we have not yet begun to explore to what degree these explanations help a user close the loop and refine a better machine concept or accomplish a task according to a set of metrics. Specifically, we will look to quantitative domain metrics (e.g., completion time, cognitive workload, situational awareness, task performance, etc.) to help us better understand how much gain is achieved via these different algorithms and their resultant graphical and/or linguistic explanations. Furthermore, in future work, we will explore how to generate a globally consistent prototype and explanation from the proposed SPWA approach.

References

[1] A. Rosinol, A. Violette, M. Abate, N. Hughes, Y. Chang, J. Shi, A. Gupta, and L. Carlone, "Kimera: from SLAM to spatial perception with 3d dynamic scene graphs," CoRR abs/2101.06894, 2021.

[2] M. Feng, S. Z. Gilani, Y. Wang, L. Zhang, and A. Mian, "Relation graph network for 3d object detection in point clouds," IEEE Transactions on Image Processing 30, pp. 92–107, 2021.

[3] X. Yang, K. Tang, H. Zhang, and J. Cai, "Auto-encoding scene graphs for image captioning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[4] P. Matsakis, J. Keller, L. Wendling, J. Marjamaa, and O. Sjahputera, “Linguistic description of relative positions in images,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 31(4), pp. 573–588, 2001.

[5] S. Wu, J. Wald, K. Tateno, N. Navab, and F. Tombari, “Scenegraphfusion: Incremental 3d scene graph prediction from RGB-D sequences,” CoRR abs/2103.14898, 2021.

[6] I. Armeni, Z. He, J. Gwak, A. R. Zamir, M. Fischer, J. Malik, and S. Savarese, “3d scene graph: A structure for unified semantics, 3d space, and camera,” CoRR abs/1910.02527, 2019.

[7] M. Haldekar, A. Ganesan, and T. Oates, “Identifying spatial relations in images using convolutional neural networks,” in 2017 International Joint Conference on Neural Networks (IJCNN), pp. 3593–3600, 2017.

[8] S. Guadarrama, L. Riano, D. Golland, D. Go¨hring, Y. Jia, D. Klein, P. Abbeel, and T. Darrell, “Grounding spatial relations for human-robot interaction,” in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1640–1647, 2013.

[9] Gardner, H. (1983). Frames of Mind: The Theory of Multiple Intelligences

[10] B. Ruprecht, D. T. Anderson, F. Petry, J. M. Keller, C. Michael, A. Buck,Scott, C. Davis, “Concept Learning Based on Human Interaction and Explainable AI,” SPIE 2021

[11] P. Matsakis and L. Wendling, “A new way to represent the relative position between areal objects,” IEEE Transactions on Pattern Analysis and Machine Intelligence 21(7), pp. 634–643, 1999.

[12] T. Havens, A. Pinar, D. T. Anderson and C. Wagner, "SPFI: Shape-Preserving Choquet Fuzzy Integral for Non-Normal Fuzzy Set-Valued Evidence," accepted for publication in WCCI, 2018

[13] J.M. Keller, X. Wang, “A fuzzy rule-based approach to scene description involving spatial relationships”, Computer Vision and Image Understanding, 80 (2000), pp. 21-41

[14] A. Barredo Arrieta et al., ‘Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI’, Information Fusion, vol. 58, pp. 82–115, 2020.

[15] Sánchez, D., Tettamanzi, A.G.B. (2006). Reasoning and Quantification in Fuzzy Description Logics. In: Bloch, I., Petrosino, A., Tettamanzi, A.G.B. (eds) Fuzzy Logic and Applications. WILF 2005. Lecture Notes in Computer Science(), vol 3849. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11676935_10

[16] Janusz Kacprzyk, Anna Wilbik, and Slawomir Zadrozny. 2008. Linguistic summarization of time series using a fuzzy quantifier driven aggregation. Fuzzy Sets and Systems 159, 12 (2008), 1485–1499.

[17] M. Delgado, M. D. Ruiz, D. Sánchez, and M. A. Vila, ‘Fuzzy quantification: a state of the art’, Fuzzy Sets and Systems, vol. 242, pp. 1–30, 2014.

[18] N. Marín, G. Rivas-Gervilla and D. Sánchez, “Fuzzy Quantified Relational Properties for Referring Expression Generation,” 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Padua, Italy, 2022, pp. 1-8, doi: 10.1109/FUZZ-IEEE55066.2022.9882695.

[19] K. Miyajima and A. Ralescu, “Spatial organization in 2d segmented images: Representation and recognition of primitive spatial relations,” Fuzzy Sets and Systems 65(2), pp. 225–236, 1994. Fuzzy Methods for Computer Vision and Pattern Recognition.

[20] A. G. and S. M. Hazarika, “Qualitative Spatial Representation and Reasoning: An Overview,” Fundamenta Informaticae, vol. 46, no. 1–2, pp. 1–29, Jan. 2001.

[21] F. Godoy and A. Rodríguez, “A quantitative description of spatial configurations,” in Spatial Data Handling, 2002, pp. 299–311.

[22] V. N. Gudivada and V. V. Raghavan, “Design and Evaluation of Algorithms for Image Retrieval by Spatial Similarity,” ACM Transactions on Information Systems, vol. 13, no. 2, pp. 115–144, Apr. 1995.

[23] H. T. Bruns and M. Egenhofer, “Similarity of spatial scenes,” in Seventh International Symposium on Spatial Data Handling (SDH ’96), 1996, no. 4A, pp. 31–42.

[24] P. Matsakis, J. M. Keller, O. Sjahputera, and J. Marjamaa, “The Use of Force Histograms for Affine-Invariant Relative Position Description,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, no. 1, pp. 1–18, Jan. 2004.

[25] Buck, A. Keller, J., and Skubic, M., “A Memetic Algorithm for Matching Spatial Configurations with the Histograms of Forces”, IEEE Transactions on Evolutionary Computation, Vol. 17, No. 4, 2013, pp.588-604.

[26] Sledge, I., Keller, J., Skubic, M., “Mapping Natural Language to Imagery: Placing Objects Intelligently”, Proceedings, IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, August, 2009, pp. 518-524.

[27] J. Scheuerman, C. Michael, B. Landreneau, D. Acklin, and J. Harman, ‘Designing Interactive Machine Learning Systems for GIS Applications’, 11 2021, pp. 147–158.