Soft Information, Dirty Graphs and Uncertainty Representation/Processing for Situation Understanding

Geoff Gross and Rakesh Nagi
Center for Multisource Information Fusion
Department of Industrial and Systems Engineering
State University of New York at Buffalo
Buffalo, New York, U.S.A.
gagross@buffalo.edu and nagi@buffalo.edu

Kedar Sambhoos
CUBRC
Buffalo, New York, USA
sambhoos@cubrc.org

Abstract - In conventional warfare as well as counter-insurgency (COIN) operations, the understanding of the situation is extremely vital to assure a sense of security. Intelligence in COIN is about people, and deployed units in the field are the best sources of intelligence. Past and present intelligence data is analyzed to look for changes in the insurgents' approach or tactics. To do this, graphical methods have proven to be effective. In recent work, [10], have developed an inexact subgraph matching algorithm as a variation of the subgraph isomorphism approach for situation assessment. This paper enhances this procedure to represent inaccurate observations or data estimates, and inaccurate structural representations of a state of interest, thus accounting for the uncertainties. Various probabilistic and possibilistic uncertainty representations, transformations between representations and methods for establishing similarities between representations have been reviewed. This comprehensible approach will give pragmatic estimates providing rigor and sound understanding during situation assessment.

Keywords: Situation assessment, uncertainty representation, inexact graph matching

1 Introduction

Situation assessment has proven to be an essential part of counter-insurgency (COIN) operations. In many COIN operations the enemy is a highly dynamic, organized force. In order to maintain stability and security in COIN operations, an organized and efficient intelligence presence is required. Situational understanding involves combining intelligence data to form a coherent picture of the current state of the COIN environment. From this representation of intelligence information, certain potentially harmful scenarios or actionable intelligence can be identified and the appropriate actions can be taken moving forward. This intelligence information takes many forms and flows from many different sources (sensors, human observation, media outlets, etc.). Associated with each piece of information is an uncertainty about its accuracy. Based on the type of observation as well as the source of the observation, the method of representing this uncertainty may be different.

Probabilistic representations are historically well studied for uncertainty representation; however these representation methods are not well suited for representing vague terms, specifically in the realm of linguistic or soft observations. Also prevalent is the field of possibility theory and fuzzy logic, which has been proposed as a method to deal with the vagueness of certain information. The combination of both probabilistic and possibilistic uncertainty representations in a unified framework necessitates the formation of transformation methods between the two representations. Some methods for performing these transformations have been proposed in literature [1]-[8].

The representation of intelligence data is facilitated by the use of a graphical structure in the form of a data graph. This graphical representation is easily updated with additional intelligence data and forms a data type which is easily traversed for the purpose of situation assessment. The situation assessment operation is performed by a subject matter expert (SME) providing the graph matching algorithm a certain scenario of hypothesis (in the form of a template graph) to search for in the data graph. Through this graph matching operation, potentially harmful or destructive plans can be identified and the appropriate actions can be taken for prevention.

The remainder of this paper is formatted as follows: Section 2 provides background on situation assessment in the area of COIN operations, Section 3 discusses probabilistic and possibilistic uncertainty representations as well as transformation methods between the two uncertainty representations, Section 4 details the graph matching procedure, Section 5 describes similarity computations and Section 6 provides an example of how the entire situational assessment procedure functions.

2 Background

In a COIN operation the enemy has many strategic advantages including the ability to operate in secrecy, provide misinformation, dictate the time and place of clashes and the ability to wait to act on counterinsurgent forces missteps. Due to these advantages it is imperative that quality intelligence information be collected and properly analyzed from as many sources as possible. This intelligence information can come in the form of sensor data (e.g. a person's height or speed) or in the form of linguistic observation data from the people deployed in the field. Often the observations provided by people in the field are the most useful. This usefulness is due to a humans advanced reasoning and judgment abilities. While a height sensor might be able to solely assess a subjects height, a human observer might be able to provide a height estimate, weight estimate and specific facial features. The difficulty in utilizing human observation is in assessing the often vague and context dependent linguistic terms provided. Although an individual sensor may be limited to capturing one dimension of the current situation, sensors are often found in sensor arrays capable of measuring many dimensions of the current situation. Physics-based sensors have the additional advantage of being well calibrated, with well defined observation confidence intervals. Human observation is difficult to calibrate due to high variation between observers.

In the past many methods have been investigated to facilitate situation assessment in intelligence analysis. These methods include Bayesian networks and graph matching. Bayesian networks are the most popular representation method found in literature. A Bayesian network consists of a set of random variables and their associated conditional probabilities. These conditional probabilities can be used in conjunction with observations when traversing a Bayesian network to assess the likelihood of a given scenario.

One of the difficulties in implementation of Bayesian networks, particularly in intelligence situation assessment, is the requirement of *a priori* or machine learned conditional probability knowledge. Without this domain wide *a priori* knowledge, it becomes difficult to maintain uniformity between observations with and without conditional probability information. The limitless scope of observations and incredibly high cost of an inappropriate situation assessment makes machine learning techniques impractical. To avoid the shortfalls of Bayesian networks the more general approach of inexact graph matching has been proposed. A further discussion of the graph matching procedure can be seen in Section's 4 and 5.

3 Uncertainty representations

With intelligence information taking many forms, it is not unreasonable to expect the methods for representing the uncertainties related to these observations to be varying. There are two main types of uncertainty representations which may be present in intelligence data, well-calibrated or probabilistic uncertainty and vague or fuzzy data.

3.1 Probabilistic uncertainty representations

Sensors are generally well-calibrated instruments which have well-defined probabilistic confidence intervals. The ability to perform calibration prior to implementation in the field allows for an accurate probabilistic representation of the variance associated with a particular sensor. Although the concept of probability has been around for hundreds of years, the axiomatic foundations for probabilistic uncertainty representations were laid by Andrey Kolmogorov in the early 1900's. The axioms are as follows:

$$0 \leq P(E) \leq 1 \tag{1}$$

$$P(\Omega) = 1 \tag{2}$$

$$P(\emptyset) = 0 \tag{3}$$

Axiom (1) states that the probability of an event or observation, 'E,' is between 0 and 1 inclusive. Axiom (2) states the probability that an observation is within the universe of discourse, Ω, is equal to 1. This can be interpreted as stating that the universe of discourse is comprehensive. The final axiom, (3), states that the probability of the empty set, ∅, is equal to 0. Another central idea to probability theory is the idea of variance. The variance of a particular probability distribution gives a method for determining the probability that the actual observed object or parameter of interest is within a certain distance from the reported observation.

Probability distributions can be broken into three sub-categories, those that are defined on bounded intervals (a bounded universe of discourse), defined on semi-infinite intervals and those that are unbounded or have no limit on potential values. A bounded interval might be appropriate for some measures such as a person's height where a reasonable observation might be between 0 and 8 feet. Alternatively, an unbounded or semi-infinite interval might be reasonable for an observation where some large magnitude observation is possible with very low probability, or the range itself is unbounded. Some examples of potential probability distributions are as follows:

- Bounded Interval Observations – beta distribution, uniform distribution, triangular distribution, etc.
- Semi-Infinite Observations – exponential distribution, chi-square distribution, gamma distribution, etc.
- Unbounded Interval Observations – normal distribution, logistic distribution, etc.

Probability distributions can be represented in two ways, a probability density function (PDF, $f(x)$) or a cumulative distribution function (CDF, $F(x)$). The PDF is required to sum to 1 across its particular universe of discourse, with the probability of a particular range of interest defined as the integral across that range. The CDF is the cumulative probability to point 'x' in the range $(-\infty, x]$. From the CDF the probability of a particular range is defined as $F(upper\ value) - F(lower\ value)$. Another probability representation of interest is a point estimate, or discrete probability distribution. These probability values and transforms to possibility values are further examined in Section 3.3.

From an exhaustive list of potential probability distributions, data associated with truthed sensor observations can be used to determine the best fit of these distributions. These distributions can in turn be used in the graph matching procedure to determine potential data graph-template graph node/edge attribute similarity scores.

3.2 Possibilistic uncertainty representations

Possibility theory is different from probability theory in that it provides an indication of the level which a particular element of the universe of discourse is possible. Possibility values (or membership values) are typically higher than a related probability value and a particular possibility function will have a wider range of possible values than an associated probability distribution. This is due to the fact that the integration of a possibility function is not required to equal 1, as well as the less strict requirement of an element simply being a possible value for the defined element of interest.

Similar to probability theory, possibility theory has its own defining axioms:

$$pos(\emptyset) = 0 \quad (4)$$

$$pos(\Omega) = 1 \quad (5)$$

$$pos(U \cup V) = \max\big(pos(U), pos(V)\big) \quad (6)$$

$$for\ any\ disjoint\ sets\ U\ and\ V$$

Axioms (4) and (5) can be jointly interpreted as stating that the universe of discourse is comprehensive. Axiom (6) states that the possibility of the union of two disjoint sets is the maximum possibility of either one of those sets.

Possibilistic distributions or membership functions are similar to probability distributions in that they can be defined either on bounded intervals or semi-infinite/unbounded intervals. Some examples membership functions are: trapezoidal functions, triangular functions, Gaussian function, Sigmoid function, etc. An example for complete membership function for a person's height is as follows:

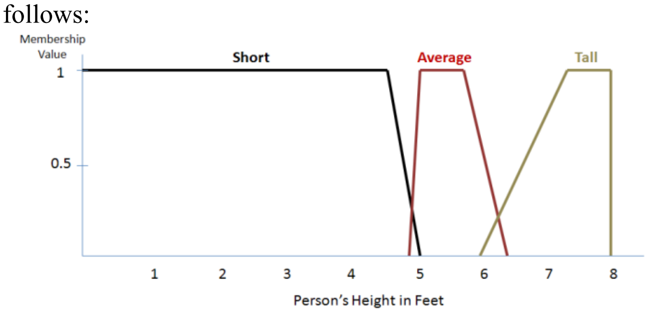

Figure 1. Membership function of linguistic terms describing a person's height

From this membership function it can be understood that any observation of a humans height under 4 feet would be classified as "Short" while an observation of 6 feet could possibly be classified as "Average" or "Tall". Similar to probability distributions, membership functions can be used to determine similarities between fuzzy data graph and template graph elements.

The previous example makes use of trapezoidal membership functions or fuzzy numbers. Trapezoidal membership functions are the most commonly used possibilistic representation due in large part to their flexibility and computational simplicity. A trapezoidal membership function has four defining points, the low base, low ceiling, high ceiling and high base. The fuzzy number for the "Average" height can be represented as $\tilde{A} = (a_1, a_2, a_3, a_4) = (4.8, 5, 5.8, 6.25)$.

Fuzzy numbers can be used in the representation of crisp/real valued observations ($\tilde{B} = (10, 10, 10, 10)$), crisp intervals ($\tilde{C} = (18, 18, 21, 21)$), as well as fuzzy intervals ($\tilde{D} = (30, 35, 35, 40), \tilde{E} = (50, 55, 65, 70)$). Figure 2 illustrates these individual functions.

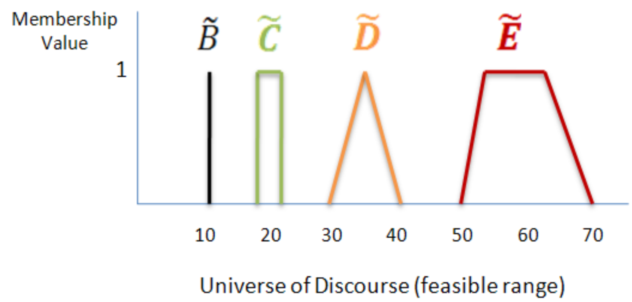

Figure 2. Individual examples of membership functions

The concept of a normalized fuzzy number is important for many fuzzy calculations. A normalized fuzzy number is one which exists in the range [0, 1] and is calculated by scaling a fuzzy number by the length of the universe of discourse. In the case of $\tilde{B}$, its normalized membership function would be: $\tilde{B}_{normalized} = \left(\frac{10}{70}, \frac{10}{70}, \frac{10}{70}, \frac{10}{70}\right)$ (assuming $\tilde{B}$'s universe of discourse is [0, 70]).

3.3 Uncertainty transformations

A review of the literature indicates that over a dozen transformation models exist for information exchange between possibilistic and probabilistic domains. Seminal papers over viewing possibility-probability transformation models are by [1]-[3]. In this section we identify four transformation models found in the literature that are most germane to the situation assessment application under investigation.

The ratio scale transformation is a straightforward transformation that is based on the notion of "normalizing" the data: see, for example, the defuzzification procedure by Yager [4]. A limitation of the ratio transformation is that it does not preserve information. This is a violation of the principle of uncertainty invariance. Information is gained when transforming from the probability domain to the possibility domain while information is lost when transforming from the possibility domain to the probability domain.

Klir's transformation [2], [5] is based on:

- Uncertainty (information) preservation between the probabilistic and possibilistic models
- Preservation of some scale
  - Ratio scale: $\pi(s_i) = \alpha \cdot p(s_i)$ (7)
  - Interval scale: $\pi(s_i) = \alpha \cdot p(s_i) + \beta$ (8)
  - Difference scale: $\pi(s_i) = p(s_i) + \beta$ (9)
  - Log-interval scale: $\pi(s_i) = \beta \cdot p(s_i)^{\alpha}$ (10)

The log-interval scale transformation possesses nice properties in terms of computation and uniqueness. We need to find the two parameters α and β such that $\pi(s_i)$ and $p(s_i)$ are assumed to be ordered in descending order.

The Dubois and Parade [6], [7] transformation method is based on the following two principles:

- Principle of preference intensity preservation
- Principle of maximum specificity

The principle of preference preservation is upheld if, prior to the transformation $p(x) > p(y)$, and after the transformation, $p(x^t) > p(y^t)$. The intensity of the preservation may be "weak" or "strong". The principle of maximum specificity is based on the specificity relation between possibility distributions. Consider possibility distributions π and $\pi^{'}$. If $\pi(x) \leq \pi^{'}(x)$ for all $x \in X$, then possibility distribution π is more specific than possibility distribution $\pi^{'}$. The transformations defined by Dubois and Prade are defined by the equations:

$$T_1: \pi(s_i) = \sum_{j=1}^{n} \min[p(s_i), p(s_j)] \quad (11)$$

$$T_1^{-1}: p(s_i) = \sum_{k=1}^{n} \frac{1}{k}\left(\pi(s_k) - \pi(s_{k+1})\right) \quad (12)$$

Asymmetric transformations are proposed by Dubois *et al.* [8]. The underlying premise for these transformations is based on the notion that probabilistic and possibilistic methodologies are not equivalent representations of uncertainty. Hence, there should be no reason to expect symmetry when transforming from probability to possibility and *vice versa*. The asymmetric transformations are based on the realization that a possibility representation is weaker than a probabilistic representation since possibility measures are structured based on ordering while the structure of a probabilistic measure is additive. Asymmetric transformations for $p \leftrightarrow \pi$ are defined by:

$$T_2: \pi(s_i) = \sum_{j \geq i} p(s_j) \quad (13)$$

$$T_2^{-1} = T_1^{-1}: p(s_i) = \sum_{k=1}^{n} \frac{1}{k}\left(\pi(s_k) - \pi(s_{k+1})\right) \quad (14)$$

The transformation from probability to possibility is based on the principle of maximum specificity. Under this principle the transformation aims at finding the most informative possibility distribution. The transformation from possibility to probability is based on the principle of insufficient reason. Here the aim is to find a probability measure that preserves the uncertainty of choice between outcomes.

4 Graph matching for situation assessment

In graph matching there are two graphs of interest, the data graph and template graph (Figure 3). The data graph contains the observational data where the nodes represent the observed entities and the arcs represent a relationship between a pair of nodes. The template graph is a pre-specified situation of interest or hypothesis, provided by a subject matter expert which has the same node-arc structure as the data graph. In both graphs the nodes and arcs can contain attributes which are the finest grain observation within the graph.

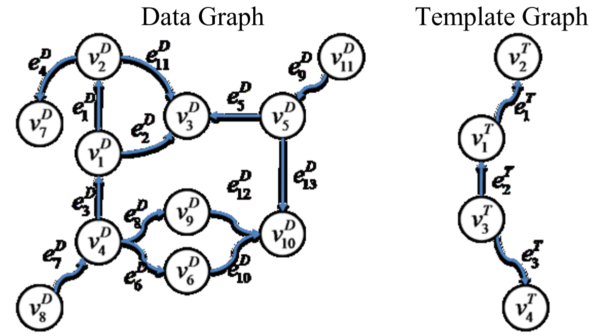

Figure 3. An example of data and template graphs

In an exact graph matching procedure the matching algorithm searches for an exact attribute-attribute match for each node-attribute or arc-attribute contained in the template graph. While this approach may be more computationally simple, it is usually impossible to find an exact match, especially when the situation of interest is a complex one.

A more useful procedure in a complex domain such as intelligence analysis is inexact graph matching. Although the results of an inexact match are more useful, the inexact graph matching problem has been shown to be NP-hard and therefore only heuristic solutions can be found for practical sized problems. A further complication is the requirement that situation assessment be done in real time, thus placing an even greater emphasis on computational efficiency. Inexact graph matching works by computing similarity scores between each node of the template and data graphs. These similarity scores can then be used in a matching algorithm which finds the best data graph to template graph match. This match is made while ensuring the topological relationships of the template graph are preserved.

In the Truncated Search Tree (TruST) algorithm [8], [9], node – node scores are calculated between each node of the template and data graphs as seen in Section 5. In addition, a neighborhood score is calculated for each node of the data graph to provide an indication of which node has the most promising topological match to the template graph. These scores can then be used in an iterative procedure to provide a set of possible matches, along with overall similarity scores (between the nodes and arcs of the template graph to matched data subgraph).

5 Similarity scoring

An important part in any graph matching procedure is establishing the degree of similarity between elements of the data and template graphs. In an attributed graph, these measures of similarity must be determined on an attribute-attribute level and then aggregated to form the overall data subgraph – template graph similarity score. The idea of a similarity measure is very much like the idea of a distance measure. Often, similarity measures are constructed by taking the compliment of the distance between two normalized observations. Many similarity measures have been suggested in literature which can be divided into two major categories, similarity measures which output a crisp similarity value [11 - 13] and those which output a fuzzy similarity value [14, 15].

Each similarity measure type can account for many different observation types. Crisp or real valued observations can be represented by considering them a trapezoidal membership function whose four defining points are all located at the crisp value. In a similar manner, crisp intervals such as "… weighed between 210 and 220 pounds" can be represented by assigning the low base and low ceiling values to 210 pounds and the high ceiling and high base values to 220 pounds.

5.1 Crisp similarity measures

Crisp similarity measures take two attributes in the form of normalized fuzzy numbers and output a point estimate for the similarity between the two. Many crisp similarity measures have been proposed in literature [11 - 13].

5.1.1 Hsieh and Chen similarity measure

Hsieh and Chen [11] suggested a computationally simple distance measure as the basis for a similarity score between two normalized fuzzy numbers. The measure is based on the "graded mean integration-representation" of distance. The distance between two fuzzy numbers, $\tilde{A}$ and $\tilde{B}$, is defined as:

$$d(\tilde{A},\tilde{B}) = |P(\tilde{A}) - P(\tilde{B})| \tag{15}$$

Where $P(\tilde{A})$ and $P(\tilde{B})$ are defined as:

$$P(\tilde{A}) = \frac{a_1+2a_2+2a_3+a_4}{6} \tag{16}$$

$$P(\tilde{B}) = \frac{b_1+2b_2+2b_3+b_4}{6} \tag{17}$$

Finally, the similarity measure, $S(\tilde{A},\tilde{B})$, is given as the compliment of the normalized distance:

$$S(\tilde{A},\tilde{B}) = \frac{1}{1+d(\tilde{A},\tilde{B})} \tag{18}$$

5.1.2 Chen similarity measure

Chen [12] proposed a similarity measure which is similar to the Hsieh and Chen measure; however it does not give a higher weight to the low and high ceiling points. The similarity value is defined by:

$$S(\tilde{A},\tilde{B}) = 1 - \frac{\sum_{i=1}^{4}|a_i-b_i|}{4} \tag{19}$$

5.1.3 Chen and Chen similarity measure

The Chen and Chen similarity measure [13] expands on the Chen measure by the addition of a center of gravity estimate of the two fuzzy numbers to be compared. More specifically it uses the simple center of gravity method (SCGM) based on [16]. From the simple center of gravity method the centroid, $(x_{\tilde{A}}^*, y_{\tilde{A}}^*)$, of a fuzzy number, $\tilde{A}$, is defined as:

$$y_{\tilde{A}}^* = \begin{cases} \frac{\left(\frac{a_3-a_2}{a_4-a_1}\right)+2}{6} & if\ a_1 \neq a_4 \\ \frac{1}{2} & otherwise \end{cases} \tag{20}$$

$$x_{\tilde{A}}^* = \frac{y_{\tilde{A}}^*(a_3+a_2)+(a_4+a_1)(1-y_{\tilde{A}}^*)}{2} \tag{21}$$

The similarity measure is composed of three [0, 1] terms which consider the effects of geometric distance and center of gravity distance. The similarity value is calculated by:

$$S(\tilde{A},\tilde{B}) = \left[1 - \frac{\sum_{i=1}^{4}|a_i - b_i|}{4}\right] \times (1 - |x_{\tilde{A}}^* - x_{\tilde{B}}^*|)^{B(\tilde{A},\tilde{B})} \times \frac{\min(y_{\tilde{A}}^*, y_{\tilde{B}}^*)}{\max(y_{\tilde{A}}^*, y_{\tilde{B}}^*)} \tag{22}$$

The $B(\tilde{A},\tilde{B})$ term is designed to indicate whether or not to use the center of gravity distance. In the case of comparing two real number or crisp observations, the center of gravity will not be considered. The appropriate $B(\tilde{A},\tilde{B})$ value is determined by:

$$B(\tilde{A},\tilde{B}) = \begin{cases} 1 & if\ S_{\tilde{A}} + S_{\tilde{B}} > 0 \\ 0 & if\ S_{\tilde{A}} + S_{\tilde{B}} = 0 \end{cases} \tag{23}$$

$$S_{\tilde{A}} = a_4 - a_1 \tag{24}$$

$$S_{\tilde{B}} = b_4 - b_1 \tag{25}$$

5.2 Fuzzy similarity measures

Fuzzy similarity measures take two normalized fuzzy numbers as input, and output a fuzzy number as the similarity between the two. It is intuitively reasonable to expect a fuzzy output when given a fuzzy input. As Voxman claimed, "if we are not certain about the numbers themselves how can we be ''certain'' about the distances among them" [14]. Based on this mindset, Voxman was the first to create a fuzzy similarity measure.

A central idea to fuzzy similarity measures is that of an α-cut. An α-cut is a method for describing which portion of the universe of discourse has a level of support greater than or equal to a particular α value. The α-cut of a fuzzy number is defined as:

$$\mu_\alpha(x) = (A^L(\alpha), A^R(\alpha)) = (a_1 + (a_2 - a_1)\alpha, a_4 - (a_4 - a_3)\alpha);\ \ \alpha \in [0,1] \tag{26}$$

5.2.1 Voxman similarity measure

Utilizing the above α-cut definition, the Voxman similarity measure [14] for two normalized fuzzy numbers is defined as:

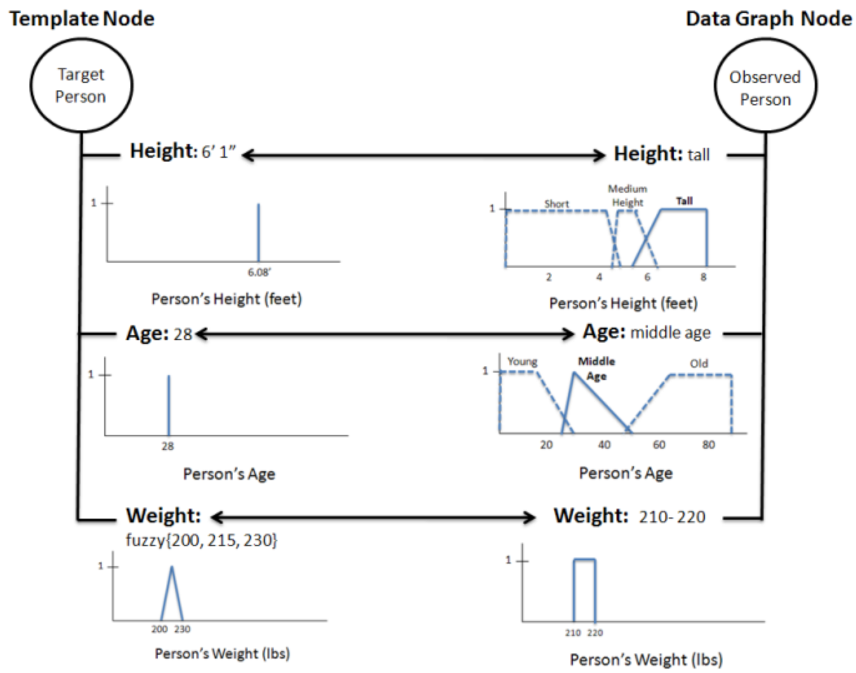

Figure 4. Example attributed data/template graph nodes with associated possibility representations

$$S_{Voxman} = 1 - d_{Voxman} \quad (27)$$
$$d_{Voxman} = (C(\alpha), D(\alpha)) \quad (28)$$

$$C(\alpha) = \begin{cases} \max\{A_2^L(\alpha) - A_1^R(\alpha), 0\} & if\ \frac{1}{2}(A_1^L(\alpha) + A_1^R(\alpha)) \leq \frac{1}{2}(A_2^L(\alpha) + A_2^R(\alpha)) \\ \max\{A_1^L(\alpha) - A_2^R(\alpha), 0\} & if\ \frac{1}{2}(A_2^L(\alpha) + A_2^R(\alpha)) < \frac{1}{2}(A_1^L(\alpha) + A_1^R(\alpha)) \end{cases} \quad (29)$$

$$D(\alpha) = \max\{A_1^R(\alpha) - A_2^L(\alpha), A_2^R(\alpha) - A_1^L(\alpha)\} \quad (30)$$

5.2.2 Guha and Chakraborty similarity measure

Like the Voxman similarity measure, Guha and Chakraborty [15] use α-cuts as the basis for their similarity measure. The similarity measure is defined as:

$$S(\tilde{A}, \tilde{B}) = (1 - d_{\alpha=1}^R - \sigma, 1 - d_{\alpha=1}^R, 1 - d_{\alpha=1}^L, 1 - d_{\alpha=1}^R + \theta) \quad (31)$$

$$d_\alpha^L = \eta[A_1^L(\alpha) - A_2^L(\alpha) + A_1^R(\alpha) - A_2^R(\alpha)] + [A_2^L(\alpha) - A_1^R(\alpha)] \quad (32)$$
$$d_\alpha^R = \eta[A_1^L(\alpha) - A_2^L(\alpha) + A_1^R(\alpha) - A_2^R(\alpha)] + [A_2^R(\alpha) - A_1^L(\alpha)] \quad (33)$$

$$\eta = \begin{cases} 1 & if\ \frac{A_1^L(1) + A_1^R(1)}{2} \geq \frac{A_2^L(1) + A_2^R(1)}{2} \\ 0 & if\ \frac{A_1^L(1) + A_1^R(1)}{2} < \frac{A_2^L(1) + A_2^R(1)}{2} \end{cases} \quad (34)$$

Guha and Chakraborty use the notion of ambiguity to compare their similarity measure with other fuzzy similarity measures proposed in literature. It is show that their similarity measure is less ambiguous than the Voxman measure, as well as other proposed similarity measures.

5.3 Example

To illustrate the graph matching process a simple node-node matching is presented here. In this example the target person has a height of 6'1", age of 28 and fuzzy weight with low base equal to 200 pounds, low ceiling equal to the high ceiling value of 215 pounds and high base value of 230 pounds. The example data graph node or observed attributes represent a tall, middle aged person with a weight between 210 and 220 pounds (see Figure 4)

Each of the observed attributes has an associated possibility function based on the type of observation (shown by solid lines). Each of the target persons attributes are also represented by a possibilistic function. The target person's weight is represented by a triangular fuzzy number to account for the natural sway of a person's weight.

The similarity scoring is shown using two methods, the Chen and Chen point estimate similarity scoring (see Table 1) and the Guha and Chakraborty fuzzy similarity scoring (see Figure 5).

Table 1. Chen and Chen attribute-attribute similarity scores

Attribute	Attribute-Attribute Score
Height	0.717
Age	0.550
Weight	0.654

Data Graph - Template Graph Score

0.640

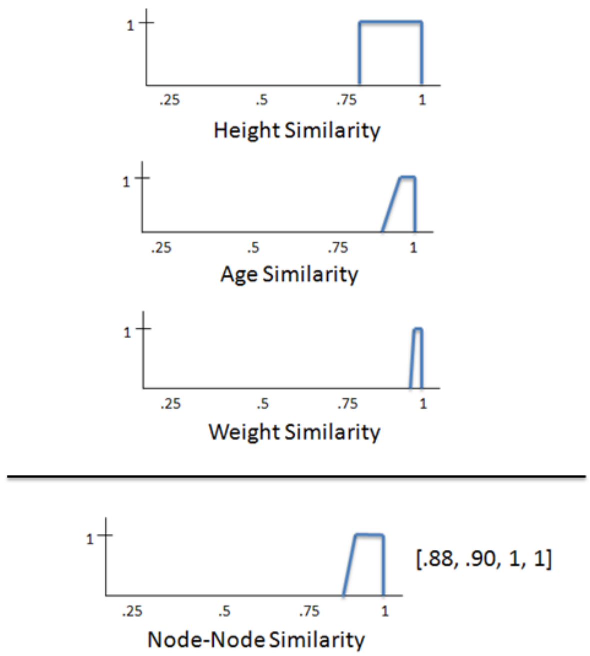

Figure 5. Guha and Chakraborty fuzzy attribute-attribute similarity scores

After calculating the similarity scores for each of the attributes in the observed node, the attributes are aggregated to form an overall node-node similarity score. In this example an equal weight is given to each of the Table 2.

The uncertainty information in the data is aligned into a single format (possibilistic representations) using the uncertainty transformation functions outlined in Section 3.3. For all the template-data graph node pairs an uncertainty based similarity measure is calculated based on the similarity scoring functions outlined in Section 5. nodes' attributes, to form the overall node-node similarity score. Other methods of aggregation have been recognized and will be examined in future work.

6 Dirty Graph Matching

The (clean) graph matching is described in Section 4. The dirty graph matching process flow diagram for TruST is shown in Figure 6 below.

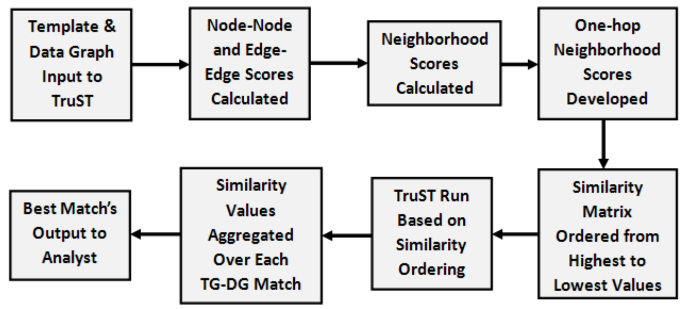

Figure 6. TruST dirty graph matching process flowchart.

The differences between crisp and fuzzy computations are illustrated in the TruST process flow shown in

The 1-hop neighborhood scores are calculated based on these node-to-node, edge-to-edge and neighborhood scores. These values are passed into the main TruST algorithm where the situations of interest (templates) are matched with observed data (data graph) to find the best match.

Table 2. Crisp and fuzzy computation process flow in TruST

	1.) Node-Node/ Edge-Edge Scoring	**2.) Neighborhood Scoring**	**3.) One-hop Neighborhood Scoring**
Crisp Computations	Crisp similarity measure for each attribute-attribute comparison (Chen, Chen)	Linear assignment problem solved to determine degree of topological promise	Crisp similarity value calculated based on weighting parameter for node-node score : neighborhood score
Fuzzy Computations	Fuzzy similarity measure for each attribute-attribute comparison (Guha, Chakraborty)	Fuzzy linear assignment problem solved to determine degree of topological promise in fuzzy sense	Fuzzy similarity value calculated based on weighting parameter for node-node score : neighborhood score

	4.) Similarity Ordering	**5.) TruST Execution**	**6.) Similarity Value Aggregation**	**7.) Best Match's Output**
Crisp Computations	Similarity matrix ordered in decreasing numeric order	Crisp value assigned as branching cost for each level of the search tree	Each TG-DG node match's score is multiplied to form overall score	Best matches presented to analyst with associated crisp score

Fuzzy Computations	Similarity matrix ordered based on fuzzy ranking method (Chen-Chen)	Fuzzy value assigned as the branching cost for each level of the search tree	Fuzzy average calculated	Best matches presented to analyst with associated fuzzy score

7 Conclusions

This paper presents a basic framework for uncertainty representation within an attributed data graph for the purpose of situation assessment. The uncertainty representations are not limited to probabilistic or possibilistic point estimates or distributions. This flexibility is accommodated by the use of transformations to convert uncertainty functions into a common format. The ability to account for uncertain information allows for a better representation of the real world and thus the ability for more accurate situational assessments. Future work in the expansion of these methods with a comparative study between the different similarity classes and investigation of the potential benefits of uncertainty representations in the data association process should also help advance the area of situation assessment with soft data.

Acknowledgement

This work has been supported by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for "Unified Research on Network-based Hard/Soft Information Fusion", issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery. We gratefully appreciate this support.

References

[1] M. Oussalah, *On the probability/possibility transformations: a comparative analysis*, International Journal of General Systems, 29(5):671–718, 2000.

[2] J.F. Geer and G.J. Klir, *A mathematical analysis of information-preserving transformations between probabilistic and possibilistic formulations of uncertainty*, International Journal of General Systems, 20:143–176, 1992.

[3] G. Klir and B. Parviz, *Probability-possibility transformations: A comparison*, International Journal of General Systems, 21(1):291–310, 1992.

[4] R. R. Yager, *Element selection from a fuzzy subset using the fuzzy integral,* IEEE Transactions on Systems, Man and Cybernetics, 23(2 Mar-Apr):467–477, 1993.

[5] G.J. Klir, *A principle of uncertainty and information invariance*, International Journal of General Systems, 17(2-3):249–275, 1990.

[6] D. Dubois and H. Prade, *On several representations of an uncertain body of evidence*, Fuzzy Information and Decision Processes Published by North-Holland Publ Co, Amsterdam, Neth and New York, NY, USA., pages 167–181, 1982.

[7] D. Dubois and H. Prade, *Fuzzy sets and statistical data*, European Journal of Operational Research, 25(3):345–356, 1986.

[8] D. Dubois, H. Prade, and S. Sandri. *On possibility/probability transformations*. Fuzzy logic: State of the art., pages 103–112, 1991. R. Lowen (Ed.), Kluwer Academic Publication.

[9] K. Sambhoos, R. Nagi, M. Sudit, and T. Rickard, *Hierarchical Higher Level Data Fusion using Fuzzy Hamming and Hypercube clustering*, ISIF's Journal of Advances in Information Fusion, JAIF, Volume 3, Number 2, December 2008.

[10] K. Sambhoos, R. Nagi, M. Sudit, A. Stotz, *Enhancements to high level data fusion using graph matching and state space search*, Information Fusion, In Press, Corrected Proof, Available online 16 December 2009, ISSN 1566-2535, DOI: 10.1016/j.inffus.2009.12.001.

[11] C.H. Hsieh, S.H. Chen, *Similarity of Generalized Fuzzy Numbers with Graded Mean Integration Representation*, in Proc. 8th Int. Fuzzy Systems Association World Congr., vol 2, (1999) 551-555

[12] S.M. Chen, *New Methods for Subjective Mental Workload Assessment and Fuzzy Risk Analysis*, Cybern. Syst. : Int. J., vol 27, no. 5, (1996) 449-472

[13] S.J. Chen, and S. M. Chen, *Fuzzy Risk Analysis Based on Similarity Measures of Generalized Fuzzy Numbers*. IEEE Transaction on Fuzzy Systems 11.1 (2003): 45-56.

[14] W. Voxman, *Some remarks on distances between fuzzy numbers*, Fuzzy Sets and Systems 100 (1998) 353–365.

[15] D. Guha, and D. Chakraborty, *A new approach to fuzzy distance measure and similarity measure between two generalized fuzzy numbers*, Applied Soft Computing (2010): 90-99. Print.

[16] P. Subasic and K. Hirota, *Similarity rules and gradual rules for analogical and interpolative reasoning with imprecise data,* Fuzzy Sets Syst., vol. 96, no. 1, pp. 53–75, 1998.