Reasoning about situations in the early post-disaster response environment

Galina L. Rogova
Encompass Consulting
9 Country Meadows Drive
Honeoye Falls, NY 14472
USA
rogova@rr.rochester.com

Peter D. Scott, Carlos Lollett,
Rashmi Mudiyanur,
Dept. Computer Science and Engineering
University at Buffalo
Buffalo NY 14260-2000
USA
peter@buffalo.edu

Abstract - The purpose of situation and impact assessment is to infer and approximate the critical characteristics of the environment in relation to the particular goals, capabilities and policies of the decision makers. The process of situation and impact assessment involves dynamic generation of hypotheses about the states of the environment and evaluation of their plausibility via reasoning about situational items, their aggregates at different levels of granularity, relationships between them, and their behavior within a specific context. This paper addresses the problem of reasoning for situation and impact assessment to support early-phase crisis management. Special attention is paid to “inference for best explanation” aimed at discovery of the underlying causes of observed situational items, their behavior, an important component of situation and impact assessment. The presented method of discovery of underlying causes is illustrated by the discovery of an unreported HAZMAT incident within an early-phase earthquake response scenario.

Keywords: situation and impact assessment, abduction, belief based argumentation system, Dempster rule of combination, post-disaster management, earthquake.

1 Introduction

This paper reports recent progress on research focused on reusable and transferable situation and impact assessment (SIA) techniques to support crisis management [1,2,3]. The core purpose of SIA is to infer and approximate the characteristics and critical events of the environment in relation to specific goals, capabilities and policies of the decision makers. The SIA processes utilize fused data about objects of interest (e.g. reports on casualties, damage of essential facilities, status of available resources and infrastructure in the post-disaster environment) dynamic databases, and expert experience, knowledge, and opinion for context processing. The goal of SIA is a coherent composite picture of the current situation along with prediction of consequences. This situational picture provides decision makers with essential information to help them to understand and respond to the situation, supporting timely and effective decisions to mitigate its impact.

Disaster response invariably engages several distinct organizations each with different tasking, competencies, technologies and scope of operation. In this case of multiple decision makers, SIA has to deliver consistent current and predicted situational pictures, which are relevant to each decision maker’s goals and functions. The process of building a situational picture comprises dynamic generation of hypotheses about the states of the environment and assessment of their plausibility via reasoning about situational items, their aggregates at different levels of granularity, relationships between them, and their behavior within a specific context. In some cases, assessment of plausibility of more complex hypotheses may require hierarchical processing, which includes not only reasoning about situational items and relationships between them but also includes relationships between hypotheses and assessments of plausibility of lower level hypotheses [4]. The SIA process is complicated by uncertainty characterizing observations, relationships, and behavior and variable reliability of the data and information sources.

An important component of situation assessment is causal inference aimed at discovery of underlying causes of observed situational items, their attributes and their behavior. Discovery of underlying causes of observed situations is the goal of abductive reasoning [4,5] or “inference for best explanations”. For example, in the early post-earthquake response phase, reasoning about situations is contingent on the assumption that most reported casualties and structural damage are the results of the primary earthquake shock incident and reported subsequent secondary incidents such as fire, flood, aftershocks and HAZMAT events. However some secondary incidents such as toxic spills may not be known for a long period of time. At the same time rapid discovery of such incidents is very important since they may have devastating consequences if not responded to quickly. These unknown secondary incidents are usually manifested by unexpected properties and behavior of situational items inconsistent with the current set of beliefs about the state of the world and therefore belief update may be required. Many belief update methods give priority to this new information and its consequences and abandon some old beliefs to preserve consequences. In the post disaster environment observations and knowledge about situational items, their behavior and relationships
are uncertain and, therefore it is necessary to account for this uncertainty while updating the current set of beliefs. In the uncertain dynamic environment belief update can be carried out by first seeking some explanations or underlying causes of these inconsistent observations and incorporating these explanations, if found, into a new set of beliefs. Possible explanations can be found as the result of abduction comprising generation of hypotheses about the underlying causes of these inconsistent observations and reasoning about plausibility of such hypotheses.

In this paper the hypothesis generation is assumed to be performed by human domain experts. For each candidate hypothesis, computation of its plausibility (belief) is based on the framework of a probabilistic argumentation system [6], the result of a combination of beliefs in supporting and refuting arguments by the unnormalized Dempster rule of combination [7,8].

The remainder of this paper comprises a description of the SIA methodology in more detail with special attention paid to the process of assessing plausible explanations of undiscovered events based on reasoning about recently received information and the behavior and relationships of situational items. The paper is organized as follows. Section 2 presents an approach to situation and impact assessment and its components. Section 3 describes the discovery of an unreported HAZMAT incident within an earthquake early-response scenario, an example illustrating the procedures described in Section 2. In Section 4 some conclusions are presented.

## 2 Higher level fusion processes

Information fusion concerns reasoning, at multiple levels of abstraction, about entities and their relationships within a given spatio-temporal domain. Lower level fusion processes focus on the detection and characterization of individual entities through association of reports with entities and determination of entity attributes through fusion of the contents of the associated reports. Higher level fusion processes focus on the relevant relationships among entities populating the domain. These higher level fusion processes include:

- Contextual understanding of the characteristics of the state of the environment critical to decision makers.
- Causal reasoning leading to recognition of possible causes or explanations of the identified events and estimated states of the world.
- Prediction of the future impact of current situations.

Figure 1 presents information flow in the SIA process.

### 2.1 Situations at different levels of granularity

SIA processes are invoked to derive essential information for multiple decision makers. The hierarchical structure of emergency management and considerations of hierarchical hypotheses call for reasoning about situations at various levels of granularity. For instance, information about the location of a casualty cluster and the expected level of injury severity of this cluster would be of interest to an ambulance dispatcher, while relationships between the capacity of the hospitals accessible from the cluster with a high expected level of injury severity have value when reported to the overall incident commander. Situation building blocks are described by inter and intra-class relationships between physical items such as casualties, buildings, and ambulances, or events such as discovery of casualties of a certain type of injury at a certain time or the presence of residual capacity of an ambulance and a hospital at specific locations within a certain time interval. These basic relational conjunctions of entities are defined as aggregates and are obtained by applying a similarity metric in the feature space.

The specific features used for aggregation depend on the information needs of a certain user or a group of users. Context dependent relationships between aggregates at a specified level of granularity define derived situations at the next coarser level of granularity. Derived intra-class situations created by the composition of basic intra-class situations at specific levels of granularity are called elementary situations. Examples of elementary situations in the disaster response context are Communication System Situation (e.g., capacity vs. demand; location, boundary); Transportation System Situation (e.g., link delay maps, impassable areas) HAZMAT Situation (e.g., secondary hazardous material release incidents; location, type, and dynamics of the possible secondary incidents), and Casualty Situation (boundary, severity, injury types, dynamics and causes). Relationships between elementary situations within a selected spatio-temporal setting and overall context comprise a composite situational picture.

Temporal reasoning about behavior of situational objects (aggregates at different levels of granularity) requires an association process, which correlates the situational objects identified at a certain time or time...
interval with situational objects identified at a different time or within a different time interval. This association process corresponds to reasoning about the identity of aggregates. Aggregates in the early post-disaster environment may be expected to constantly lose or gain members and change their attributes and characteristics but retain their identity. For example, clusters of casualties may lose and gain members due to a certain percentage of casualties being picked up by ambulances, a certain percentage being transported by private vehicles and others moving off on foot, while at the same time new members are being reported. Simultaneously, the location, growth rate or area of the casualty cluster can change. The reasoning about cluster identity is complicated by the fact that the information on the identity of members of aggregates is not known with certainty, and their characteristics and therefore cluster characteristics are also uncertain.

Following [9], in which the author was concerned about relationships between vague spatial regions, we conduct temporal cluster association (temporal reasoning about cluster identity) by reasoning about such topological relationships as disjoint, touch, overlap (strong and weak), covers, covered by, contains and contained by. Unlike the authors of [9], we define these relationships not in the physical space but in the cluster characteristic space, which includes such features as area and distribution of cluster members. As in [9] we call two aggregates identical if they are disjoint, touch, or weakly overlap, and distinct otherwise (see Figure 2).

![Spatial relationships between aggregates](from [9])

The temporal association thus requires a criterion for distinction between the weak and strong overlap, which is defined in [9] by the ratio of the area of the regions intersection to the area of the smallest region. We use the following criterion to classify overlap as weak or strong. Let $Ci_t = (Ci_{1,t},...,Ci_{m,t})$ be a set of clusters identified at time $t$ and $Ci_{t-1} = (Ci_{1,t-1},...,Ci_{m,t-1})$ be a set of clusters identified at time $t-1$. $\forall Cl_t$ a set of clusters $O(Cl_t) = (Cl_{i}^{m} | Cl_{i}^{m} \cap Cl_t \neq \emptyset)$ contains all clusters at time $t-1$, which overlap with cluster $Cl_t$.

Cluster $Cl_{1,t-1}$ discovered at time $t-1$ and cluster $Cl_t$ discovered at time $t$ strongly overlap if $Cl_{1,t-1} \cap Cl_t \neq \emptyset$ and $F(Cl_{1}, Cl_{1,t-1}) > \alpha$. $\alpha (0.5 < \alpha < 1)$ is a selected threshold and

$$F = \min\left(\frac{|Cl_{i}^{m} \cap Cl_t|}{\min(|Cl_{i}^{m} |, |Cl_t|)}, \frac{P(Cl_{min})}{\min(P(Cl_{i}^{m}),P(Cl_t))}\right),$$

where $P(Cl)$ is the expected number of members in cluster $Cl$ given all fused data reported for that cluster (cluster population), $Cl_{min} = \arg\min(P(Cl_{i}^{m}),P(Cl_t)))$, $P(Cl_{min})$ is the expected number of members of cluster $Cl_{min}$ in $Cl_{1,t-1} \cap Cl_{1}$, $|Cl|$ the volume of cluster $Cl$ in the feature space. If clustering is conducted in 2-dimensional space $(x,y)$ $|Cl|$ is the area of cluster $Cl$.

If $Cl_{1,t-1} \cap Cl_t = \emptyset$, $\forall Cl_{1,t-1}, Cl_t$ is a new cluster.

If $Cl_{1,t-1} \cap Cl_t = \emptyset \forall Cl_t$, cluster $Cl_{1,t-1}$ is said to have terminated by time $t$.

Cluster identity, and the behavior of its characteristics along with spatial relationships between clusters and their behavior, are used for casualty and damage assessment, resource allocation, discovery of possible underlying causes for assessed behavior, and impact prediction.

2.2. Reasoning about situations and causes

Let $\Theta$ be a set of possible states of the environment, $\Theta^i$ be a subset of possible states of the environment relevant to decision maker $k$, and $Pl$ be a plausibility structure on $\Theta$. At each time $t$, a situational picture relevant to decision maker $k$ can be described as a set of the plausible states of the environment:

$$S^k(t) = \{\theta^k(t) \in \Theta | Pl(\theta^k(t)) > 0\}.$$  

It is assumed here that decision makers do not have complete knowledge about all relevant states of the environment, that is, we adopt the open world assumption [8].

Dynamic reasoning about plausible states of the environment is performed based on characteristics and behavior of situational items at different level of granularity. These characteristics and behaviors are constantly updated by newly processed observations. The current set of plausible states is constantly revised and new hypotheses about the plausible state of the environment capable of explaining new characteristics and changes in the behavior of situational items are regularly generated and evaluated.

A reasoning framework appropriate for SIA in the post-disaster environment may be constructed based on the Probabilistic Argumentation System (PAS) (see, e.g. [6]) augmented with the set of relevant domain specific models such as hospital models and dynamic dispatch/routing models. Following [6] PAS can be described as an approach to non-monotonic reasoning under uncertainty, combining symbolic logic with probability theory for judging hypotheses about the unknown or future world by utilizing given knowledge. Logic is used to find arguments in favor of and against a
hypothesis about possible causes or consequences of the current state. An argument is a defeasible proof built on uncertain assumptions, that is, a chain of deductions based on assumptions that make the hypothesis true, or false. Every assumption is linked to an a priori probability that the assumption is true. The probabilities can be understood in the traditional Kolmogorov-axiomatized way but also can represent subjective probabilities. The probabilities that the arguments are valid are used to compute the credibility of the hypothesis, which can then be measured by the total probability that it is supported by the totality of supportive and refuting arguments. The resulting degree of support corresponds to belief of the theory of evidence and is used to make a decision whether a hypothesis should be accepted, rejected, or whether the available knowledge is insufficient to form a satisfactory judgment at this time.

In the post-disaster environment accurate a priori probabilities that the assumptions are true are rarely available and expert subjective beliefs have to be used. Moreover, due to the high uncertainty characterizing the post-disaster environment $P(A)$, expert subjective belief that assumption $A$ is true, is not in general, equal to $1 - P(\neg A)$ and therefore PAS has to be generalized to utilize sub-additive subjective belief measures: $\text{Bel}(A) + \text{Bel}(\neg A) \leq 1$. These subjective beliefs can be expressed in linguistic form, e.g., very high, high, low, very low with subsequent quantization of these linguistic values. The belief measures can be also represented numerically and be approximated by a function of the values assigned to attributes and relationships characterizing the state of environment and related to the assumptions. In some cases these belief measures can be the result of a combination of beliefs based on different characteristics with the Dempster rule. Beliefs in assumptions are combined to obtain beliefs in arguments, which favor and refute the hypotheses. These beliefs in turn are used to gauge the overall credibility of the hypothesis, measured by the total belief that is supported by arguments.

As was mentioned in the introduction, an important component of situation assessment is causal inference aimed at discovery of the underlying causes of observed situational items, their attributes and their behavior (abductive inference). This abductive process of reasoning from effect to cause requires [4,10]:

- constructing or postulating possible hypotheses and the states of the world explaining observations
- computing plausibility of these hypotheses
- selecting the most plausible hypothesis from among these.

Automatic hypothesis generation is the most difficult process to implement within SIA and is not assumed here, so that hypotheses are assumed to be generated by human experts and users. The process of hypothesis evaluation has to take into account the following considerations [10]: to what degree is the hypothesis to be selected better than alternatives?

- How credible is the hypothesis by itself, independently of considering the alternatives, i.e. one should be cautious about accepting a hypothesis even if it is clearly the best one we have if it is not sufficiently plausible in itself.
- Reliability of incoming data, which requires explanations.

Abductive inference starts with discovery of characteristics inconsistent with the current state of knowledge and behavior of attributes and relationships between the associated situational items. In the present model this anomalous behavior, or data inconsistency is detected by significant deviation in the behavior of attributes and relationships of situational items from expected, given the current state of knowledge. Examples of such behavior may include discovery of a new aggregate or situation, a specific pattern of spatial and/or temporal relationships between aggregates, or a significant deviation of the behavior of one or several characteristics of an aggregate or a situation from the expected average behavior of the characteristics of similar aggregates or situations. Classes of similarity of aggregates can be defined by clustering of the environmental features related to aggregate formation. For example, the expected number of casualties in a cluster depends on the severity and type of damage in the area, time of the day and the rate of casualty discovery, which in turn depends on the possible number of civilians, police, and ambulances reporting the casualties (density of roads, proximity to the hospital or density of population in the area).

Discovery of a deviation from the expected is followed by construction of a set of hypotheses (possible causes of the discovered deviation) about the situation. Then beliefs in each of these hypotheses are evaluated by identifying and combining with the Dempster rule of combination pro and contra arguments for them. Resulted beliefs are used to decide whether there is enough information to select one of the hypotheses and which hypothesis to be selected.

Let $\Theta' = \{\theta'_1, \ldots, \theta'_k\}$ be a set of hypotheses under consideration at time $t$ and $\text{Bel}'(\theta_i)$ be beliefs assigned to these hypotheses. Given the open world assumption, this hypothesis set is not exhaustive and $\text{Bel}'(\emptyset) \neq 0$. Emergency management in post-disaster environment acts under severe time and resource constraints and a high cost associated with certain errors. Selection of a certain hypothesis may correspond to a specific state of the environment, which may require specific, often urgent, response. Waiting may result in unacceptable decision latency leading to either wasted resources or lost lives. At the same time the cost of false alarm can be very high since valuable resources might be diverted from the location where they are critically needed. Therefore cost of waiting for additional incoming information to obtain data of better quality and reduce false alarm, has to be justified by the benefits of achieving results of higher certainty. These considerations lead to the following decision criteria:

- If $\text{Bel}'(\emptyset) \geq \text{max}\{\text{Bel}'(A)\}, \forall A \subseteq \Theta$ than an expert is engaged to revaluate a set of hypotheses considered and/or a sensor management process is initiated, for example an expert observer can be dispatched to verify the incoming information.
is a hypothesis that the excessive dispersion of the material is modeled by a Gaussian plume driven by the wind field, resulting in primarily respiratory injuries (7% in our case). These relationships can describe the form of the threshold $\theta(t)$ is context specific. It is considered within the class of decreasing convex functions and is equal to zero when it achieves a certain maximum value (a deadline). In certain situations, when decisions based on the resulting decision state estimations have very serious consequences, a sensor management process can be rapidly employed.

The next section illustrates these ideas in the context of the DIsaster Relief Environment. The process of discovery of an unreported HAZMAT incident during early response to an earthquake is considered.

3 Secondary event discovery

3.1 The DIRE simulation test bed

The DIsaster Relief Environment, DIRE [1-3], is a synthetic task environment designed to simulate key features of the conditions during, and responding agencies’ actions taken in, a range of earthquakes inspired by the January 17, 1994 Northridge, California earthquake. Systematic testing of the ideas presented here using DIRE is underway, and a brief description follows.

Classes of entities modeled in the DIRE environment include agents, mobile objects, stationary objects and secondary events. Agents include police, ambulance personnel, civilian reporters, hospital engineers, call center personnel and casualties. Casualties are seeded in the environment by a Monte Carlo laydown of the statistics reported by HAZUS, a consequences prediction package developed by FEMA. Agents generate reports on casualties, which include attributes such as type of injury (e.g. trauma, respiratory, cardiovascular). These attributes are all subject to random reporting error, with error probabilities varying by reporter type. Reports are subject to a randomized delay before reception at a call center. The report stream is associated and fused. Fused reports are utilized to dispatch and route available ambulances, pick up casualties and deliver them to the hospitals with available resources.

Secondary events include aftershocks, spills of hazardous materials, fires, and delayed structural failures. DIRE version 2.0 is configured to model a HAZMAT incident in which a colorless, odorless toxic gas is vented to the atmosphere as the result of the accidental or malicious rupture of a chemical storage tank. Dispersion of the material is modeled by a Gaussian plume driven by the wind field, resulting in primarily respiratory casualties. An excess of respiratory casualties in a given cluster, and its growth with the prevailing wind, is supportive of a hypothesis of a secondary HAZMAT incident not yet discovered. Discovery of such an incident, as described in the next section, permits impact prediction and may drive targeted evacuations as well as additional constraints to the ambulance and police movement.

3.2 Reasoning processes

The example presented in this section illustrates the process of discovery of a possible unreported event, namely an occult toxic spill. At fixed time intervals, shrink clustering [3] is used to identify the current set of casualty clusters. Each cluster consists of a connected set of cells, which are used for the hierarchical cluster routine. Discovery of an unreported toxic spill is invoked by detection of unusually high percentage of respiratory injuries within certain casualty clusters at time $t$, and corroborated by reports of new respiratory casualties in spatial progression downwind of the discovered but unexplained respiratory cluster. In the absence of uncertainty, this new information calls for update of the current belief that an expected percentage of respiratory injuries due to building damage are not higher than an a priori known value. In our case we select this value to be 10%, the number characterizing the fraction of non-hazmat-related respiratory injuries reported during the 1994 Northridge earthquake. Due to uncertainty of observations and the current knowledge base it is advantageous to look for a possible underlying reason for this unusually high level of respiratory injuries before updating the current beliefs. We do not ignore the possibility that this high level of respiratory injury is due to building damage as a result of the earthquake.

We consider a two hypothesis frame of discernment $\Theta = \{\theta_1, \theta_2\}$, where $\theta_1$ is a hypothesis that a toxic spill occurred and $\theta_2$ is a hypothesis that the excessive respiratory injuries are the result of structural damage only. We also assume that there might be an unknown cause (open world assumption) and that the plausibility that there is more than one toxic spill within a certain time interval is negligible.

The arguments used to compute beliefs supporting or rejecting a hypothesis represent a conjunction of propositions and uncertain assumptions about characteristics and behavior of “suspicious subclusters” and spatio-temporal relationships between such subclusters as well as between such subclusters and other clusters. Suspicious subclusters at time $t$ are subclusters comprising connected cells with the expected number of respiratory injuries in each cell above the threshold defined by the expected value and the deviation of respiratory injuries (7% in our case).

A set of suspicious subclusters at time $t$, $SC_t = \{SCI_t\}'$, is represented as a union of 3 subsets: $SC_t = P^1_t \cup P^2_t \cup P^3_t$, where $P^1_t$ is a set of subclusters formed at time $t$, $P^2_t$ is a set of subclusters formed before time $t$ but not suspicious at time $t - 1$, and $P^3_t$ is a set of suspicious subclusters, which were suspicious at $t - 1$. Below are definitions of the relationships “between”, “close”, and “neighbors” used in the reasoning processes. These relationships can describe intra relationships...
between subclusters, between clusters as well as \textit{inter} relationships between clusters and subclusters of other clusters [3]:

1. Clusters $C_{i}^l$ and $C_{j}^l$ are considered \textit{neighbors} at time \( t \) if the line connecting their centroids does not intersect any other clusters. Relationship \textit{neighbors} is reflexive, symmetric but not transitive. \( N_{i}^l \) denotes a set of neighbors of cluster \( i \ (C_{i}^l) \) at time \( t \).

2. Clusters $C_{i}^l$ and $C_{j}^l$ are called "\textit{close}" if the distance between the centroids of these clusters is less then a threshold: $\text{Dist}_{ij} = \max(W \cdot \Delta t, a \cdot \max(D_{i},D_{j}))$, where \( W \) is the wind speed, $\Delta t$ is the time step considered, \( D_{i},D_{j} \) are maximum diameters of $C_{i}^l$ and $C_{j}^l$, respectively, and \( a \) is a constant.

3. Cluster $C_{i}^l$ is said to be \textit{between} clusters $C_{k}^l$ and $C_{l}^l$ if $C_{i}^l \in N_{k}^l \cap N_{l}^l$ and $C_{i}^l$ is within the box around both $C_{k}^l$ and $C_{l}^l$, and clusters $C_{k}^l$ and $C_{l}^l$ are close.

Specific propositions considered for Hazmat spill discovery include propositions characterizing wind topology and topology temporal behavior (e.g., new disappearing clusters and subclusters):

\begin{itemize}
    \item \textit{P1:} wind direction
    \item \textit{P2(SCI$_{i}^l$):} SCI$_{i}^l$ \( \in P_{i}^l \)
    \item \textit{P3(SCI$_{j}^l$):} SCI$_{j}^l$ \( \in P_{j}^l \)
    \item \textit{P4(SCI$_{i}^l$):} SCI$_{i}^l$ \( \in P_{i}^l \)
    \item \textit{P5(C$_{i}^l$,SCI$_{i}^l$):} C$_{i}^l$ \( \in N_{i}^n \) (cluster \( n \) is a neighbor of suspicious subcluster \( m \) at time \( t \))
    \item \textit{P6(Cl$_{i}^l$):} Cl$_{i}^l$ \( \not\in SC_{i} \) (cluster \( j \) is discovered at time \( t \))
\end{itemize}

It is necessary to note that in the uncertain environment cluster topology and topology behavior declarations are uncertain and represent assumptions. In our pilot study we assume that their truth is known with certainty and consider them propositions.

Assumptions about suspicious subcluster characteristics and their behavior include:

\begin{itemize}
    \item \textit{A1:} The expected fraction of respiratory injuries in a subcluster indicate HAZMAT (how "suspicious" is this subcluster?)
    \item \textit{A2:} The expected fraction of respiratory injuries in a subcluster is increasing.
    \item \textit{A3:} Subcluster area is growing.
    \item \textit{A4:} Subcluster center is moving downwind.
    \item \textit{A5:} Subcluster center is moving upwind.
    \item \textit{A6:} Subcluster centroid is moving downwind.
    \item \textit{A7:} Subcluster centroid is moving upwind.
    \item \textit{A8:} Cluster $C_{i}^l$ (subcluster SCI$_{i}^l$) is located downwind from subcluster SCI$_{j}^l$.
    \item \textit{A9:} SCI$_{i}^l$ is \textit{between} clusters $C_{i}^l$ and $C_{l}^l$.
    \item \textit{A10:} Cl$_{i}^l$ is \textit{between} subclusters SCI$_{i}^l$ and SCI$_{j}^l$.
\end{itemize}

Each assumption is assigned a belief measure, which represents expert belief that this assumption is true. In our example these belief measures are modeled as functions of the behavior of values of suspicious cluster characteristics and mereotopological \textit{intra} relationships between neighboring subclusters and \textit{inter} relationships between subclusters and neighboring clusters. Let \( \Omega = \{\alpha_{l}^{(i)}, \alpha_{l}^{(j)}\} \), where $\alpha_{l}^{(i)}$ is a hypothesis that assumption \( i \) is true and $\alpha_{l}^{(j)}$ is a hypothesis that assumption \( j \) is not true. Then for each assumption we model the measures of belief as follows.

For assumptions A1 – A3 :

\[
\text{bpa}(\alpha_{l}^{(i)}) = \frac{\lambda_{i}}{1 + \alpha_{i} e^{-\beta_{i}X_{l}}}, \quad \text{bpa}(\alpha_{l}^{(j)}) = 0, \tag{1}
\]

where $\alpha_{i}, \beta_{i}, \lambda_{i}$ are parameters, $l=1,2,3,9,10$.

For $l=1$ \( X_{l}^i \) is the fraction of respiratory injuries and belief is based on the "level of suspiciousness".

For $l=2-3$ \( X_{i} \) is the relative difference between the change of the subcluster characteristics under consideration \( (\Delta Y_{i}^l) \) at time \( t \) and the absolute value of an average change of magnitude of these characteristics \( (avg_{i-l}) \) up to and including time \( t-1 \):

\[
X_{i} = \left| \frac{\Delta Y_{i}^l - avg_{i-l}^{j-l}}{avg_{i-l}} \right|, \tag{2}
\]

where \( Y_{i}^l \) is the fraction of respiratory injuries, if \( l=2 \) and the suspicious subcluster area if \( l=3 \). For \( l=9 \) \( X_{l}^i \) is the distance between the centroid of SCI$_{i}^l$ and the line connecting centroids of $C_{i}^l$ and $C_{l}^l$.

For \( l=10 \) \( X_{l}^i \) is the distance between the centroid of $C_{l}^l$ and the line connecting centroids of SCI$_{l}^l$ and SCI$_{j}^l$.

For assumptions A4 – A8 :

\[
\text{bpa}(\alpha_{l}^{(i)}) = \frac{1 + (1-Y_{i}^l)^2 \cos(\phi_{l})}{2} \chi_{l}, \quad \text{bpa}(\alpha_{l}^{(j)}) = 0, \tag{3}
\]

where $\phi_{l}$ is the angle between the wind direction and the direction of movement of the geometrical center \( (l=4,5) \), or the center of gravity \( (l=6,7) \), or the vector from the center of gravity of SCI$_{l}^l$ to the center of gravity of $C_{l}^l$ \( (l=8) \) and \( \chi_{l} \) is a scaling parameter.

Finally, arguments built from these propositions and assumptions corroborating and refuting the toxic spill hypothesis are composed. Sets of arguments differ slightly for \( P_{i}^l, P_{j}^l \), and \( P_{k}^l \) because of the temporal difference in behavior of their characteristics. Below are assumptions considered for subclusters SCI$_{i}^l \in P_{i}^l$ \( (P_{2}) \).

Corroborative arguments \( (\text{Arg}P_{2}) \) include:

\textit{ArgP}_{2}: The expected fraction of respiratory injuries in subcluster SCI$_{i}^l \in P_{i}^l$ indicate a toxic spill, this cluster is a \textit{neighbor} of subcluster SCI$_{j}^l \in P_{j}^l$ and is located downwind from SCI$_{j}^l \in P_{j}^l$.
A suspicious subcluster area is growing downwind: $A3(SCl_i) \land AA(SCl_i) \land P2(SCl_i)$

$ArgP_2$: Respiratory injuries are growing downwind: $A2(SCl_i) \land A6(SCl_i) \land P2(SCl_i)$

$ArgP_1$: A suspicious subcluster area is growing upwind: $A3(SCl_i) \land AA(SCl_i) \land P2(SCl_i)$

$ArgC_2$: Respiratory injuries are growing upwind: $A2(SCl_i) \land A7(SCl_i) \land P2(SCl_i)$

$ArgC_3$: A suspicious subcluster $SCl_i \in P_r^2$ is between clusters $C_l^s$ and $C_l^p$, which are located along the wind direction and do not contain suspicious subclusters: $Pr2(SCl_i) \land A9(SCl_i, C_l^s, C_l^p) \land A8(SCl_i, C_l^s)$

Beliefs in support of each assumption $i$ invoke simple support functions on a frame of discernment $\Omega = \{\omega_i, \omega_j\}$, with a single focal element $\omega_i$ (assumption $i$ is true). A direct sum of the simple support functions over a set $\{\Omega^i \mid \forall i = ArgP_r, \forall k\} \text{ is mapped then into a simple support function } \mu_k \text{ with focus } \theta (\text{pro HAZMAT}):$

$$\mu_k(\theta) = \prod_{i = ArgP_r} bpa(\omega_i), \quad \mu_k(\Theta) = 1 - \mu_k(\theta). \quad (4)$$

Similarly, a direct sum of the simple support functions over the set $\{\Omega^i \mid \forall i = ArgC_j, \forall j\} \text{ is mapped into a simple support function } \nu_j \text{ with focus } \theta (\text{contra toxic spill}):$

$$\nu_j(\theta) = \prod_{i = ArgP_r} bpa(\omega_i), \quad \nu_j(\Theta) = 1 - \nu_j(\theta). \quad (5)$$

Then belief in each hypothesis, based on arguments built for each suspicious subcluster is computed as a combination of $\mu_k$ and $\nu_j$ for all $k$ and $j$ with the Dempster rule of combination. The final decision is based on the combination of beliefs obtained for subclusters belonging to clique of neighboring subclusters. Selection of a certain hypothesis is based on the decision process described in Section 2.2.

### 3.3 Experiments

Systematic testing of the proposed higher level data fusion inferencing system in an environment with a large number of degrees of freedom, such as DIRE, requires a great many runs to secure an acceptable degree of statistical significance in the factorial analysis against selected test metrics. These studies are currently underway, and are not yet available to be reported. Here we present the results of a set of preliminary pilot studies, in which an abstracted and simplified DIRE-like environment, containing many fewer degrees of freedom, is employed. The goal of these studies is to determine if the proposed inferencing scheme can discover secondary HAZMAT incidents from evidence hidden in the "report cloud" generated subsequent to a primary earthquake event, without also producing an unacceptably high rate of false alarms.

In the pilot studies, fused casualty reports were laid down in the $X-Y$ plane using a mixture distribution containing five components: three distinct Gaussian components representing clusters of non-HAZMAT-related earthquake casualties, a HAZMAT component to be described, and a uniformly distributed component of additional scattered non-HAZMAT casualties. At each time step, approximately 65 new reports were laid down by randomly selecting a component then sampling that component's distribution to create a single realization. This HAZMAT component begins at the onset of the secondary incident as a single Gaussian, and at each time step an additional Gaussian with mean displaced according to the assumed wind field (to model atmospheric plume transport), and standard deviation increased by a factor of 1.05 (to model dispersion), is added. This models a steady, rather than instantaneous bolus, release of the toxic material following rupture of the storage tank.

Each casualty from the HAZMAT component is reported to be of respiratory type, while 10% of all others, selected randomly, are so reported.

Three cases were considered, simulating different source conditions. In the first case, the HAZMAT spill was eliminated, so all HAZMAT clusters discovered would be false positives. In the second case, the HAZMAT source is well separated from the non-HAZMAT focal Gaussian casualty sources. In the third case, the initial HAZMAT Gaussian source distribution exactly matches one of the non-HAZMAT Gaussians, both in location and variance. This models a situation leads to a cluster of non-HAZMAT trauma injuries superposed with a cluster of HAZMAT injuries. The results of the experiments utilizing arguments based on the value and behavior of cluster characteristics only are shown in Table 1. The results are averaged over 5 Monte-Carlo runs.

In Case 1, in which there was no HAZMAT incident to be discovered, all clusters found at all times shown were correctly labeled non-HAZMAT. In Case 2, in which the HAZMAT location was isolated from the focal non-HAZMAT casualty sources, a single false negative was found at most times in one of the five Monte Carlo runs. It should be noted that at each of these times in that run a true positive was also found, so the HAZMAT incident would be identified, if not precisely localized, without delay at all times in all runs. In Case 3, the most difficult discovery case because the HAZMAT source was collocated with a non-HAZMAT source, once again there were false negatives, but always accompanied by true positives, thus enabling HAZMAT discovery. Note that a strict definition of false negative was used in this study: any cluster containing one or more casualties from the HAZMAT source was considered a positive cluster, even if it contained numerous non-HAZMAT respiratory and non-respiratory casualties making it very difficult to infer
the correct label. Case 3 also exhibited a single false positive in one run due to the random clustering of respiratory casualty labels from the uniformly distributed source component. In all, 534 clusters were found, of which 487 were correctly labeled, and the presence of a HAZMAT incident was discovered within a few time steps in all runs of all those cases in which a HAZMAT incident actually occurred.

<table>
<thead>
<tr>
<th>Time</th>
<th>Subclusters</th>
<th>True pos.</th>
<th>True neg.</th>
<th>False pos.</th>
<th>False neg.</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>2.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>4.4</td>
<td>0.0</td>
<td>4.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>12</td>
<td>7.0</td>
<td>0.0</td>
<td>7.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>16</td>
<td>6.2</td>
<td>0.0</td>
<td>6.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>20</td>
<td>5.4</td>
<td>0.0</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>24</td>
<td>5.8</td>
<td>0.0</td>
<td>5.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>Case 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>0.4</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.50</td>
</tr>
<tr>
<td>8</td>
<td>2.4</td>
<td>0.4</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>12</td>
<td>7.0</td>
<td>1.2</td>
<td>5.6</td>
<td>0.2</td>
<td>0.0</td>
<td>0.03</td>
</tr>
<tr>
<td>16</td>
<td>11.4</td>
<td>2.0</td>
<td>9.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.04</td>
</tr>
<tr>
<td>20</td>
<td>10.2</td>
<td>3.0</td>
<td>6.4</td>
<td>0.0</td>
<td>0.8</td>
<td>0.08</td>
</tr>
<tr>
<td>24</td>
<td>10.2</td>
<td>4.0</td>
<td>4.4</td>
<td>0.2</td>
<td>1.6</td>
<td>0.18</td>
</tr>
<tr>
<td>Case 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>2.6</td>
<td>1.8</td>
<td>0.4</td>
<td>0.0</td>
<td>0.4</td>
<td>0.15</td>
</tr>
<tr>
<td>8</td>
<td>3.4</td>
<td>2.4</td>
<td>0.8</td>
<td>0.0</td>
<td>0.2</td>
<td>0.06</td>
</tr>
<tr>
<td>12</td>
<td>5.6</td>
<td>2.4</td>
<td>3.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>16</td>
<td>7.4</td>
<td>1.8</td>
<td>4.4</td>
<td>0.2</td>
<td>1.0</td>
<td>0.16</td>
</tr>
<tr>
<td>20</td>
<td>6.6</td>
<td>1.8</td>
<td>3.2</td>
<td>0.0</td>
<td>1.6</td>
<td>0.24</td>
</tr>
<tr>
<td>24</td>
<td>8.0</td>
<td>2.2</td>
<td>3.4</td>
<td>0.0</td>
<td>2.4</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 1: HAZMAT cluster labeling accuracy

4 Conclusions

This paper reports progress in designing a general methodology for higher level fusion and its application to situation assessment and risk prediction in a dynamic post-disaster environment. The dynamic situational picture is built by analyzing spatial and temporal relations of the situational entities and entity aggregations at different levels of granularity, and their dynamics provided within the overall situational context. Special attention is paid to “inference for best explanation” aimed at discovery of the underlying causes of observed situational entities and their behavior, an important component of situation and impact assessment. A reasoning framework is constructed based on the generalized Probabilistic Argumentation System approach, which allows for allocating rational belief in hypotheses about the environment by utilizing given knowledge to find arguments in favor of and against a hypothesis. Subjective expert opinions based on the observed characteristics and behavior of the situational items are used to assign sub-additive measure of beliefs to uncertain assumptions used to construct the arguments. These beliefs are combined to produce the degree of support for each hypothesis further used in an anytime decision process. The presented method of discovery of underlying causes is illustrated by the discovery of an unreported HAZMAT incident within an early-phase earthquake response scenario. Implementation of this process within a synthetic task environment based on an earthquake early response scenario is described. Preliminary results obtained during a pilot study show the feasibility of the approach described.

Acknowledgements

This work was supported by the AFOSR under award F49620-01-1-0371. The authors express their gratitude to James Scandale for his valuable comments and help with simulations.

References