Using Subjective Logic to Handle Uncertainty and Conflicts

Murat Sensoy*, Jeff Z. Pan*, Achille Fokoue†, Mudhakar Srivatsa†, Felipe Meneguzzi ‡
*University of Aberdeen, Aberdeen, UK
†T. J. Watson IBM Research Center, New York, USA
‡Carnegie Mellon University, Pittsburgh, USA

Abstract—**In coalition operations, information from different sources belonging to different organisations have to be gathered and aggregated. Such information may not be consistent, and inconsistencies in the gathered information creates severe uncertainties that hinders the usefulness of the information. In this paper, we propose a Subjective Logic based approach for modelling the trustworthiness of information sources within a specific context. This model is used to handle inconsistencies through filtering information from less trustworthy sources.**

I. Introduction

In modern coalition operations, new sensing opportunities and information collection paradigms have proliferated, as exemplified by trends in Internet of Things [1] and crowd/participatory-sensing applications. In these cases, the closed, single-administrative-domain association between the information producers and consumers is challenged by more open and, hence, more complex and unpredictable, collaborative multi-administrative (and even no-administrative) domain associations. Knowledge of the capabilities of sensing entities may be unavailable and unknown, or policy-constrained, and certainly of questionable reliability. In addition, shared data may be deliberately manipulated for various reasons. As a result, it becomes harder to quantify the value of the fused data and the risks associated with acting on subsequent inferences.

Many tasks, such as reasoning about action, are critically dependent on making correct inferences about the state of the world, and therefore depend on the availability of specific data. However, this data is often not available, and must instead be obtained through third party data providers, which may conflict. Without resolving these conflicts, the resulting uncertainty may hamper critical inferences required to achieve one's goals. The reliability of data from these providers depends on the degree of *trust* in those providers. One way of handling conflicts is to filter conflicting information based on the reliability of their source.

In this paper, we consider information represented using an OWL ontology based on Description Logics [2]. Facts about the state of the world are gathered from different sources. Then, the following steps are followed: a) gathered information is fused into a KB, b) conflicting sets of facts are detected using off-the-shelf tools, c) trustworthiness of each conflicting set of facts is computed based on their source, and d) the facts with lower trustworthiness are removed to resolve conflicts.

II. Trust and Subjective Logic

We can define *trust* broadly as the willingness of one party (trustor) to rely on the actions of another party (trustee) [3]. Several approaches have been proposed to model trust.. A number of these approaches are based on *Subjective Logic* (SL), which is a belief calculus that allows agents to express opinions as degrees of belief, disbelief and uncertainty about propositions. Let ρ be a proposition such as "information source y is trustworthy in context c". Then, the binary opinion of agent x about ρ is equivalent to a Beta distribution. That is, the binomial opinion about the truth of a proposition ρ is represented as the tuple (b, d, u, a), where b is the belief that ρ is true, d is the belief that ρ is false, u is the uncertainty, and a is the base rate (a priori probability in the absence of evidence), as well as $b + d + u = 1.0$ and $b, d, u, a \in [0, 1]$. Opinions are formed on the basis of positive and negative evidences, possibly aggregated from different sources. Let r and s be the number of positive and negative past observations about y respectively, regarding ρ. Then, b, d, and u are computed based on Equation 1.

$$b = \frac{r}{r+s+2},\ d = \frac{s}{r+s+2},\ u = \frac{2}{r+s+2} \tag{1}$$

Then the opinion's probability expectation value is computed using Equation 2. Considering ρ, the computed expectation value can be used by x as the trustworthiness of y in the context c [4], [5].

$$t^x_{y:c}(r, s, a) = b + a \times u = \frac{r + a \times 2}{r+s+2} \tag{2}$$

The base rate parameter a represents a priori degree of trust x has about y in context c, before any evidence has been received. The default value of a is usually set to 0.5 in the literature [4], [5], which means that before any positive or negative evidence has been received, both outcomes are equally likely. While x has more evidence to evaluate trustworthiness of y, the uncertainty u, so the effect of a, decreases.

III. Knowledge Representation

In order to represent knowledge, we have used Description Logics (DLs). For convenience, we have selected to address the semantics of $\mathcal{SHIQ}$, which is equivalent to OWL-DL 1.0 [1] minus nominals and datatype reasoning, as shown in Table I

[1] http://www.w3.org/2001/sw/WebOnt

TABLE I
SHIQ DESCRIPTION LOGIC

Definitions	Semantics
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$
$\neg C$	$\Delta^{\mathcal{I}} \backslash C^{\mathcal{I}}$
$\exists R.C$	$\{x \mid \exists y. <x,y> \in R^{\mathcal{I}}, y \in C^{\mathcal{I}}\}$
$\forall R.C$	$\{x \mid \forall y. <x,y> \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}}\}$
$\leq nR$ C	$\{x \mid \mid \{<x,y> \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \mid \leq n\}$
$\geq nR$ C	$\{x \mid \mid \{<x,y> \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \mid \geq n\}$
R^{-}	$\{<x,y> \mid <y,x> \in R^{\mathcal{I}}\}$

(a) Constructors

Axioms	Satisfiability conditions
Trans(R)	$(R^{\mathcal{I}})^{+} = R^{\mathcal{I}}$
$R \sqsubseteq P$	$<x,y> \in R^{\mathcal{I}} \Rightarrow <x,y> \in P^{\mathcal{I}}$
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
$a : C$	$a^{\mathcal{I}} \in C^{\mathcal{I}}$
$R(a,b)$	$<a^{\mathcal{I}}, b^{\mathcal{I}}> \in R^{\mathcal{I}}$
$a \dot{\neq} b$	$a^{\mathcal{I}} \neq b^{\mathcal{I}}$

(b) Axioms

(We assume the reader is familiar with Description Logics [2]). Let $\mathcal{N}_C$ be the set of atomic concepts, $\mathcal{N}_R$ be the set of atomic roles, and $\mathcal{N}_I$ be the set of individuals. $\mathcal{N}_C$, $\mathcal{N}_R$, and $\mathcal{N}_I$ are mutually disjoint. Complex concepts and roles are built using constructs presented in Table I(a).

A $\mathcal{SHIQ}$ knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ consists of a Tbox $\mathcal{T}$ and an Abox $\mathcal{A}$. A Tbox $\mathcal{T}$ is a finite set of axioms, including:

- transitivity axioms of the form Trans(R) where R is a role.
- role inclusion axioms of the form $R \sqsubseteq P$ where R and P are roles. $\sqsubseteq^{*}$ denotes the reflexive transitive closure of the $\sqsubseteq$ relation on roles.
- concept inclusion axioms of the form $C \sqsubseteq D$ where C and D are concept expressions.

An Abox $\mathcal{A}$ is a set of axioms of the form $a : C$, $R(a,b)$, and $a \dot{\neq} b$.

As for First Order Logic, a model-theoretic semantic is adopted here. In the definition of the semantics of $\mathcal{SHIQ}$, $\mathcal{I}$= $(\Delta^{\mathcal{I}}, .^{\mathcal{I}})$ refers to an interpretation where $\Delta^{\mathcal{I}}$ is a non-empty set (the domain of the interpretation), and $.^{\mathcal{I}}$, the interpretation function, maps every atomic concept C to a set $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, every atomic role R to a binary relation $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and every individual a to $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$. The interpretation function is extended to complex concepts and roles as indicated in the second column of Table I(a).

An interpretation $\mathcal{I}$ is a model of a knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, denoted $\mathcal{I} \models \mathcal{K}$, iff. it satisfies all the axioms in $\mathcal{A}$, and $\mathcal{T}$ (see Table I(b)). A knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is consistent iff. there is a model of $\mathcal{K}$. Let α be an axiom, a knowledge base $\mathcal{K}$ entails α, denoted $\mathcal{K} \models \alpha$, iff. every model of $\mathcal{K}$ satisfies α.

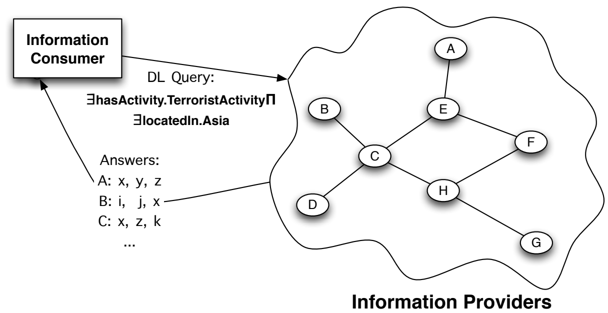

Fig. 1. A simple scenario demonstrating the interactions between consumers and providers

IV. COMPUTING TRUSTWORTHINESS

In this paper, we consider a simple scenario, where an information consumer is in need of a certain information and queries a set of information providers to access it. We assume each information provider has a knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and both information consumer and providers share the same $\mathcal{T}$ but have a possibly different $\mathcal{A}$. The queries sent by the consumer are represented as DL-queries, which are DL class descriptions based on the shared $\mathcal{T}$. Therefore, these queries can be clearly interpreted by the providers. Figure 1 demonstrates the scenario with a simple query example. In this figure, the consumer queries a set of information providers for learning locations in Asia where terrorist activities exist, i.e., $\exists hasActivity.TerroristActivity \sqcap \exists locatedIn.Asia$.

A. Context Representation

Trustworthiness of an information source is evaluated within a specific context c. We use DL class expressions to represent context with respect to a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$. A class expression describes a concept using the constructs introduced in Table I. For instance, a class expression composed of top concept $Thing$ refers to the most general context, while one like $\exists locatedIn.Asia$ refers to a context representing a query relevant to things located in Asia. That is, in this work, we assume each query constitutes a context in which trustworthiness of the queried nodes should be evaluated.

B. Context Hierarchies and Similarity

Given a set of context concepts, we can reason about the relationships between these concepts using a DL reasoner and can infer a taxonomy. The inferred taxonomy for a set of context concepts are shown in Figure 2. This taxonomy provides valuable information such as subsumption relationships or similarity between the contexts. There are different methods for computing similarity between two concepts within a taxonomy. Our model does not depend on a specific similarity metric. Once a taxonomy between concepts is derived, a specific method can be chosen by the agent to compute similarity between the search concepts.

In a concept taxonomy, semantic similarity between two concepts can be estimated by calculating the distance between these concepts. The length of the path between any two

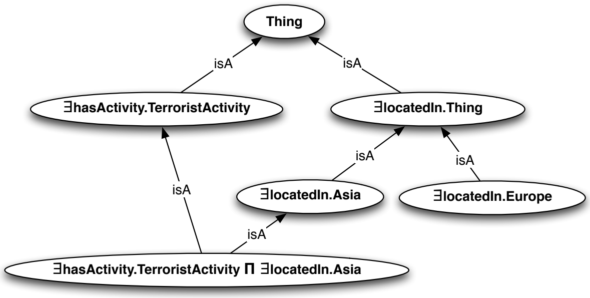

Fig. 2. Context hierarchy

concepts indicates how similar these concepts are. There are various distance-based semantic similarity metrics [6]–[8] that can be used for this purpose.

In this work, we describe the similarity metric proposed by Wu and Palmer [6] because of its intuitiveness and simplicity. Accordingly, we compute the similarity between c_1 and c_2 using Equation 3. Let c_0 be the most specific concept subsuming both c_1 and c_2. In the equation, N_1 is the length of the path between c_1 and c_0; N_2 is the length of the path between c_2 and c_0; lastly, N_0 is the length of the path between c_0 and the root of the concept taxonomy.

$$sim(c_1, c_2) = \frac{2 \times N_0}{N_1 + N_2 + 2 \times N_0} \tag{3}$$

An interesting property of distance-based similarity metrics is their instant response to the changes in the taxonomy. That is, while new concepts are added to or existing concepts are removed from a concept taxonomy, similarity between two specific concepts in the updated taxonomy may change immediately. This is because, in the updated taxonomy, the distance between these two concepts may increase or decrease after the insertions and deletions.

C. Parameter Estimation

For a given information source y, trustworthiness of y in a context c is computed using Equation 2. However, this computation requires three parameters: r, s, and a. Here, r and s refer to the number of times y has provided correct and incorrect information in the context c, respectively. In many settings, both r and s are zero since we do not have any evidence in c. So, the trust is solely based on the base rate a. Therefore, it is critical to select the right base rate a to compute trustworthiness of y in c when we have little or no evidence. We can assign a specific value to base rate for a specific context manually. For instance, if we set base rate of the context $Thing$ to 0.5, the apriori trust for y in the most general context becomes 0.5. Although we can assign base rates to some contexts, we cannot do it for all. Let us assume that a context $c' \in \mathcal{C}$ is not assigned a base rate. In this case, the information consumer x needs to estimate $a^x_{y:c'}$, the base rate for y in c', iteratively using Equation 4, where $sim(c, c')$ is the similarity function, $Super(c')$ is the set of direct super concepts of c' in the context hierarchy, and $t^x_{y:c}$ is the trustworthiness of y in c computed using Equation 2.

$$a^x_{y:c'} = \frac{\sum_{c \in Super(c') \wedge c \neq c'} sim(c, c') \times t^x_{y:c}}{\sum_{c \in Super(c') \wedge c \neq c'} sim(c, c')} \tag{4}$$

Using Equation 4, an information consumer may estimate the apriori trust (i.e., base rate) for an information provider in a new context c'. While doing this, the trustworthiness values of y in other contexts are used. In some settings, the consumer may not have any previous interaction with the provider. In these situations, the base rate estimated by Equation 4 would be determined by the default base rate values. If the base rate value for the most general context $Thing$ is set to 0.5, this would be the default base rate for any context for an unknown information provider.

Given a query Q of information consumer x, the trustworthiness of answers from an information provider y is equivalent to y's trustworthiness in the context Q, i.e., $t^x_{y:Q}$ computed based on Equation 2.

V. Conflict Resolution

Adding information from diverse information sources to a KB may make it inconsistent. An inconsistency in the KB can be resolved as follows. First, the justification of the conflict is derived using off-the-shelf reasoners such as Pellet [9]. Computing a single justification can be done fairly efficiently by 1) using some tracing technique to obtain a significantly small set S of axioms that is responsible for an inconsistency discovered by a single consistency test, and 2) performing additional $|S|$ consistency check on KBs of size at most $|S|-1$ to remove extraneous elements from S.

Unfortunately, computing all justifications is known to be intractable even for small and medium size expressive KBs. Kalyanpur [10] establishes a connection between the problem of finding all justifications and the hitting set problem (i.e., given n sets S_i, find sets that intersect each S_i). The intuition behind this result is the fact that in order to make an inconsistent KB consistent at least one axiom from each justification must be removed. Therefore, starting from a single justification a Reiters Hitting Tree can be constructed in order to get all justifications as illustrated in Figure 3. Starting from the first justification $J = \{2, 3, 4\}$ computed in the KB $\mathcal{K}$ (J is set to be the root v_0 of the tree), the algorithm arbitrarily selects an axiom in J, say 2, and creates a new node w with an empty label in the tree and a new edge $\langle v_0, w \rangle$ with axiom 2 in its label. The algorithm then tests the consistency of the $\mathcal{K} - \{2\}$. If it is inconsistent, as in this case, a justification J' is obtained for $\mathcal{K} - \{2\}$, say $\{1, 5\}$, and it is inserted in the label of the new node w. This process is repeated until the consistency test is positive in which case the new node is marked with a check mark. As an important optimization, we stop exploring super set of path discovered earlier and marked the node with 'X'.

The hitting set tree described above allows us discover different ways of resolving inconsistencies in a KB. For instance,

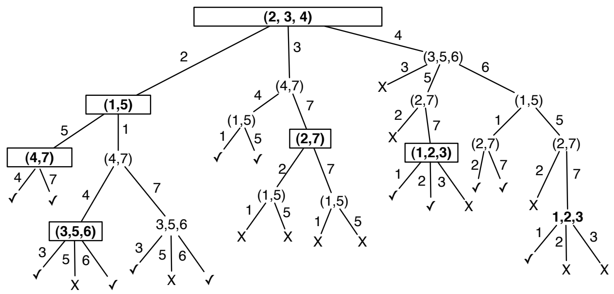

Fig. 3. Computing all justifications using Reiters Hitting Set Tree Algorithm from [10].

in Figure 3, we can resolve a conflict by removing one of several sets of axioms, e.g., $\{2, 5, 4\}$, $\{2, 5, 7\}$, $\{2, 1, 4, 3\}$ and so on. Let these sets be $s_0, s_1, \ldots, s_m$, where each set s_i contains only ABox axioms $A_{i0}, A_{i1}, \ldots, A_{ik}$. To resolve a conflict, we need to select one of these sets and remove all axioms in the set from the KB. Here, we propose to select this set based on the trustworthiness of axioms it has. That is, we select a set so that we remove the less trustworthy axioms from the KB to resolve a conflict. Trustworthiness of an axiom A_{ij} for information consumer x is computed using Equation 5, where $sources(A_{ij})$ is the set of sources that have provided A_{ij} to x and $t^x_{y:c(A_{ij})}$ is the trustworthiness of the source y in the context of A_{ij}.

$$A_{ij} = \arg \max_{y \in sources(A_{ij})} t^x_{y:c(A_{ij})} \tag{5}$$

The context $c(A_{ij})$ of the ABox axiom A_{ij} can be determined at different levels of granularity. For instance, if A_{ij} is a type assertion such as $\langle a, type, C\rangle$, then $c(A_{ij})$ can be C. If A_{ij} is an object property assertion such as $\langle a, P, b\rangle$, then $c(A_{ij})$ can be $\exists P.C$, where C is the type of b.

VI. Conclusion

In this paper, we have proposed a novel approach for estimating trustworthiness of information sources within a specific context using Subjective Logic. Then, we have shown how conflicts during information fusion can be resolved using the proposed trust model. In this preliminary work, we only propose the approach and leave its evaluation with real-life datasets and scenarios as a future work. Also, we would like to integrate the proposed approach into distributed query processing settings where the queried nodes in a data network have different level of trustworthiness in different contexts.

References

[1] N. Gershenfeld, R. Krikorian, and D. Cohen, "The Internet of Things," *Scientific American*, vol. 291, no. 4, pp. 76–81, 2004.

[2] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook*. Cambridge University Press, 2003.

[3] R. C. Mayer, J. H. Davis, and F. Schoorman, "An Integrative Model of Organizational Trust," *Academy of Management Review*, vol. 20(3), pp. 709–734, 1995.

[4] A. Jøsang, *Subjective Logic*. Book Draft, 2011. [Online]. Available: http://folk.uio.no/josang/papers/subjective_logic.pdf

[5] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decis. Support Syst.*, vol. 43, pp. 618–644.

[6] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, 1994, pp. 133–138.

[7] P. Resnik, "Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language," *Journal of Artificial Intelligence Research*, vol. 11, pp. 95–130, 1999.

[8] E. Blanchard, M. Harzallah, and P. Kuntz, "A generic framework for comparing semantic similarities on a subsumption hierarchy," in *Proceeding of 18th European Conference on Artificial Intelligence (ECAI 2008)*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008, pp. 20–24.

[9] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, and Y. Katz, "Pellet: A practical owl-dl reasoner," *Web Semant.*, vol. 5, pp. 51–53, June 2007.

[10] A. Kalyanpur, "Debugging and repair of owl ontologies," Ph.D. dissertation, College Park, MD, USA, 2006, aAI3222483.