Accepted in
Joint 3rd International Conference on Soft Computing and Intelligent Systems and
7th International Symposium on advanced Intelligent Systems (SCIS & ISIS 2006)

Fuzzy Trust and Combining Information, A Blueprint for the Semantic Web Trust Layer

Mohsen Lesani, Hassan Abolhassani, Saeed Bagheri
Sharif University of Technology
Azadi Avenue, Tehran, Iran
mohsen_lesani@mehr.sharif.edu, abolhassani@sharif.edu, s_bagheri@sharif.edu

***Abstract*— World Wide Web can be said to be the most prominent media of information offering and obtaining. In our time the start point for acquiring information from the web is seeking information using the search engines that direct us to many possibly relevant web pages. The user is responsible then for estimating trustworthiness of a page on the subject and combining the information form different pages to arrive at a unified unique knowledge about the subject. The semantic web envisions a time when machines do such jobs as trust estimation and information composition. This work proposes a combining procedure for information obtained from different sources according to estimated trust to the sources. The trust to sources plays an important role in integration of information from them. According to the vague nature of human language statements and trust to source estimation, an extended fuzzy description logic where a membership value and a numeric type individual can also be a fuzzy set is suggested to represent the formal representation of assertions. Two fuzzy rules are proposed for combining fuzzy description logic assertions and are illustrated in action through a practical example. It is shown that the composition procedure performs well at integrating even contradictory information from different sources by taking the trust to sources into account.**

Trust, Combining Information, Fuzzy DL, Fuzzy Implication, Semantic Web

I. Introduction

Some weeks before, I heard some news from one of my friends and he did not exactly know about its truth. It was very important for me to know if it is true so I started searching the news in Google. Google listed some pages that the news was been found there. I myself was responsible about knowing the trustworthiness of the page providers many of those I did not know. While knowing the most trustworthy news providers is not that much hard, the same also happens for any other bit of information you search on the web. The user him or herself should browse different pages and combine them according to the trustworthiness of providers. The task gets harder when you find contradictory information in different pages. The semantic web promise is to employ machines to do these tasks for us.

Obviously World Wide Web is an important source of information. When the web is searched for a question answer, many pages are faced with, that are likely to be relevant. A mechanism is needed to integrate information gotten from different sources and to yield a unified conclusion that decision can be made upon. The conclusion is desired to be more influenced by more trusted information. The question of how much a statement is trusted in the web is a hard question. It can be reduced to the question of how much the source of the statement is trusted. Different sources i.e. URIs are trusted to different degrees in a certain subject. The sources should be rated in trustworthiness on the searched subject according to some trust measures. As trust is of imprecise nature and the sources are subject to change, trust to a source can most practically be mentioned approximately so in this work trust measures are proposed and assumed to be fuzzy ones.

We receive information from different parties and are able to combine them and benefit from the result in our reasonings. The attempts to simulate the human reasoning ability have led to the currently logic based reasoning engines such as Racer DL Reasoner [1]. They are used because the information needed may be implied and not expressed explicitly and need to be interpreted semantically or may be also contradictory from a source to another. Although the human reasoning process may be different, at least by the at hand reasoning engines the information in a document should be converted to a logical form such as Description Logic before the reasoning can be done. While structured data i.e. some standard ontology compliant document is more readily usable information format at this phase, unstructured data is what that is faced more in the real world. Natural language documents are the predominant bulk of information on the web not only because of it being traditional but also because of hardness if not impossibility of forcing an ontology standard because of the distributed nature of the web and the world from a wider view.

Natural Language Processing is an ongoing research that helps in deciphering text and so the web pages and extracting the desired information. The information in the parsed sentences is converted to a logical form to be used in later

M. Lesani is AI MSc student in Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
Dr. Abolhassani and Dr. Bagheri are faculty members of the Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

Written in June 06.

natural language understanding systems phases. The information automatically interpreted from a text is hardly precise just like it may be when a human performs the interpretation of a text or a talk due to natural language imprecise nature. The knowledge representing logical form must well accommodate these imprecisenesses. On the other hand we also live in an uncertain world and the concepts encountered in the real world do not have precisely defined criteria of membership. We may say that an individual is an instance of a concept only to a degree depending on the individual's properties. For instance, Romantic is such a concept: we may say that an individual Titanic is an instance of the concept Romantic only to a certain degree depending on the film. And also the relations or roles may be true to a degree for example being skilled in jobs for individuals can have different degrees. These clues lead to modeling concepts that are sets of individuals and also roles that are sets of paired individuals as fuzzy sets. While the need to handling contradictory and vague information led to deployment of Fuzzy Description Logic in fact, the information obtained from a natural language text can be represented properly in FDL.

While coping with inconsistent and vague knowledge form different parties was discussed about the web in the preceding paragraphs because of its importance, the same problem occurs in ontology merging when two or more ontologies and knowledge bases are to be united or one is to be populated or supplemented from the others and also in multi agent systems when an agent should get knowledge from other agents according to their expertise in a subject [2].

Dempster's Rule of Combination is a classical evidence combination [3]. Josang has proposed more elaborated compositional methods as a logic [4]. Golbeck [5] has conducted a research on trust in social networks and has proposed the TidalTrust algorithm for trust inference in crisp trust graphs. Her proposed method for information combining is to average the numerical values weighted with the trust number values. The averaging performs poorly at handling conflicting information. Lesani [6] has extended the work from crisp modeling of trust to a fuzzy modeling and has proposed the FuzzyTrust algorithm for trust inference in the fuzzy trust graphs. The fuzzy trust inference was shown to well handle contradictory information. The issue that is seen in both of the works is the lack of distribute ness because they proposed the algorithms for the social networks and have assumed that the trust graph is centrally at hand. While this assumption is accepted within a social network, it is not applicable to the web where distributivity is a main characteristic.

While the Semantic Web down layers provide the ontologies and the knowledge bases represented according to them, the upper layers provide the reasoning techniques on the knowledge where trust is an issue that influences the reasoning. The trust layer should be built on top of the other in development layers mentioned in the layers cake, but some roadmaps should be made to track them until a desired design is gained. This work proposes the fuzzy model of trust and describes the procedure of combining fuzzy DL assertions according to the trust values to the sources where the concepts and roles are modeled as fuzzy sets, individuals of numeric type concepts are fuzzy sets and the trust values are also fuzzy sets. It is shown that the composition procedure performs well at unifying the information coming from different sources with magnifying more trusted while fading less trusted information in the result.

This paper starts with a brief explanation of DL, Fuzzy DL and an extension to Fuzzy DL which establishes a foundation for describing the assertion combining procedure in the next section. A practical example is illustrated afterwards and the conclusions and future works will conclude the paper.

II. Extended Fuzzy Description Logic

An assertion (denoted by *a*) in DL is an expression of type *C(a)* (*a* is an instance of *C* concept), or an expression of type *R(a, b)* (*a* is related to *b* by means of *R* role). For instance, *Tall(mohsen)* asserts that mohsen is tall, whereas *Friend(mohsen, mostafa)* asserts that *mostafa* is a friend of *mohsen* [7].

The previously proposed Fuzzy DL lets the assertions to be fuzzy i.e. they have a membership value in the [0, 1] range. From a syntax point of view, in fuzzy DL we are dealing with expressions of type *n a*, where *a* is a DL assertion and *n* called the membership is a number in the [0, 1] range. So DL is extended by allowing expressions of the form *n C(a)* with the meaning of "the membership degree of individual *a* being an instance of concept *C* is at least *n*" and also *n R(a, b)* with meaning of "individual *a* is related by *R* to individual *b* at least with the *n* degree" [8]. For example the *0.8 Tall(mohsen)* and *0.6 Friend(mohsen, mostafa)* assertions.

This work proposes to have the membership values as fuzzy sets over the belief range of [0, 1] (in the [0, 1] support set). Second order fuzzy sets are dealt with in fact as membership fuzzy sets for an assertion indicate the belief value to each crisp membership value (in the [0, 1] belief range) of concept or role assertion to the concept or role. This means having the assertions of the form *T a* (that are *T C(a)* or *T R(a, b)*) where *T* is a fuzzy term. For instance consider the assertions of *High Tall(mohsen)* and *Medium Friend(mohsen, mostafa)* meaning that *mohsen* is highly tall and *mohsen* relation with *mostafa* is a mediocre neither good nor bad one. The fuzzy sets for low, medium and high fuzzy terms are defined in fig. 1 for instance while any other number of fuzzy terms and different membership function definitions are possible.

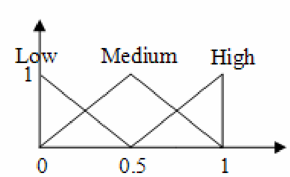

Fig. 1. Low, Medium and High Trust Term Fuzzy Sets

As the membership values are fuzzy sets in general, so they can be singleton fuzzy sets (a number in the [0, 1] range) as the previously proposed fuzzy DL assertions and hence what is proposed here is a generalization of previously proposed fuzzy DL. The membership values even can be proposed to be fuzzy sets with higher orders but that seems to have little practical application.

Also this work suggests using fuzzy values for continuous numeric type concepts such as concepts that indicate a real continuous finite or infinite range. These concepts have a continuous not a crisp nature, their values have the potential to be vague and so fuzzy set values are proposed to represent their individuals. Thus fuzzy set terms will be present in role assertions that have domains or ranges of numeric type concepts. For example salary, age, GPA concepts and so on all have a numeric type and can have fuzzy values in addition to crisp number values. It is usual that you get high or medium values for salary, very young or old values for age and high or low values for GPA. Role assertions of the forms *hasSalary(mohsen, high)*, *hasAge(mohsen, very young)* and *hasGPA(mohsen, high)* are possible with meanings of "mohsen has a high salary", "mohsen is very young aged" and "mohsen has a high GPA" respectively.

Research on fuzzy databases also proposes tuples with membership values and fuzzy values for the fields in accordance with what is called here as assertion membership and fuzzy values for numeric type domains or ranges in roles.

III. Fuzzy Trust based Combining of Fuzzy DL Assertions

A. Concept Assertions

If we have n different sources and from the source i we have assertion $SourceBelievedMembership_i$ $C(a)$ and source i is trusted by $TrustToSource_i$ fuzzy set value then a fuzzy set membership value $BelievedMembership_i$ is obtained from each source by inference with the following fuzzy rule:

$$if\ \left(TrustToSource\ is\ \ Acceptable\right)$$
$$then\ \begin{pmatrix} BelievedMembership\ is \\ SourceBelievedMembership \end{pmatrix} \quad (1)$$

The final resulting fuzzy set membership value is *BelievedMembership* fuzzy set the fuzzy union of the $BelievedMembership_i$ fuzzy sets. The result of the composition is thus *BelievedMembership C(a)*.

The Acceptable fuzzy set used in the rule is shown in fig. 2. An assertion that is trusted by 1 belief value is completely acceptable while one that is trusted by 0 belief value is completely unacceptable. The belief values in between to an assertion are partially accepted and increase linearly from 0 to 1. The Acceptable fuzzy set is defined so that the firing rate of the rule gives an importance to information of a source proportional to the trust to the source.

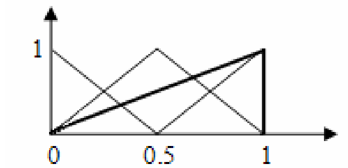

Fig. 2. The Acceptable Fuzzy Set (the darker lines)

Note the similarity of the rule to how people get a friend opinion as their own about an unknown matter if they acceptably trust the friend.

B. Role Assertions

If we have n different sources and from the source i we have assertion M_i $C(a, SourceBelievedValue_i)$ and source i is trusted by $TrustToSource_i$ fuzzy set value then there are two trust values in sequence that are $TrustToSource_i$ that is trust to the source and M_i that is trust of the source to the value. The two in sequence trust values should be combined to arrive at a unique trust to the value. This inferred trust value is used in the next step to perform combining different values according to trust to the values. The membership function of $Trust_i$ fuzzy set that is the trust to the value that is provided by the source i is taken from (2) with considering M_i and $TrustToSource_i$ fuzzy set membership functions as $\mu_A(x)$ and $\mu_B(x)$.

$$\mu_w(x) = \underset{w=\min(u(x),v(x))}{Sup} \min\{\mu_u(x), \mu_v(x)\} \quad (2)$$

Note that the intersection of two second order fuzzy sets A and B with the membership functions of
$\mu_A(x) = \{(u(x), \mu_u(x)) \mid u(x), \mu_u(x) \in [0, 1]\}$ and
$\mu_B(x) = \{(v(x), \mu_v(x)) \mid v(x), \mu_v(x) \in [0, 1]\}$
for each x in the discourse is
$\mu_{A\cap B}(x) = \{(w(x), \mu_w(x)) \mid w(x), \mu_w(x) \in [0, 1]\}$
where $\mu_w(x)$ is defined by (2) [9].

Then a fuzzy set $BelievedValue_i$ is obtained from each source by inference with the following fuzzy rule.

$$if\ \left(Trust\ is\ \ Acceptable\right)$$
$$then\ \begin{pmatrix} BelievedValue\ is \\ SourceBelievedValue \end{pmatrix} \quad (3)$$

The final resulting *BelievedValue* is the fuzzy union of the $BelievedValue_i$s. The result of the composition is thus *One C(a, BelievedValue)* where *One* is a singleton fuzzy set at 1. The role combining procedure is depicted in fig. 3.

Two different rules were proposed for concept and role assertions composition but a concept membership can also be viewed as a role.

M Concept (c) = One conceptHasMember (c, M)

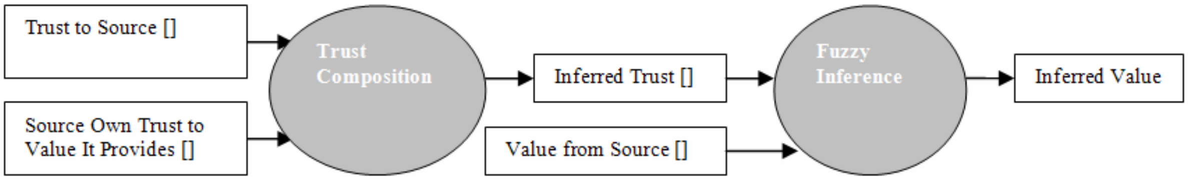

Fig. 3. The Role Combining Procedure

And so the only rule for role assertions composition can be employed for all assertion compositions. As the result of (2) with *One* singleton fuzzy set and any trust term fuzzy set as inputs is the trust term fuzzy set itself, the role assertions fuzzy rule can be reduced to concept assertions fuzzy rule when the faked proposed role assertion is used instead of a concept assertion for concept assertions integration.

What the fuzzy inferences return is the inferred trust fuzzy set. If the result is fed to an artificial deciding agent, currently fuzzy set form is the desired form, but if the client is a user the result should be converted to a more readable format. To make the inferred fuzzy set comprehendible to the user, it should be approximated to the most similar fuzzy set of the known terms or the known terms with hedges such as very and somehow and even a disjunction of them and report the corresponding linguistic expression such as "medium or somehow high" to the user.

A simple procedure for yielding approximating linguistic expressions is to prepare all the possible conjunction pairs of the terms with hedges and find the most similar one by an exhaustive similarity computation for them and the inferred fuzzy set, although more intelligent algorithms may be possible. The similarity of two fuzzy sets A and B is defined as:

$$S = \frac{|A \cap B|}{|A \cup B|} \tag{4}$$

Where || is the cardinality of a fuzzy set that is:

$$|A| = \int_{x \in SupportSet} \mu_A(x) \tag{5}$$

IV. Stock Price Change, A Practical Example

Consider a practical situation with an ontology with the concepts Stock and NumericValue and also the role hasPriceChange (Stock, NumericValue). Consider that the following fuzzy DL assertions are gained from different sources where the fuzzy sets are defined as in fig. 4.

Source 1: High, hasPriceChange (goldStock, Zero)
Source 2: High, hasPriceChange (goldStock, positiveMedium)
Source 3: High, hasPriceChange (goldStock, negativeLarge)
Source 4: High, hasPriceChange (goldStock, positiveLarge)

The membership of an assertion can be viewed as the truth belief of the source to the information it provides. It is assumed that the sources completely believe in their own information and so assertions are considered to have the default value of High or 1.0 for the membership value while they can have any other fuzzy set value too. It can be shown that result of (2) on any of the trust term fuzzy sets with the high term fuzzy set is the term fuzzy set itself, as expected. Assume the following trust values are assigned to the sources by a fuzzy trust measure:

Source 1: Medium
Source 2: High
Source 3: Low
Source 4: Medium

It is hard to judge about the stock price change with these at hand information as they are numerous and contradictory. There is a need for an integration to unify the at hand information to a unique integrated one consisting of all the available information according to trust values to them.

The resulting composed fuzzy set value for the gold stock price change obtained by the fuzzy Larsen implication method is shown in fig. 5. While the membership function of the result fuzzy set is ideal for a machine decision process, the fuzzy set can be approximated by a fuzzy set of a linguistic expression that is reported to a human user. The fuzzy set tells that the possibility of price raise is more than its fall so it may be better to buy some stock.

It is obviously seen that the result is influenced more with the information from the more trusted sources. For instance consider that the contribution of the second source with the high trust value is the most and the maximum possibility value in the result is derived from it. The counterpart argument can be done for the third source with the low trust value. It I seen that that the information from a source that is trusted with a low, medium and high value is scaled by 1/3, 2/3 and 1 factors respectively.

The superiority of fuzzy inference and a fuzzy set value for the result over weighted averaging and a number for the result is its ability to handle and integrate contradictory information and that the input information that are integrated are still present in the result while their contribution is scaled. Thus

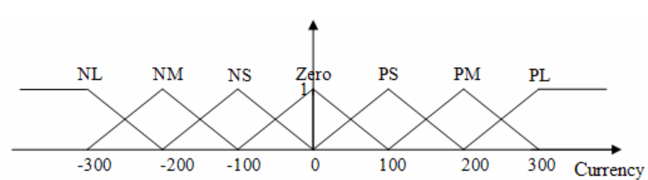

Fig. 4. NegativeLarge, NegativeMedium, NegativeSmall, Zero, PositiveSmall, PositiveMedium, PositiveLarge fuzzy sets

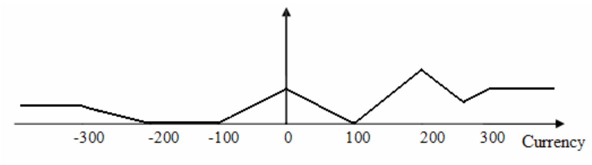

Fig. 5. The Result Fuzzy Value for Stock Price Change

richer information i.e. a sense of the contributing integrated input information is conveyed to the decision process that is absent in a reported number.

V. Conclusion And Future Work

The proposed procedure and fuzzy inference rules are appropriate for combining even conflicting information that are expressed in the proposed extended fuzzy Description Logic. The information coming from a source contributes in the resulting integrated information proportional to the trust value to the source. The trust to source is modeled to be a belief fuzzy set. Some fuzzy trust measures should be devised to derive such fuzzy trust ratings from some metadata about the source that is the future work of this research.

References

[1] V. Haarslev and R. Moeller. Racer: A core inference engine for the Semantic Web. In 2nd International Workshop on Evaluation of Ontology-based Tools (EON-2003)

[2] Majid Nili Ahmadabadi and Masoud Asadpour. Expertness Based Cooperative Q-Learning. IEEE Transaction on Sys, Man and Cybernetics – Part B: Cybernetics, Vol. 32, No. 1, pp. 66-76, Feb. 2002.

[3] Glenn Shafer. A Mathematical Theory of Evidence, Princeton University Press, Chapter 3, 1967.

[4] A. Jøsaang, L. Gray and M. Kinateder. Simplification and Analysis of Transitive Trust Networks. Web Intelligence and Agent Systems Journal (to appear), 2006.

[5] Jennifer Golbeck. Computing and Applying Trust in Web-based Social Networks. 200 pages, 2005, Doctor of Philosophy Dissertation, University of Maryland, College Park.

[6] Mohsen Lesani, Saeed Bagheri. Applying and Inferring Fuzzy Trust in Semantic Web Social Networks. Canadian Semantic Web Working Symposium Conference Proceeding (to appear), 2006.

[7] Franz Baadaer, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, 2003.

[8] Umberto Straccia. A Fuzzy Description Logic. In Proc. of the 15th Nat. Conf. on Artificial Intelligence (AAAI-98), pages 594–599, Madison, USA, 1998.

[9] Zimmerman. Fuzzy Set Theory and Its Applications. Kluwer Academic Publishers, Chapter 5, 1996.