COMPARABILITY BETWEEN FUZZY SETS AND CRISP SETS: A SEMANTIC WEB APPROACH

KAMALUDDEEN USMAN DANYARO 1, JAFREEZAL JAAFAR 1, SHAHIR LIEW 2

1 Department of Computer & Information Sciences
Universiti Teknologi PETRONAS, Seri Iskandar, 31750 Tronoh, Perak, MALAYSIA
2 Department of Civil Engineering
Universiti Teknologi PETRONAS, Seri Iskandar, 31750 Tronoh, Perak, MALAYSIA
kamaluddeen_g01325@utp.edu.my, {jafreez, shahir_liew}@petronas.com.my

Abstract: - Fuzzy set theories facilitate the extensions of today’s Web structure, especially in the context of Web data. Currently, the increase or sharing of data from different sources: individuals, industries, academia, etc are not knowing/ considering the origin of Semantic Web transition. As the extension of Web grows, the thorny problem of data heterogeneity is also increasing. However, the review/ comparability of the fuzzy sets, crisp sets and Semantic Web technologies were given based on their extensions. The authors provided the approaches that will help in knowing the uncertainty within the Semantic Web environment.

Key-Words: - Fuzzy sets, Crisp sets, Semantic Web, Description logics

1 Introduction
With proliferation of logically interconnected data from different sources in Web environment, data is moving in an uncertainty way. People tend to expect that the knowledge transition of Web has come by itself. The answer is no. Ordinarily, reasoning and knowledge representation are structurally and symbolically exist in the area of Artificial Intelligence (AI) [7], [10]. These built the Semantic Web transition which linked/coined fuzzy sets, crisp sets, Description Logics (DLs) and other Semantic Web technologies together.

Theoretically and logically, the Semantic Web technologies consist of eXtensible Markup Language (XML), Web Ontology Language (OWL), Resource Description Framework (RDF) and RDF Schema (RDFS). These allow people to leverage and share data across the globe using machine-readable or intelligent machine of Web pages [5], [8], [11]. As such, it enhances the understanding and transfer of knowledge over deductive reasoning processes. Therefore, to understand the reasoning processes, it is necessary to consider the uncertainty aspects which are the essential tools for modeling/ comparing fuzzy sets and crisp sets [2].

Fuzzy sets are the type of sets that deal with representations of vague Web services with the expressed concepts in natural language. In particular, fuzzy sets are the type of sets in set theory that are imprecise and no boundaries [2]. Accordingly, they only relay on the degree of their memberships.

While nonfuzzy (classical or crisp) set [2] is the traditional set theory with the concept uncertainty. But however [15] it is preserved as special fuzzy set theory. Also, it can be considered as the driver of the concepts and properties of ontologies in Semantic knowledge representation of uncertainty [7].

Thus, in order to capture the uncertainty using fuzzy logic, it is also of good advantage to compare the variables between fuzzy sets and crisp sets.

2 Background
Knowledge representation, uncertainty, imprecision, logics, categorization, expert systems and other soft computing disciplines are structurally developed under AI that deal with human interactions. In this regard, it is to measure the uncertainty with respect to set theories. Pal and Shiu [4] cited the following procedure for describing and measuring fuzzy sets over crisp sets based on retrieving and indexing data:

- Numerical characteristic could be change to fuzzy terms to ease comparison.
- It allows multiple indexing on one characteristic with different degrees of membership.
- It increases the flexibility in case retrieval with the help of modifiers.
- It makes simpler to transfer knowledge across domain.
2.1 Definition
Let \( U \) be the set of universal set and \( A \) be a fuzzy set then with continuum grade membership, \( \forall x (x \in U) \) assigned to be a membership value with \( x \) as imprecise concept [2], [4], [14]. Then \( A \) will be:

\[
A = \{ \mu_A(x) / x, x \in U \}
\]  

(1)

Where \( \mu_A(x) \), is the membership (characteristic) function, which defined by:

\[
\mu_A : U \rightarrow [0, 1]
\]  

(2)

If \( \mu_A(x) = 1, x \) is considered as membership of \( A \) and if \( \mu_A(x) = 0, x \) s considered as nonmembership of \( A \).

In other word, suppose \( P(x) = \) Proposition of the form “\( x \) has property \( P \)”. Then \( A = \{ x | P(x) \} \)  

(3)

Therefore, the rule base method.

It is good to note that the membership grades are usually represented in the interval between 0 and 1. However, graphically the membership function can be represented as:

![Membership Graph](image)

2.2 Definition
Crisp sets can be defined in a way that the individual in the group are:

- The members and belong to the group certainly
- The nonmembers and do not belong to the group certainly

3 Semantic Web Concepts
The ubiquitous and intrinsic reality of data heterogeneity makes Semantic Web more essential [13]. Consider its extension tools; RDF, RDFS and OWL. They allow agents or even programs to formalize and Work intelligently. Nevertheless, this can be done by supporting machine understandable information through the Uniform Resource Identifier (URI) of WWW [3]. Based on the standardization of RDF and crisp set theory, the closed association of bivalent semantics has given birth of new model-theoretic semantics [3].

4 Reasoning and knowledge representation
Fuzzy set theory has been used in soft computing environment for matching and reasoning. In Semantic Web context, the triples in RDF and the matching/mapping in OWL are the reasoning tools.

In particular, Reasoning over comprehensive data is important. Because the difficulties in data expression and vague knowledge the use of reasoning DLs [1]. Also, [3] noted that without the help of Semantic it is difficult to understand, analyze or optimize the behavior of reasoning systems.

Of course, the ultimate goal for reasoning is to enable "global knowledge-base" through intelligent machine over queried data. In fact, it can be achieved through capturing behavior and creating the algorithms. But, some authors; [1], [6], [7], [16] noted that, the capturing behavior in linguistic way is simple and it has to be comprehended based on the knowledge base agent near to human thinking. Significantly, to enhance the strength of Semantic Web over the real world knowledge, a fuzzy set structure with human interpretation must be incorporated [16].

Thus, the knowledge could be published explicitly on the Web and then implicitly inferable by machine. It is also a critical challenge for reasoning algorithms.

4.1 Description logic
As Linus Pauling said “The best way to get a good idea is to have a lot of ideas” [12]. Indeed, the idea of description logics (DLs) in Web environment comes through many ideas from fuzzy set theories. For example; OWL DL and OWL Lite were axiomatically OWL standards after the work from description logics [5]. In particular, looking at RDF(S) and OWL vocabularies, they were known as “T-Box” and “A-Box” respectively [7]. Also, [17] pointed that Description Logics (DLs) are the representation of knowledge languages mainly for particularization of ontologies which brings the knowledge bases and reasoning together. This idea is considered as the theoretical foundation of logical reasoning in Web services.

4.1.1 Syntax
The Semantic Web technologies and systems analyze the documents in RDF/XML formats. This enables the relationship between Web resources in form of values and properties [11] [18]. However, in reasoning concepts, the syntax of Semantic DLs which represent the OWL Lite and OWL DL are \( \text{SHIF(D)} \) and...
respectively [8]. They deal with the relationships of classes, membership and datatypes for the DLs knowledge base concept expressions [8], [9]. In particular, \( \text{SHOIN(D)} \) works with the reasoning concepts of concrete datatypes and it has a very expressive DL structure [14] [18], [19]. Therefore, the reasoning representation of fuzzy sets extension of \( \text{SHOIN(D)} \) have more power and capability than the classical (crisp sets) of \( \text{SHOIN(D)} \).

5 Fuzzy variables
The fuzzy variables are the tools that represent the linguistic concepts in the fuzzy theories.

Consider the example below given by Klir and Yuan [2] in Fig.2

![Fig.2 Examples of fuzzy sets that expressed the concepts of little educated (●), highly educated (◆), and very highly educated (■) people.](image1)

Where the temperature can be found within \([T_1, T_2]\).

However, to distinguish the fuzzy variable and crisp variable, it is by comparing the temperatures in Fig.3 which shows that the five linguistic concepts; very low, low, medium, high and very high as membership functions.

\[
[T_1, T_2] \rightarrow [0,1]
\]

![Fig.3 Fuzzy variable in \([T_1, T_2]\) range.](image2)

6 Conclusion
The review of this paper shows the basic incorporation between reasoning and knowledge for uncertainty in Semantic Web domain. The ideas presented are importantly to the Semantic Web knowledge representation in DLs. For this purpose, the fuzzy set is an extension of crisp (classical) set and Semantic Web is an extension of World Wide Web (WWW).

Thus, the knowledge representation integrates description logics and description logics integrate the Semantic Web. As such, these transitions solve the plight of data or datatypes representations in DLs which can be done by describing the vocabularies of imprecise information related to the degree of membership. However, this can be used to also solve the paradoxes in logical reasoning of new areas of AI (such as Semantic Web, expert system applications, etc).
6.1 Future and Recommendation
For future work, the authors will consider the use of fuzzy theories regarding the relationships between fuzzy sets and crisp sets which involve the use α-cuts and strong α-cuts principles, membership cardinality functions and designing the algorithms with respect to approaches of Semantic Web.

ACKNOWLEDGMENT
The authors would like to thank Universiti Teknologi PETRONAS for funding the presentation of this paper at the conference.

References: