Description logic-based knowledge merging for concrete- and fuzzy-domain ontologies

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was accepted for publication in Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture [SAGE Publications / © IMechE]. The definitive version is available at: http://dx.doi.org/10.1177/0954405414564404

Metadata Record: https://dspace.lboro.ac.uk/2134/17023

Version: Accepted for publication

Publisher: SAGE Publications / © IMechE

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Description Logic based Knowledge Merging for Concrete and Fuzzy Domain Ontology

Sri Krishna Kumar* and J. A. Harding**†
*Wolfson School of Mechanical Engineering, Loughborough University, UK
†Communicating Author: J.A.Harding@lboro.ac.uk

Abstract:
Enterprises, especially virtual enterprises (VEs) are nowadays getting more knowledge intensive and adopting efficient Knowledge management (KM) systems to boost their competitiveness. The major challenge for KM for VEs is to acquire, extract and integrate new knowledge with the existing source. Ontologies have been proved to be one of the best tools for representing knowledge with class, role and other characteristics. It is imperative to accommodate the new knowledge in the current ontologies with logical consistencies as it is tedious and costly to construct new ontologies every time after acquiring new knowledge. This paper introduces a mechanism and a process to integrate new knowledge in to the current system (ontology). Separate methods have been adopted for fuzzy and concrete domain ontologies. The process starts by finding the semantic and structural similarities between the concepts using Wordnet and Description logic (DL). DL-based reasoning is used next to determine the position and relationships between the incoming and existing knowledge. The experimental results provided show the efficacy of the proposed Method.

Keywords: Description Logic, Fuzzy Logic, Fuzzy-DL, Knowledge Merging, Knowledge Management, Ontology
INTRODUCTION:

In the current era of globalization industries are facing stiff competition through shorter product life-cycles, volatile markets and swift technological advancement. Under these circumstances, enterprises are collaborating to form large supply chain networks or virtual enterprises (VE). Such networks or VEs allow enterprises to provide advanced and multifaceted products or services to customers whilst focusing on their core-competencies and collaborating for other complementary aspects to remain competitive in the market. However, managing and operating a successful VE can be more complex than managing the individual member enterprises and strong communication, cooperation & collaboration in-line with interoperability is required between the member enterprises.

A successful VE needs to develop a mechanism for seamless transfer of data, information and knowledge among member enterprises¹. Information and communication technology (ICT) can help to achieve collaboration in VEs at a technical level² whilst ontologies have been proved to be important tools at the semantic level. Generally each individual enterprise builds its own ontology, based on the domain of their operation, to represent the enterprise’s knowledge. It is imperative that enterprise ontology should holds two features: 1. Interoperability (current aspect): to be able to collaborate with other enterprises and 2. Maintenance (continuous or on-going aspect): to accommodate new knowledge. Interoperability means that information and knowledge transferred using ontologies need to be understood accurately, i.e. with correct intention and extension³. However, as ontologies may be developed independently to suit personnel requirements, it is impossible to avoid heterogeneity in the terminology used for concepts and their relation and mappings between ontologies are required to interrelate different concepts and to achieve interoperability. Many mapping techniques have been adopted and proposed, in the literature, to achieve uniformity and to tackle interoperability of the enterprise ontologies (current aspect)⁴⁻⁵.

Nowadays knowledge has become one of the most precious resources for any enterprise. However, this knowledge is more valuable if it can be made inferable and deducible. The future success of enterprises is coupled with their knowledge assets so enterprises need to accumulate knowledge (or create knowledge) from information e.g.
by updating their knowledge in the form of ontology. According to Mo and Zhou, knowledge is power and its proper management is necessary to preserve valuable content, learn new things, solve problems, consolidate core competency and discover and implement new technologies. Enterprises should be able to maintain their ontologies to accommodate new knowledge to stay competitive and successfully collaborate in VEs not only in the current time but also in the future. For this reason, maintenance of ontology is termed as a continuous or on-going aspect of virtual enterprises.

Ontologies definitely play important roles in knowledge management, but the knowledge discovery possess is equally important to identify and accommodate new knowledge within existing ontologies. The discovered knowledge will not be useful unless it is mapped semantically and structurally with the existing ontologies. To merge knowledge correctly, both the syntax and semantics must be considered, in order to:

1. Deduce similar or new concepts
2. Deduce the possibility of merging concepts, i.e. by restructuring an ontology.
3. Achieve logically consistent mappings.

This paper tackles all three of these problems and develops a mechanism for ontology mapping in the same domain i.e. by enhancing the enterprises’ knowledge by accommodating new knowledge into an existing ontology. Moreover, this paper tackles the above problem by using the Description Logic (DL) paradigm as enterprises are increasingly using OWL (web ontology language) to store, use and transfer data and knowledge through the web and OWL is based on DL which is a fragment of first order logic (FOL). The proposed approach can be widely applicable in E-Commerce, product design, product development and medical domain, where new information is being gathered with time. Ontology based knowledge merging approach, proposed here, will help in making their knowledge bases coherent. Furthermore, nowadays enterprises are moving from traditional product lifecycle management (PLM) to knowledge based PLM, in which ontologies play a crucial role. This approach can help enterprises in updating their knowledge bases for improved and efficient knowledge based PLM.

The next section reviews in detail the current progress in the area. Section 3 gives the preliminaries about ontology construction methods (DL, fuzzy logic, fuzzy-DL) and
Wordnet used to identify the meaning and relation of the words used for concept creation. Section 4 describes a method for finding similarity between concepts. The process of merging and reconfiguration is dealt with in section 5. The proposed techniques have been implemented and they are demonstrated through an example which is presented in section 6 and section 7 concludes the paper.

**Literature Survey:**

Exhaustive surveys have been carried out on KM\textsuperscript{13-14} and its tools\textsuperscript{15}. KM in enterprises is mostly tackled at the subjective level and this can be divided in three different stages: 1. Knowledge creation 2. Ontology development for new knowledge and 3. Merging new knowledge in the existing sources.

Knowledge plays a significant role in the organizational performance\textsuperscript{16}. Due to the widespread application of different information systems, a large amount of different knowledge is accumulated during collaboration between enterprises. One of the most important factors in knowledge management is knowledge discovery. Proliferation of data has created a completely new and different area of knowledge management\textsuperscript{17} requiring the extraction of knowledge from abundant data and the organization and merging of this knowledge with existing knowledge. Existing knowledge supports organizations in creating new knowledge and updating the overall knowledge base\textsuperscript{18}. Knowledge discovery includes discovering implicit knowledge from the data, often using Data Mining techniques to extract knowledge from data sources. Exhaustive literature surveys illustrate that Knowledge Management frameworks, Knowledge-based systems (KBS), information and communication technology (ICT), artificial intelligence and expert systems, database technology etc. have all been adopted by enterprises to exploit knowledge in order to solve their current problems and enhance their expertise. A detailed review has been done by Liao\textsuperscript{15}. Pollalis and Dimitriou\textsuperscript{4} first proposed the different initiatives needed for knowledge creation and then developed the requirements at each stage of the KM-lifecycle.

Ontology based frameworks have been proven to be the ideal tools for knowledge representation as they provide uniform frameworks to identify similarities and differences between different entities in the specific domain\textsuperscript{9}. Many researchers have proposed different methodologies for ontology creation from new knowledge. Huang
and Diao\textsuperscript{19} proposed a methodology for creating a Concept Map based ontology construction method for knowledge integration. This accumulates knowledge in the business processes and rules and constraints are implemented using SWRL (semantic web rule language). However to implement this in the VE scenario, enterprises need to reconstruct their ontology every time they move to a new collaboration. Ling et al.\textsuperscript{20} proposed an ontology-based method to build an integrated knowledge base from heterogeneous sources operating in a single domain. Rajsiri et al.\textsuperscript{21} developed a knowledge-based ontology model for the collaborative business process model. A distributed enterprise system framework for KM is developed by Ho et al.\textsuperscript{22} (2004). Pirro et al.\textsuperscript{23} developed a framework for creating, managing and sharing knowledge within an organization with a distributed functional system. Mo and Zhou\textsuperscript{6} developed tools and methods for managing the intangible knowledge of VE. Ling et al.\textsuperscript{20} proposed an ontology-based method for knowledge integration in a collaborative environment. They used heterogeneous ontologies to build domain ontology, i.e. by merging them and through inconsistency elimination. Chen et al.\textsuperscript{24} used Wordnet and fuzzy formal concept analysis for merging domain ontologies. Raunich and Rahm\textsuperscript{25} proposed the ATOM (Automatic Target-driven Ontology Merging) for integration of multiple ontologies. The process was based on the equivalent relation between source and target taxonomy and merging them preserving the target taxonomy. PROMPT\textsuperscript{26} uses the class-name similarities and relies on user for specific merge operation whereas, OntoMerge\textsuperscript{27} uses the bridge ontology concept for ontology merging.

It has been widely reported that classical ontologies are not appropriate to deal with imprecise and vague knowledge inherent to several real-world domains\textsuperscript{28}. It is necessary to merge knowledge in an enterprise, not only for concrete domains, but also for fuzzy domains. Recently approaches have been reported for extending and reasoning with ontologies in fuzzy domains\textsuperscript{28-30}.

It is clear from the literature survey that the 3\textsuperscript{rd} stage of the KM in enterprises, i.e. merging new knowledge in the existing ones has been given little or no attention. This paper, introduces a method to map discovered knowledge with existing knowledge using an ontology and, if needed, reconfiguring the ontology.
**THEORY OF ONTOLOGY:**

This section describes the description logic (DL), fuzzy logic, fuzzy-DL and Wordnet used in this paper. DL is a decidable fragment of first order logic which acts as a backbone for ontology development. Fuzzy logic and consequently fuzzy-DL deal with the vague knowledge. Wordnet is helpful in finding semantic similarity between words. A detailed description of each of these approaches is given in the following section.

**Description Logic (DL):**

Description logic (DL) provides a logical construction for knowledge bases (KB) and is comprised of Concepts, Roles and Individuals as basic building blocks. DL has been proved to be most promising for processing, sharing and interpreting knowledge especially using the web. Ontologies play a key role in constructing KBs in a hierarchical manner of concepts and roles in a particular domain.

The formation of a KB in DL starts by defining the atomic concepts and atomic roles. Atomic concepts and roles generally represent the domain specific, self-explainable entities that are not defined using other concepts and roles (for more detail see Baader 31). Other general concepts and roles are defined using atomic concepts and general concepts, atomic roles and general roles and constructors (like union, intersection, quantifiers etc.). For example the concepts \( C \) are formed from atomic concepts using top concept \( (T) \), bottom concept \( (\bot) \), negation \( (\neg A) \), union \( (C_1 \cup C_2) \), intersection \( (C_1 \cap C_2) \), existential quantifier \( (\exists R.C) \), universal quantifier \( (\forall R.C) \), cardinality restriction \( (\geq_n R.C, \leq_n R.C) \) etc. Similarly Roles (R) are constructed from atomic roles \( (P) \), negation \( (\neg R) \), transitive \( (R^+) \), inverse roles \( (R^-) \) etc. Concepts and roles, in DL are seen as unary and binary relations such as: \( C(x) \) and \( R(y, z) \), where \( x \) satisfies the concept \( C \) and \( y \) and \( z \) are in relation \( R \).

However, simple DL is less appropriate in cases of imprecise definition of concepts and relations (roles) and therefore fuzzy-DL, which is a mixture of fuzzy logic and DL has been invented.

Integrating KBs does not simply mean joining an existing KB with a new one. Rather it requires a unified representation of entities (Concepts, Roles and Individuals) in the
merged KB. Moreover, an integrated KB must contain all the valuable knowledge and must be free from inconsistencies.

**Fuzzy Set and Fuzzy logic:**

Fuzzy set theory and fuzzy logic\(^3\) are widely adopted for capturing vague knowledge. Unlike the crisp set, where an element is either a member of a set or not, i.e. the binary (0 and 1) relation, a fuzzy set \(X\) and its members \(x_1, x_2, \ldots\) are related with a membership function, \(\mu: X \to [0,1]\). In other words, an element is a member of the set with a degree between 0 to 1. In a broader sense the crisp set can also be considered as a fuzzy set which takes only the boundary values 0 and 1.

Like DL, fuzzy logic also supports operations like complement, union, intersection, transitivity etc. with the help of strong mathematical principles. A Fuzzy complement \(c\) is a unary function defined by \(c: [0,1] \to [0,1]\) with interpretation \(I(\lnot x) = \Theta I(x)\). A Fuzzy complement satisfies the boundary conditions, i.e. \(c[0]=1, c[1]=0\) and is monotonically increasing, i.e. \(x \leq y \Rightarrow c(x) \geq c(y)\). There are many complement functions defined in the literature, among them are Lukasiewicz negation: \(c(x)=1-x\) and Godel complement \(c(x)=1\) if \(x>0\) else \(c(x)=0\).

Fuzzy intersection, termed as t-norm, is defined by the function \(t:[0,1] \times [0,1] \to [0,1]\) with interpretation \(I(x \land y) = I(x) \otimes I(y)\). Fuzzy intersection satisfies the boundary conditions: \(t(x,1)=x\) and \(t(x,0)=0\), monotonicity: \(y \geq z \Rightarrow t(x,y) \geq t(x,z)\) and other set theoretic properties like, commutativity: \(t(x,y)=t(y,x)\), associativity: \(t(x,t(y,z))=t(t(x,y),z)\). Most widely used t-norm functions are Lukasiewicz t-norm: \(t(x,y)=\max(0,x+y-1)\) and product t-norm: \(t(x,y)=x \cdot y\).

Fuzzy union, termed as t-conorm, is a function defined by \(u:[0,1] \times [0,1] \to [0,1]\) with interpretation \(I(x \lor y) = I(x) \oplus I(y)\). Similar to fuzzy complement and fuzzy intersection, fuzzy union also satisfies the boundary conditions: \(u(x,0)=x\), \(u(x,1)=1\) and monotonicity. It also follows the commutative and associative rules as in case of fuzzy intersection. Most commonly used t-conorm functions are Lukasiewicz: \(u(x,y)=\min(1,x+y)\), Godel: \(u(x,y)=\max(x,y)\).
One of the most important operations in fuzzy logic relating to classical logic is fuzzy implication. Fuzzy implication is defined by the function: \( I : [0,1] \times [0,1] \rightarrow [0,1] \) with interpretation \( I(x \rightarrow y) = I(x) \Rightarrow I(y) \). In classical logic implication \( A \rightarrow B \) is equivalent to \( \neg A \lor B \) or \( \max \{ t \in [0,1] \mid A \land t \leq B \} \). In classical logic both are equivalent but their extension in fuzzy logic leads to S-implication and R-implication respectively (ref.). Fuzzy implication functions are Luckasiewicz: \( I(x,y) = \min(1, x + y - 1) \), Godel: \( I(x,y) = y \) if \( x > y \), else \( I(x,y) = 1 \).

Although, there are many functions related to fuzzy logic, Luckasiewicz, Godel, product etc., Bobillo and Straccia\(^{29}\) showed the benefit of using Luckasiewicz function and this paper uses these functions for fuzzy interpretation.

**Fuzzy Description Logic:**

Fuzzy DL is an extended version of DL where concepts (unary relation) and roles (binary relations) are extended to fuzzy set and fuzzy binary relations. DL-axioms are also extended into fuzzy set using degree of truth. Similar to DL, fuzzy-DL consists of fuzzy-Tbox and fuzzy-Abox. Fuzzy-Tbox consists of the concepts (C, D) and role names (P, R) along with the general inclusion axioms i.e. \( \langle C \subseteq D \rangle \geq \alpha \) which means that concept C is sub-concept of concept D with truth value \( \alpha \) where, \( \alpha \in [0,1] \). Similarly for roles \( \langle P \subseteq R \rangle \geq \beta \) with \( \beta \in [0,1] \). The fuzzy A-box consists of the fuzzy assertion of individuals with concept and roles with fuzzy membership value in the form \( \alpha \beta \), where \( \alpha \in \{\geq, >, \leq, <\} \) and \( \beta \in [0,1] \). Like DL, interpretation of fuzzy-DL \( I_i \) is a pair \( (\Delta^i, \cdot^i) \), where \( \Delta^i \) is the fuzzy domain of interpretation and \( \cdot^i \) is interpretation function with the following characteristics:

For individual a, \( a \in \Delta^i \). For concept C, \( C^i : \Delta^i \rightarrow [0,1] \). For role R, \( R^i : \Delta^i \times \Delta^i \rightarrow [0,1] \). Comparison between DL and fuzzy-DL interpretation with basic concepts, roles and constructors is been shown in the table 1. (For more detailed explanations read Stoilos et al.\(^{28}\)).
<table>
<thead>
<tr>
<th>Concepts / Roles</th>
<th>DL Interpretation</th>
<th>Fuzzy-DL Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Concept : A</td>
<td>$A' \subseteq \Delta'$</td>
<td>$A' : \Delta' \rightarrow [0,1]$</td>
</tr>
<tr>
<td>Top concept: T</td>
<td>$\Delta'$</td>
<td>$T'(a) = 1$</td>
</tr>
<tr>
<td>Bottom concept: $\bot$</td>
<td>$\phi$</td>
<td>$\bot'(a) = 0$</td>
</tr>
<tr>
<td>Concept conjunction: $C \sqcap D$</td>
<td>$C' \cap D'$</td>
<td>$(C \circ D)'(x) = C'(x) \otimes D'(x)$</td>
</tr>
<tr>
<td>Concept disjunction: $C \sqcup D$</td>
<td>$C' \cup D'$</td>
<td>$(C \circ D)'(x) = C'(x) \oplus D'(x)$</td>
</tr>
<tr>
<td>Concept negation: $\neg C$</td>
<td>$\Delta' \setminus C'$</td>
<td>$\neg C'(x)$</td>
</tr>
<tr>
<td>Atomic role: $R$</td>
<td>$R' \subseteq \Delta' \times \Delta'$</td>
<td>$R' : \Delta' \times \Delta' \rightarrow [0,1]$</td>
</tr>
<tr>
<td>Inverse Role: $R^-$</td>
<td>${(y,x) \in \Delta' \times \Delta' \mid (x,y) \in R'$</td>
<td>$R'^-(y',x') \equiv R'(x',y')$ : $\Delta' \times \Delta' \rightarrow [0,1]$</td>
</tr>
<tr>
<td>Concept assertion: $a : C$</td>
<td>$a' \in C'$</td>
<td>$C'(a') \rightarrow [0,1]$</td>
</tr>
<tr>
<td>Concept Subsumption: $C \subseteq D$</td>
<td>$C' \subseteq D'$</td>
<td>$\inf{C'(x) \Rightarrow D'(x)}$</td>
</tr>
<tr>
<td>Role Assertion: $(a,b):R$</td>
<td>$(a',b') \in R'$</td>
<td>$R'(a',b') \rightarrow [0,1]$</td>
</tr>
</tbody>
</table>

Table 1: Comparison between DL and Fuzzy-DL

**Wordnet:**

Wordnet (wordnet API)\(^{33}\), created by Princeton university, is a dictionary of semantically similar English words, arranged structurally. Words are characterized based on the parts of speech- noun, verb, adjective etc. and linked together and categorized as synonyms, hyponyms etc.

**Ontology Similarity:**

An ontology is the explicit specification of shared conceptualization\(^ {34}\) (Gruber, 1993). In simple words, an ontology is a domain specific knowledge representation specified in terms of concepts and their relations. An ontology can be represented as $O := \{C, R, A\}$ where $C$ is the set of concepts, $R$ is the set of roles and $A$ is the set of axioms. Similarly a fuzzy ontology can be represented as $O^f := \{C, R^f, A^f\}$ where $R^f$ and $A^f$ additionally associate fuzzy membership values between $[0, 1]$. 
In this paper, an ontology $O_1$ is defined as the existing knowledge and $O_2$ as the new knowledge. Let $C_i^1$ and $C_j^2$ be the $i^{th}$ and $j^{th}$ concepts of two ontologies $O_1$ and $O_2$ respectively such that $C_i^1 \in O_1$ and $C_j^2 \in O_2$. All other notations used in this section are as follows:

- $Syn_{i}^1$: Synonym set of $C_i^1$,
- $Syn_{j}^2$: Synonym set of $C_j^2$
- $Hyper_{i}^1$: Hypernym set of $C_i^1$,
- $Hyper_{j}^2$: Hypernym set of $C_j^2$
- $Hypo_{i}^1$: Hyponym set of $C_i^1$,
- $Hypo_{j}^2$: Hyponym set of $C_j^2$
- $SC_{i}^1$: Set of super concepts of $C_i^1$,
- $sC_{i}^1$: Set of sub concepts of $C_i^1$,
- $SC_{j}^2$: Set of super concepts of $C_j^2$,
- $sC_{j}^2$: Set of sub concepts of $C_j^2$

The methodology adopted in this paper for knowledge merging, i.e. ontology mapping and ontology reconfiguration, is based on two steps. In the first step, a similarity matrix or index is calculated. In the second step merging and reconfiguration are carried out based on logical arguments to get a consistent final ontology.

For calculating the similarity matrix, two parameters have been taken into account: semantic similarity and structural similarity. Semantic similarity determines how closely two concept names are linguistically associated, whereas structural similarity determines the hierarchical relationship (equivalent, super and sub) between new concepts and concepts of existing ontology. The next section illustrates the process of calculating semantic similarity, structural similarity and hence the similarity matrix.

**Semantic Similarity:**

The Semantic similarity between concepts is defined by the function $\psi : \{C_i^1, C_j^2\} \rightarrow [0,1]$, where $C_i^1 \in O_1$ & $C_j^2 \in O_2$. In language two words can be related to each other in various ways e.g. same root, antonyms etc. However, the synonyms, hypernyms and hyponyms of two words imitate the equivalent, super and sub relationship of ontological concepts, and therefore only synonym, hypernym and hyponym relations have been taken into
account for calculating the semantic similarity. As concept names in general do not contain any fuzziness, rather instances and relations are fuzzy, this paper does not take fuzziness in the semantic similarity into account. The procedure for semantic similarity calculation is as follows:

**Synonym:** $\psi_1(C_i^1, C_j^2) = 1$, if $\exists t_1, t_2 \mid t_1 = t_2 \& t_1 \in \text{SynC}_i^1 \& t_2 \in \text{SynC}_j^2$

$$\psi_1(C_i^1, C_j^2) = 0$$ otherwise

**Hypernym:** $\psi_2(C_i^1, C_j^2) = \beta_2, \{ \beta_2 \in [0,1] \}$, if $\exists t_1, t_2 \mid t_1 = t_2 \& t_1 \in \text{HyperC}_i^1 \& t_2 \in \text{HyperC}_j^2$

$$\psi_2(C_i^1, C_j^2) = 0$$ otherwise

**Hyponym:** $\psi_3(C_i^1, C_j^2) = \beta_3, \{ \beta_3 \in [0,1] \}$, if $\exists t_1, t_2 \mid t_1 = t_2 \& t_1 \in \text{HypoC}_i^1 \& t_2 \in \text{HypoC}_j^2$

$$\psi_3(C_i^1, C_j^2) = 0$$ otherwise

Here, $\beta_2$ and $\beta_3$ are weights given to Hypernym and Hyponym relation. The final semantic similarity index will be the maximum of all, i.e. $\psi(C_i^1, C_j^2) = \max\{\psi_1(C_i^1, C_j^2), \psi_2(C_i^1, C_j^2), \psi_3(C_i^1, C_j^2)\}$

**Structural Similarity:**

The structural similarity between the concepts of two ontologies is the measurement of their association in terms of equivalence, super and sub relationships. The structural similarity is measured at three levels (equivalence, super and sub relation). As relationships between the concepts can be fuzzy, this paper considers both the concrete domain (instance 1) and the fuzzy domain (instance 2) for structural similarity calculation. The procedure is explained next.

**Equivalence relation Similarity (ER):**

**Instance 1 (Concrete domain):** In the concrete domain, an equivalence relation between two concepts is closely associated with the equivalence between their super and sub
concepts respectively. Mathematically, the equivalence relation between concept \( C_i \) and \( C_j \) can be given as:

\[
ER(C_i, C_j) = 0.5 \left\{ \frac{\text{Sim}(SC_i, SC_j)}{|SC_i \cup SC_j|} \right\}^2 + 0.5 \left\{ \frac{\text{Sim}(sC_i, sC_j)}{|sC_i \cup sC_j|} \right\}^2
\]

Here, function \( \text{Sim}(.,.) \) determines the number of similar elements in the two sets and \(|A|\) is the cardinality of the set \( A \). The first part of the equation calculates the similarity in terms of super concepts and the second part calculates the similarity in terms of sub concepts. Squaring the function gives more weightage to the structurally equivalent concepts as the ratio will never exceed the value 1. Equal weightage has been given to both the parts as two concepts are equivalent if their super and sub concepts are equivalent respectively.

**Case 2 (Fuzzy domain):** In the fuzzy domain the equivalence concept relation can be given as:

\[
ER(C_i, C_j) = 0.5 \left\{ \frac{\sum \sum \mu \text{Sim}(SC_i, SC_j)}{|SC_i \cup SC_j|} \right\}^2 + 0.5 \left\{ \frac{\sum \sum \mu \text{Sim}(sC_i, sC_j)}{|sC_i \cup sC_j|} \right\}^2
\]

Where, \( \mu \text{Sim}(.,.) \) determines the fuzzy value or truth value of the equivalence relation of the two input concepts. Unlike concrete domain square root of the function has been taken in the fuzzy case as the fuzzy value will not exceed the value 1 which will give more weightage to the structurally similar concepts.

**Super Relation Similarity (SR):**

A concept is said to be in a super concept relationship with another concept when its sub concepts match with the super concepts of the other. In this paper super relation similarity between the two concepts has been identified for both the cases as follows:

**Instance 1 (Concrete domain):** In the concrete domain, super relation similarity between two concepts is the similarity of their super and sub concepts. Mathematically this can be stated as:
\[
\text{SupR}(C_i^1, C_j^2) = \left\{ \text{Sim}(\{sC_i^1 \cup C_i^1\}, \{SC_j^2 \cup C_j^2\}) \right\}^2 \left\{ \frac{1}{sC_i^1 \cup SC_j^2 \cup C_i^1 \cup C_j^2} \right\}
\]

The super relation function (\text{SupR}) includes both concepts (\(C_i^1\) and \(C_j^2\)) in the denominator to get a closer evaluation.

**Instance 2 (Fuzzy Domain):** For the fuzzy case super relation similarity is

\[
\text{SupR}(C_i^1, C_j^2) = \sqrt{\sum \sum \left\{ \mu \text{Sim}(\{sC_i^1 \cup C_i^1\}, \{SC_j^2 \cup C_j^2\}) \right\}^2 \left\{ sC_i^1 \cup SC_j^2 \cup C_i^1 \cup C_j^2 \right\}}
\]

and is a measurement of the equivalent fuzzy value or truth value for the super relation between the two input concepts.

**Sub Relation Similarity (sR):**

In contrast with super relation similarity, a concept is in a sub relationship similarity with another concept if its super concepts match with the sub concepts of the other. For both cases this can be calculated as follows:

**Instance 1 (Concrete domain):** In line with the argument given in the super relation, a concrete domain sub relation can be given as:

\[
\text{subR}(C_i^1, C_j^2) = \left\{ \text{Sim}(\{SC_i^1 \cup C_i^1\}, \{sC_j^2 \cup C_j^2\}) \right\}^2 \left\{ \frac{1}{SC_i^1 \cup sC_j^2 \cup C_i^1 \cup C_j^2} \right\}
\]

**Instance 2 (Fuzzy Domain):** For fuzzy domain it will be:

\[
\text{SubR}(C_i^1, C_j^2) = \sqrt{\sum \sum \left\{ \mu \text{Sim}(\{SC_i^1 \cup C_i^1\}, \{sC_j^2 \cup C_j^2\}) \right\}^2 \left\{ SC_i^1 \cup sC_j^2 \cup C_i^1 \cup C_j^2 \right\}}
\]

Now the overall mapping relation based on semantic and structural similarity can be given as:

1. **Equivalence relation:**

\[
\mathcal{E}\mathcal{R}(C_i^1, C_j^2) = k\psi(C_i^1, C_j^2) + (1-k)ER(C_i^1, C_j^2)
\]
2. **Super relation:**

\[
\mathcal{S}_\mathcal{R}(C_i^1, C_j^2) = k\psi(C_i^1, C_j^2) + (1 - k)\text{Sup}(C_i^1, C_j^2)
\]

3. **Sub relation:**

\[
\mathcal{S}_\mathcal{R}(C_i^1, C_j^2) = k\psi(C_i^1, C_j^2) + (1 - k)\text{Sub}(C_i^1, C_j^2)
\]

The constant \( k \in [0, 1] \) is the weight given to the semantic relation. The relational matrix obtained here is used for ontology merging and reconfiguration (explained in the next section). The relational matrix serves two purposes as it not only relates the closeness of two concepts from different ontologies but also explores the kind of relation i.e. equivalence, super and sub. The next section describes the process of ontology merging and reconfiguration.

**Ontology merging and Reconfiguration:**

The relational matrix obtained in the previous section is used for ontology merging and, if necessary, for ontology reconfiguration. This approach first determines the greatest similarity in terms of equivalence, super and sub relations between the concepts of the ontologies. The next step involves establishing logical consistency, i.e. the formation of a logically consistent merged and reconfigured ontology. This process is different for both the concrete and fuzzy domains. This section begins by explaining the process for the concrete domain and later deals with the fuzzy domain.

**Concrete Domain:** In the consistency checking part, two concepts of two different ontologies are compared. This process first finds the maximum of \( \mathcal{E}_\mathcal{R}(C_i^1, C_j^2), \mathcal{S}_\mathcal{R}(C_i^1, C_j^2), \mathcal{S}_\mathcal{R}(C_i^1, C_j^2) \) (see figure 20). In case two or more concepts have the maximum value arbitrary selection is carried out. Separate reasoning is carried out for each equivalence, super and sub relation (as explained next). In case, no consistent relation is derived for the maximum value, then the process selects the next best value and so on until a solution is reached or no relation is found. In case no relation is found it is added as a new concept. The detailed explanation is given next.
**Equivalence relation:** The easiest case is when the equivalence relation matrix is one (case 1) and the new concept ($C^2_j$) can be established in the existing ontology as shown in fig (1).

Case 2 (fig. 2) is where the equivalence relation matrix is less than one and a possible position for the new concept ($C^2_j$) is as a sibling of concept ($C^1_i$). This case arises when $C^2_j$ is the sub-concept of $SC^1_i$ but $C^1_i$ and $C^2_j$ do not have any common sub-concepts. In this case the equivalence relational matrix will have a greater value than the sub and super relational matrix.

**Sub relation:** This is where the existing concept is in a sub-concept relation according to relational matrix i.e. $C^1_i \subseteq C^2_j$. In this case, three positions are possible where the new concept can be merged in the ontology, as shown in the figures 3 to 5.

The first condition (case 3) arises when the new concept ($C^2_j$) is equivalent to the super concept ($SC^1_i$) of the compared concept ($C^1_i$). The second condition (case 4) arises when the new concept ($C^2_j$) is a super-concept of the compared concept ($C^1_i$) and is also
a sub-concept of the super concept \((SC^1_i)\). This situation arises when a concept in an ontology is further subdivided or refined. The third condition (case 5) arises when the new concept \((C^2_j)\) is super concept of \((SC^1_i)\), as shown in the figure (5).

![Figure 3: Case 3](image1)

![Figure 4: Case 4](image2)

In this case, the position of \(C^2_j\) is above \(SC^1_i\), but to get the exact place \(C^2_j\) must be compared with \(SC^1_i\) and then the conditions (3) and (4) should be checked again.

![Figure 5: Case](image3)
**Super Relation:** This is where the existing concept is in a super-concept relation according to the relational matrix i.e. $C^S_j \subseteq C^I_i$. In this case, possible positions where the new concept can be merged in the ontology are shown in figures 6 to 9.

The first condition (case 6) arises when $C^S_j$ is a sub concept of $C^I_i$ and a super concept of $sC^I_i$. The second condition (case 7) is when $C^S_j$ is a sub concept of $C^I_i$ and is equivalent to $sC^I_i$. The third condition (case 8) arises when $C^S_j$ is a sub concept of $C^I_i$ and is disjoint with $sC^I_i$. This scenario describes the condition when a concept is redefined with the addition of new concepts (or new characteristics).

The last condition (case 9) describes the situation when $C^S_j$ is a subclass of $sC^I_i$ (fig 9). In this condition, to get the exact position of $C^S_j$, it must be compared with $sC^I_i$ and further evaluated for conditions 6-8.
Although this approach has considered all possible conditions for equivalence, sub and super relations, it may also be possible that the new concept has no defined position or possibly has no relation with existing concepts (including case 5 and 9). In this scenario merging and reconfiguration is carried out using the super concept of the new concept. Let $SC_j^2$ be the super concept of $C_j^2$ and the relational matrix is obtained in the same manner as in the case of $C_i^1$ and $C_j^2$. The following conditions can be obtained in line with the previous explanations (Prefix ‘Super’ ($S$) has been used to emphasise that super concept of a new concept is compared to get the relational matrix):

**Super equivalence relation:** In a super-equivalence relation, the mapping of a new concept ($C_j^2$) in terms of its super-concept ($SC_j^2$) with respect to $C_i^1$ follows the same procedure as the mapping between $C_i^1$ and $C_j^2$. The simplest condition is when $SC_j^2$ is equivalent to $C_i^1$ (Condition $S1$). This is depicted in figure (10). As no relation is found between $C_j^2$ and the existing ontology, this is simply added as a sub concept of $SC_j^2$ in the merged ontology.
Similar to condition 2, condition $\tilde{S}2$ arises (fig 11) when a new concept $SC_j^2$ is added in
the ontology as a sub-concept of $SC_i^1$ and $C_j^2$ is added in the ontology accordingly.

![Figure 10: Case $\tilde{S}1$](image1)

![Figure 11: Case $\tilde{S}2$](image2)

**Super-Sub-relation:** Super-Sub-relation mapping is carried out when the relational
matrix entails that $SC_j^2$ is closer to the sub concept of the $C_j^1$. Condition $\tilde{S}3$ (fig 12) and
$\tilde{S}4$ (fig 13) have same logical base as cases 3 and 4 respectively. Similar to condition 5,
in the case of condition $\tilde{S}5$ (fig. 14), the super concept of $SC_j^2$ is checked with $C_j^1$ for
conditions $\tilde{S}3$ and $\tilde{S}4$. 
**Super-super relation:** Super-super relation mapping occurs when the relational matrix intimates that $C^i_j$ is the super class of $SC^2_j$. All possible places where $SC^2_j$ and $C^2_j$ can fit...
have been shown in the figures 15-18. Conditions §6, §7 and §8 have similar logical explanations as conditions 6, 7 and 8 respectively.

Similar to condition 9, in condition §9 the exact position of $SC^2_j$ cannot be determined. It is compared with $sC^1_i$ and condition 6, 7 and 8 are checked with respect to $sC^1_i$ and $SC^2_j$.

If none of the conditions (Case 1 to Case 9 and Case §1 to Case §9) are satisfied in this process then it is clear that a new concept needs to be added in the ontology. For this, the process finds the super most concepts, unrelated to the existing ontology, as a new concept and adds its sub-concepts accordingly.

Figure 15: Case §6

Figure 16: Case §7
Fuzzy domain:

In fuzzy-DL two concrete concepts or even an assertion of individual in concrete concepts and roles are related with the fuzzy value or truth value as described in the section 3. Unlike the concrete domain, as explained earlier, a fuzzy domain ontology not only requires mapping and reconfiguration of the ontology but also requires a fuzzy value or truth value to be calculated for the merged or reconfigured concept with the concepts of the existing ontology. The first step for fuzzy domain ontology mapping and reconfiguration is similar to the concrete domain (i.e. finding the maximum). The second step is implemented in two stages: the first stage determines the position of the new ontology within the existing one (similar to case 1 to case § 9) and the second stage recalculates the fuzzy value or truth value for different relations among the concepts.

Considering the case (4) as shown in figure (4), following fuzzy values are available:

\[ \text{SubR}(C_i^1, C_j^1) \equiv C_i^1 \sqsubseteq C_j^1 = \mu \]  \hspace{1cm} (A.1)
SupR(\(SC_j^i, C_j^i\)) \(\equiv C_j^2 \subseteq SC_j^i = \omega\) .............................................. (A.2)

From A.1 and A.2, it is clear that \(C_j^1 \subseteq C_j^2 \subseteq SC_j^i\). This step merges or reconfigures the ontologies and introduces the intermediate concept \(C_j^2\). This reconfiguration or introduction of the new concept needs to identify the fuzzy sub-concept value between \(C_j^2\) and \(SC_j^i\). Assuming,

\[C_j^2 \subseteq SC_j^i = \lambda\] ......................................................... (A.3)

Now the sub-concept relation \(C_j^1 \subseteq SC_j^i\) is the implication of two sub-concept relations \(C_j^1 \subseteq C_j^2\) and \(C_j^2 \subseteq SC_j^i\), i.e.

\[(C_j^1 \subseteq SC_j^i) \equiv (C_j^1 \subseteq C_j^2) \rightarrow (C_j^2 \subseteq SC_j^i)\] ......................................... (A.4)

Using the Lukasiewicz implication function and A.1 – A.4,

\[\omega = \min(1, 1 - \mu + \lambda)\]

The method shown above for truth value recalculation has considered only truth values equal to some constant, but a similar approach can be used in cases where the truth value is greater than (\(\geq\)) or less than (\(\leq\)) some constant value. The rest of the ontology mapping and reconfiguration in fuzzy case can be obtained for all cases (case 1 to case §9) in a similar way.

**Implementation Method:**

The proposed methodology for ontology merging and reconfiguration for both concrete and fuzzy domains can be created in an OWL API such as Protege\textsuperscript{35} and its plugin fuzzyDL\textsuperscript{36}. Merging and reconfiguration is carried out in Java. The overall implementation method is summarized in figure (19). Jena parser, a Java API is used as the Ontology API to get the concept names without the namespace. Wordnet API\textsuperscript{33} (Wordnet) is used to get the synonym, hyponym and hypernym of the concepts for carrying out the word similarity and finally calculating the Lexicon similarity matrix. A Structural similarity matrix is calculated as previously described. Pellet-reasoner\textsuperscript{37} is
used to find the relationship between concepts (i.e. equivalence, super, sub etc.). As the two ontologies considered here are from the same domain, the assumption that they are built on same base ontology is valid. This assumption has been used for building the Tbox and Abox for reasoners.

The Semantic similarity matrix and Structural similarity matrix are used to calculate the relational similarity matrix. The next steps, merging, reconfiguration and consistency checks are illustrated in figure (20).
Example:

**Concrete domain**

In order to illustrate the overall procedure, two ontologies in the concrete domain of car manufacturing have been developed from different car parts website available on
the internet. As shown in figure (21), The Car ontology describes the current knowledge of the field, whereas the New Car ontology represents new knowledge in the field of car manufacturing. In order to merge the two ontologies, reconfigure an existing ontology (Car Ontology), or incorporate the new knowledge (New Car ontology) the process as described in the previous section is carried out.

Figure 21. Input Ontologies
With the help of the Jena parser, concept names are identified. The next step involves calculating the similarity matrix. This step comprises of calculating the Lexicon similarity matrix and Structural similarity. Consider the two concepts: Water from the Car ontology (C: Water) and Oil from the New Car ontology (NC: Oil). As there is no similarity between the two concepts in terms of synonyms, hypernyms and hyponyms, their lexicon similarity:

$$\psi(C: Water, NC: Oil) = 0.$$  

For calculating the structural similarity, super and sub concepts need to be identified.

$$S(C: Water) = \{\text{Cooling, Engine, Car}\}$$

$$s(C: Water) = \emptyset$$

$$S(NC: Oil) = \{\text{Cooling, Engine, Car}\}$$

$$s(NC: Oil) = \emptyset$$

As both child concepts are empty sets,

$$ER(C: Water, NC: Oil) = 0.5 \left\{ \frac{\text{Sim}(SC_i^1, SC_j^2)}{|SC_i^1 \cup SC_j^2|} \right\}^2 + 0.5 \left\{ \frac{\text{Sim}(sC_i^1, sC_j^2)}{|sC_i^1 \cup sC_j^2|} \right\}^2$$

$$= \left( \frac{3}{3} \right)^2 = 1$$

As both child concepts are empty sets it follows that:

$$\text{Sup}(C: Water, NC: Oil) = 0 \text{ and}$$

$$\text{Sub}(C: Water, NC: Oil) = 0$$

Now, considering equal weightage for semantic and structural similarity

$$\mathcal{E}R(C: Water, NC: Oil) = 0.5 \ \psi(C: Water, NC: Oil) + 0.5 \ \text{ER}(C: Water, NC: Oil)$$

$$= 0.5$$

Similarly,
In a similar manner the similarity matrix is calculated between ‘NC:Oil‘ and all the concepts of Car Ontology. The non-zero values obtained are:

\[ \text{SR}(C:Water, NC: Oil) = 0 \text{ and } \text{SR}(C:Water, NC: Oil) = 0. \]

The next step involves the merging of the new concept into the existing concept. As both the similarity matrix indexes have the same value i.e. 0.5, the algorithm arbitrarily selects one of them and tries to merge it logically into the existing concept as explained in section 5. Taking ‘C: water‘ the logical relations obtained are:

\[ C: Cooling \equiv NC: Cooling \text{ (From TBox similarity) } \]
\[ NC: Oil \subseteq \{ NC: Cooling \equiv C: Cooling \} \]
\[ NC: Oil \cap C: Water \equiv \bot \]

Clearly, conditions (a), (b) and (c) lead to the case (2) and ‘NC: Oil \subseteq C: Cooling‘ is established as shown in the figure 22. Table 2 depicts the overall result obtained in this process. In the case of the same similarity index occurring between more than one concept, random selection process has been adopted.
Figure 22. Merged / Reconfigured Ontology
<table>
<thead>
<tr>
<th>New Concept</th>
<th>Max similarity with Existing Concept</th>
<th>Type of similarity</th>
<th>Logical Case</th>
<th>Final relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC: Car</td>
<td>C: Car</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Brake</td>
<td>C: Brake</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Drum_Brake</td>
<td>C: Drum_Brake</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Single_Leading_edge</td>
<td>C: Single_Leading_edge</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Drum_Brake</td>
</tr>
<tr>
<td>NC: Power_Brake</td>
<td>C: Power_Brake</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Air_Suspended</td>
<td>C: Air_Suspended</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Power_Brake</td>
</tr>
<tr>
<td>NC: Vaccum_Suspended</td>
<td>C: Vaccum_Suspended</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Power_Brake</td>
</tr>
<tr>
<td>NC: Engine</td>
<td>C: Engine</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Cooling</td>
<td>C: Cooling</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Oil</td>
<td>C: air</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Cooling</td>
</tr>
<tr>
<td>NC: Direct_Injection</td>
<td>C: Direct_Injection</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Fuel.Injector</td>
</tr>
<tr>
<td>NC: Valve</td>
<td>C: Valve</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Spring_valve</td>
<td>C: Spring_valve</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Valve</td>
</tr>
<tr>
<td>NC: Safety</td>
<td>C: Safety</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Seat_belt</td>
<td>C: Seat_belt</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Safety</td>
</tr>
<tr>
<td>NC: Air_bags</td>
<td>C: Air_bags</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Safety</td>
</tr>
<tr>
<td>NC: Automated_Braking</td>
<td>C: Automated_Braking</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Safety</td>
</tr>
<tr>
<td>NC: Airbags</td>
<td>C: Airbags</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Safety</td>
</tr>
<tr>
<td>NC: Rack_and_pinion</td>
<td>C: Rack_and_pinion</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Rack_and_pinion</td>
</tr>
<tr>
<td>NC: Cam_and_Lever</td>
<td>C: Cam_and_Lever</td>
<td>TBox</td>
<td>-</td>
<td>Equivalent</td>
</tr>
<tr>
<td>NC: Worm_and_Nut</td>
<td>C: Worm_and_Nut</td>
<td>Equivalent</td>
<td>Case 2</td>
<td>Ø C: Worm_and_Nut</td>
</tr>
</tbody>
</table>

Table 2: Merging process outcome

Although, example presented here describes the knowledge merging process within an enterprise but this process can be extended in case of merging of two different enterprises with different ontology based data bases built on same domain.

**Fuzzy Domain:**

In the case of a fuzzy ontology, the process of finding the similarity index and determining the position of merging of a new concept with the existing ontology has been explained in sections 4 and 5. In this process Fuzzy domain ontologies differ from concrete domain ontologies only in the step comparing the fuzzy values of the new concept with the existing ones.
Taking an example as shown in figure 23, a concept 'Engine failure' has different sub-concepts with different fuzzy values e.g. 'Low Fuel $\delta_{0.3}$ Engine Failure', and in the new knowledge as shown in figure 24, 'Low air pressure $\delta_{0.9}$ Clogged air filter'. These two ontologies are first compared for similarity and then to determine the position of the new concept within the new merged ontology, the same process is followed as described earlier. The results obtained identify 'Low air pressure $\delta$ Clogged air filter $\delta$ Engine Failure'. Regarding the fuzzy values the following information is available:

Low air pressure $\delta_{0.8}$ Engine Failure .........................................................(x)

Low air pressure $\delta_{0.9}$ Clogged air filter ......................................................... (y)

The unified ontology needs to find the fuzzy value of 'Clogged air filter $\delta$ Engine failure' (assume $\beta$). As explained in section 5,

Low air pressure $\delta_{0.8}$ Engine failure = (Low air Pressure $\delta_{0.9}$ Clogged air filter) $\rightarrow$ (Clogged air filter $\delta_{\beta}$ Engine failure).
Using the Luasiewicz implication function: 

\[ 0.8 = \min (1, 1-0.9 + \beta) = \min (1, 0.1+\beta), \]

clearly, \( \beta = 0.7 \) as shown in the 25. Using this process all other fuzzy values for the fuzzy relation can be defined.

![Diagram](image)

**Figure 25: Fuzzy merged knowledge**

**CONCLUSION:**

This paper addresses issues related to combining new and existing knowledge in the form of an ontology. The advantage of the proposed method is that enterprises need not reconstruct their existing ontology to accommodate newly acquired knowledge. This methodology merges the new knowledge into an existing ontology using ontology merging and reconfiguration. It also checks for any inconsistencies. The process of merging and reconfiguration first identifies the similarity between the concepts of two ontologies and then, with the help of reasoning, identifies the positions where the new concepts will fit in to the existing ontology without any inconsistencies. This approach not only considers the concrete domain but also the fuzzy domain. The proposed approach can also be used in any ontology based knowledge base with different application domains such as product data management (PDM), product lifecycle management (PLM)\textsuperscript{12} and product development\textsuperscript{10}. OWL representation of knowledge
has been shown by many researchers and in all such cases new knowledge can be merged in the existing database in the form of an ontology, build in the same domain, using this approach.

REFERENCES:


