

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

**ANSI***net*

Asian Network for Scientific Information  
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

## Composing Semantic Relations among Ontologies with a Description Logics

Manel Kolli and Zizette Boufaida

Department of Computer Science, LIRE Laboratory, Mentouri University of Constantine 25000, Algeria

**Abstract:** This study presents a formal solution based on the techniques of description logics to solve the problem of ontology matching process which consists of detecting the semantic mappings between two given ontologies. Composing the relations between their subsumers does the deduction of the relations among the semantic entities of different ontologies. It also, presents some performed experiments to assess the effectiveness of the composition operation and to verify that the proposed solution returns complete and precise results. The results obtained look very promising, comparing with expert mappings.

**Key words:** Ontology matching, semantic entities, semantic web, mappings, formal alignment, subsumption

### INTRODUCTION

The semantic Web is considered as the new vision to the web that tries to give semantic to the web resources (Djaghloul and Boufaida, 2007). Since it is based on a logical foundation, the web semantic and its language OWL provide a good solution to semantic problems and ontology ouration (Djaghloul and Boufaida, 2005). With this Web, many related but heterogeneous ontologies are being created. In such an open context, there is no reason why two domain-related applications would share the same ontologies (David and Euzenat, 2008). In order to facilitate the exchange of knowledge between such applications, it must discover the set of semantic relations among entities from these ontologies. The ontology matching process consists of finding the semantic mappings between two given ontologies.

Despite its pervasiveness, today ontology matching is still contains a fair amount of errors or only covers a small part of the ontologies involved (Caracciolo *et al.*, 2008). According to Schorlemmer *et al.* (2007) part of the problem is due to the fact that the majority of work in ontology mapping or database schema matching is based on techniques that use syntactic and structural features of ontologies (Noy and Musen, 2003). The emphasis is on automation, scalability and (re-)use of alignment algorithms but there is an apparent lack of formal foundations for most of this work.

In the literature, there are very few studies about formal matching. In the study of Tang *et al.* (2006) ontology mapping is formalized as a problem of decision making. In this way, discovery of optimal mapping is cast as finding the decision with minimal risk. Sotnykova *et al.* (2005) uses spatial relation algebra for the purpose of expressing correspondences between spatio-temporal ontologies but they do not consider using disjunction of

relations as alignment relations. Smiljanic *et al.* (2005) suggests to transform the XML schema matching problem into a constraint optimization problem. Constraint optimisation problems are constraint satisfaction problems whose solutions maximise a quantity. Though the authors do not consider algebras of relations, this study suggests to use them, since Allens constraint propagation algorithm (Allen, 1983) is an instance of a constraint propagation algorithm (which can be generalized as arc-consistency). In database schema matching, the notion of mapping composition is prominent and has been thoroughly investigated (Madhavan and Halevy, 2003). The problem is to design a composition operator that guarantees that the successive application of two mappings yields the same results as the application of their composition (Fagin *et al.*, 2005). The approach is relatively different since relations in this context are always subsumption ones and their applications involve manipulating the ontology language instead of the alignment language (here the alignment relations). It is even shown that in general the result of composition may require a stronger language than the alignment language (hence the actual result of the composition is not an alignment). The approach taken with composition in algebra of relations is weaker-results are always, in the alignment language-but is not complete.

In response to the above challenge, Kolli and Boufaida (2010) have developed a new formal approach for the ontology matching problem. This approach is based on the techniques of description logics and the semantic distance inspired by Cullot and Jouanot (2003) and adapted to the description logics by Kolli and Boufaida (2004) for the detection of the semantic relations among the semantic entities of two different ontologies.

In the proposed approach, the ontologies are expressed as Description Logics (DLs) knowledge bases. It is well known that, the semantic entities are organized according to the mechanism of subsumption in their ontologies. This organization implies the existence of these entities in the intentional definitions of their subsume. Thus, the detection of the semantic relations among semantic entities can be done through the combination of the semantic relations of their subsumers which is already deduced before. This combination can be realised by the composition of the relations between their subsumers. The results obtained by these operations are validated by the DLs mechanisms. This allows having complete and precise results. This approach returns a generic and extensible ontology of the semantic relations among the semantic entities (ontology of alignment).

This composition is realized in this study by suggesting a set of formal rules. The application of these rules on the semantic entities allows avoiding the calculation of the measure of the semantic distance to every liver because they exploit the mechanisms of the description logics reasoning. In order to estimate the effectiveness of the proposed approach, this study aims at implementing the composition operation by OWL API language.

### BASICS OF DESCRIPTION LOGICS

**Definition 1 (interpretation notion):** Let us consider  $C$  and  $D$  two concepts. An interpretation  $I = (\Delta, \cdot^I)$  consists of a set  $\Delta$  (the domain of  $I$ ) and a function  $\cdot^I$  (the interpretation function of  $I$ ) that maps every concept to a subset of  $\Delta^I$  such that:

- $\emptyset^I = \emptyset$
- $(C \cap D)^I = C^I \cap D^I$
- $(C \cup D)^I = C^I \cup D^I$
- $(\neg C)^I = \Delta^I \setminus C^I$

**Definition 2 (subsumption, equivalence, disjunction and overlapping):**

- A concept  $D$  is subsumed by a concept  $C$  (respectively  $C$  subsumes  $D$ ) which is denoted by  $D \sqsubseteq C$  (respectively  $C \supseteq D$ ) if and only if  $D^I \subseteq C^I, (\forall I)$
- The concepts  $C$  and  $D$  are equivalent which is denoted by  $C = D$  if and only if  $C^I = D^I, (\forall I)$
- A concept  $C$  is disjoint from a concept  $D$  which is denoted by  $D \perp C$  if and only if  $C^I \cap D^I = \emptyset, (\forall I)$
- The concepts  $C$  and  $D$  are overlapped which is denoted by  $C$  and  $D$  if and only if  $C^I \cap D^I \neq \emptyset, (\forall I)$

### ONTOLOGY MATCHING

This study elaborates on the work proposed by Kolli and Boufaïda (2009) which developed a new formal solution to the matching problem based on CBR mechanism for the detection of semantic relations among the semantic entities. Present system consists of retrieving the source cases from the case-base (ontology of alignment) the cases which contain the subsumers of the concerned concepts. Then, it provides a solution to the target case from the solution of the selected source cases which are adapted in order to satisfy the constraints of the posed problem which represents here the detection of the semantic relation between two entities (Fig. 1).

For present system, a case is composed of the following two elements:

- **The problem:** It is to determine the semantic relation between two concepts  $C_i$  and  $C_j$  which belong to two different ontologies
- **The solution:** It is the detection of the semantic relation between these two concepts. We restrict the semantic relation  $rel$  to be one of the relations from the set:  $\{=, \perp, \text{and}, \subseteq, \supseteq\}$ . Intuitively,  $rel(C_i, C_j) = \supseteq$  means that  $C_i$  is more general than  $C_j$ . Thus, we can say that  $C_i$  is the subsumer of  $C_j$ .  $rel(C_i, C_j) = \subseteq$  means that  $C_i$  is less general than  $C_j$ . So,  $C_i$  is the subsumee of  $C_j$ .  $rel(C_i, C_j) = =$  means that  $C_i$  is equivalent to the  $C_j$ . The equivalence relation is defined as the subsumption in both directions.  $rel(C_i, C_j) = \text{and}$  means that  $C_i$  is overlaid with  $C_j$ . Finally,  $rel(C_i, C_j) = \perp$  means that  $C_i$  is disjoint from  $C_j$ .

The cycle of the CBR that we have used is divided into five operations: The elaboration, the retrieving, the reuse, the revision and the memorization. The adaptation of the retrieved cases is done in the reuse step which consists of combining the solutions of the retrieved cases and adapting them in order to find a solution to the target case.

In this step, we rely on the model of adaptation proposed by D'Aquin *et al.* (2006). This model is based on the techniques of Knowledge Discovery from Databases (KDD). The objective of the KDD is to obtain knowledge from data. A KDD session usually relies on two main steps: Data preparation and data-mining (Fig. 1).

Data preparation step aims at formatting and filtering data. The formatting operations transform the data into an acceptable format to the chosen data-mining operations. The filtering operations are used to eliminate noisy data and to concentrate the data-mining operations on a relevant subset of concepts.

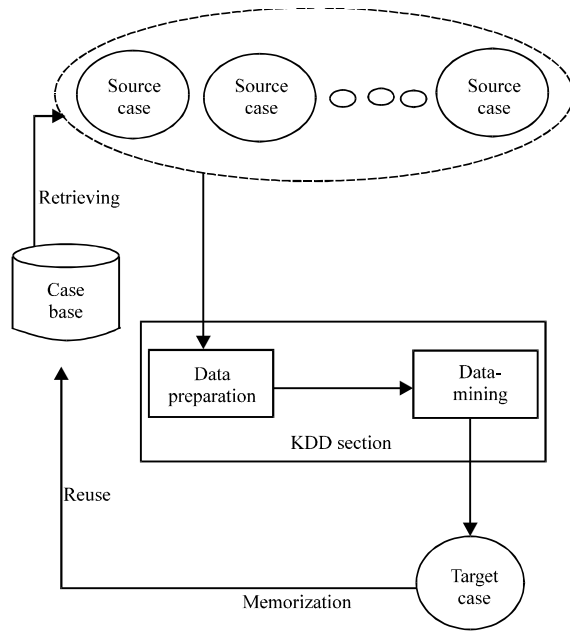


Fig. 1: The detection of the semantic relations in ontologies with a CBR system

Data-mining step allows extracting elements of information from the data. We realize this extraction by applying the Adaptation Knowledge (AK) which we have suggested in the previous work in a formal way by proposing a new formal approach to infer the semantic relations between concepts. As we have already said, in this approach, the deduction of the relations among the semantic entities of different ontologies is done by composing or aggregating the relations between their subsumers.

The result of the comparison is a gradual increase in size of the case-base which demonstrates the need of the organization and the maintenance of the case-base throughout the system life (Kolli and Boufaida, 2010). To reply to this need, we use the ontology notion whose role is to model this kind of knowledge. This generic and extensible ontology is called ontology of alignment  $O_A$ . The formalism that we use for the description of this ontology is the description logics. The set of the deduced relations represents the ontology alignment.

**Composition:** In the Semantic Web infrastructure that rests on the DLs for the ontology construction, we note that the semantic entities are organized according to the subsumption relation which allows the deduction of subsumption relations ( $\sqsubseteq$ ,  $\supseteq$ ) among the concepts of the same ontology in an easy and direct way.

This observation permits us to say that the reuse of the stored case in the ontology  $O_A$  can be deduced by the

composition operation. Thus, if there exists a semantic relation between concept  $C_1$  of ontology  $O_1$  and concept  $C_2$  of ontology  $O_2$  and another semantic relation which is the subsumption relation between concept  $C_2$  and concept  $C_3$  of the same ontology ( $O_2$ ), then it should be possible to obtain the semantic relation between the concepts  $C_1$  and  $C_3$ .

To do this, we have developed a set of formal rules allowing this composition in order to infer the semantic relations among semantic entities. The composition rules are presented as:

- Rule 1:** If the two concepts  $C_1$  and  $C_2$  are disjoint then their subsumee  $C_3$  and  $C_4$  are disjoint too. For example: The concepts man and woman are disjoint then their subsumee boy and girl are disjoint
- Rule 2:** If the two concepts  $C_1$  and  $C_2$  are disjoint then the concept  $C_1$  is disjoint from  $C_4$  the subsumee of the concept  $C_2$  and the concept  $C_2$  is disjoint from  $C_3$  the subsumee of the concept  $C_1$
- Rule 3:** If the concept  $C_1$  is equal to the concept  $C_2$  and this latter subsumes the concept  $C_3$  then  $C_1$  also subsumes the concept  $C_3$
- Rule 4:** If the concept  $C_1$  is subsumed by the concept  $C_2$  and this latter subsumes the concept  $C_3$  then  $C_1$  also subsumes the concept  $C_3$
- Rule 5:** If the two concepts  $C_1$  and  $C_2$  are overlapped then the concept  $C_1$  is overlapped with  $C_4$  the subsumee of the concept  $C_2$  and the concept  $C_2$  is overlapped with  $C_3$  the subsumee of the concept  $C_1$

It is easy to proof these rules by the interpretation notion of DLs as follows:

- (If  $(C_1 \perp C_2)$  then  $(C_1^I \neq C_2^I)$  and if  $(C_3 \sqsubseteq C_1)$  then  $(C_3^I \sqsubseteq C_1^I)$  and if  $(C_4 \sqsubseteq C_2)$  then  $(C_4^I \sqsubseteq C_2^I)$ ) this implies that  $(C_3^I \neq C_4^I)$ . So, we can deduce that:  $(C_4 \perp C_3)$ ,  $(\forall I)$
- (If  $(C_1 \perp C_2)$  then  $(C_1^I \neq C_2^I)$  and if  $(C_4 \sqsubseteq C_2)$  then  $(C_4^I \sqsubseteq C_2^I)$ ) this implies that  $(C_1^I \neq C_4^I)$ . So, we can deduce that:  $(C_1 \perp C_4)$ ,  $(\forall I)$
- In the same way, we can proof that  $(C_2 \perp C_3)$
- (If  $(C_1 = C_2)$  then  $(C_1^I = C_2^I)$  and if  $(C_2 \sqsubseteq C_3)$  then  $(C_2^I \sqsubseteq C_3^I)$ ) this implies that  $(C_1^I \sqsubseteq C_3^I)$ . So, we can deduce that:  $(C_1 \sqsubseteq C_3)$ ,  $(\forall I)$
- (If  $(C_1 \sqsubseteq C_2)$  then  $(C_1^I \sqsubseteq C_2^I)$  and if  $(C_2 \sqsubseteq C_3)$  then  $(C_2^I \sqsubseteq C_3^I)$ ) this implies that  $(C_1^I \sqsubseteq C_3^I)$ . So, we can deduce that:  $(C_1 \sqsubseteq C_3)$ ,  $(\forall I)$
- (If  $(C_1 \text{ and } C_2)$  then  $(C_1 \cap C_2)^I \neq \emptyset$  and if  $(C_3 \supseteq C_1)$  then  $(C_3^I \supseteq C_1^I)$ ) this implies that  $(C_2 \cap C_3)^I \neq \emptyset$ . So, we can deduce that:  $(C_2 \text{ and } C_3)$ ,  $(\forall I)$
- In the same way, we can proof that  $(C_1 \text{ and } C_4)$

### IMPLEMENTATION AND EXAMPLE

To validate the composition operation, we carried out an implementation on two ontologies  $O_1$  and  $O_2$  from the same domain to be able to represent the special cases. Figure 2 and 3 represent the two ontologies that we have chosen. These ontologies are implemented with the well known ontology editor: Protege2000 version 4.1\_beta.

Present system must be used with known cases in order to instantiate the case-base (ontology of alignment  $O_A$ ) and provide a reasoning base. To do this, we take some different entities from the ontologies  $O_1$  and  $O_2$  which are: The concepts person, male and young from  $O_1$  and the concepts human, female and adult from  $O_2$ . We note that person is equivalent to human, male is disjoint from female and adult is disjoint from young. So, these three cases are implemented with Protégé 2000 (Fig. 3).

**Prototype:** The prototype that we have implemented follows the following algorithm:

Algorithm 1: Composition of semantic relations

Data:

1.  $O_1$  and  $O_2$ : two ontologies to align
2.  $O_A$ : initial case base

Results:  $O_A$ : ontology of alignment

```

1. Begin 2
   /*Parse all classes of the ontology  $O_1$ */
2. forall (cls1  $\in$   $O_1$ ) do {
   /* read the subsumers of cls1*/
3. forall (sup1? super_class(cls1)) do {
   /*Parse all classes of the ontology  $O_2$ */
4. forall (cls2  $\in$   $O_2$ ) do {
   /* read the subsumers of cls2 */
5. forall (sup2  $\in$  super_class(cls2)) do {
6. get the relation between sup1 and sup2 from  $O_A$ 
7. if (sup1 is equivalent to sup2) then {
   /* Add cls1 as subsumed by sup2 and cls2 as subsumed by sup1 to  $O_A$ */
8. Add ((cls1, sup2),  $\subseteq$ ),  $O_A$ )
9. Add ((cls2, sup1),  $\subseteq$ ),  $O_A$ )
10. }
11. Else {
12. If (sup1 is disjoint from sup2) then {
   /* cls1 and cls2 are disjoint*/
13. Add ((cls1, cls2),  $\perp$ ),  $O_A$ )
14. }
15. Else {
16. If (sup1 is subsumed by sup2) then {
   /* cls1 is subsumed by sup2*/
17. Add ((cls1, sup2),  $\subseteq$ ),  $O_A$ )
18. }
19. Else {
20. If (sup2 is subsumed by sup1) then {
   /*cls2 is subsumed by sup1*/
21. Add ((cls2, sup1),  $\subseteq$ ),  $O_A$ )
22. }
23. return ( $O_A$ )
24. }}}}
25. End
    
```

The algorithm 1 allows composing the semantic relations among concepts to deduce new semantic relations between other concepts of two ontologies. It takes as input two ontologies  $O_1$  and  $O_2$  to align, described in DLs and the initial alignment ontology  $O_A$ . It returns the final alignment ontology  $O_A$ . It takes two concepts  $cls_1$  and  $cls_2$  belonging, respectively to the ontologies  $O_1$  and  $O_2$  (c.f., lines 2 and 4 of algorithm 1). It extracts their subsumers respectively  $sup_1$  and  $sup_2$  (c.f., lines 3 and 5). Then, depending on the relations between  $sup_1$  and  $sup_2$  specified in the alignment ontology  $O_A$  it deduces the relation between  $cls_1$  and  $cls_2$  (c.f., lines 6 to 22). For example,  $sup_1$  is equivalent to  $sup_2$  then the rule 3 is used in order to infer the semantic relation between  $cls_1$  and  $cls_2$ . Finally, it adds the resulting case to the ontology  $O_A$  (c.f., lines 8, 9, 13, 17 and 21).

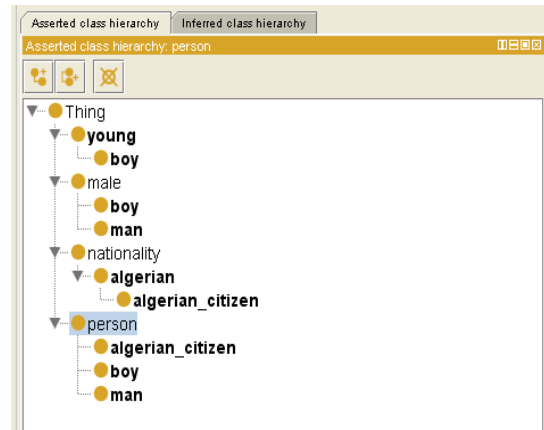


Fig. 2: Concepts hierarchy of the ontology  $O_1$  with Protégé 2000



Fig. 3: Concepts hierarchy of the ontology  $O_2$  with Protégé 2000

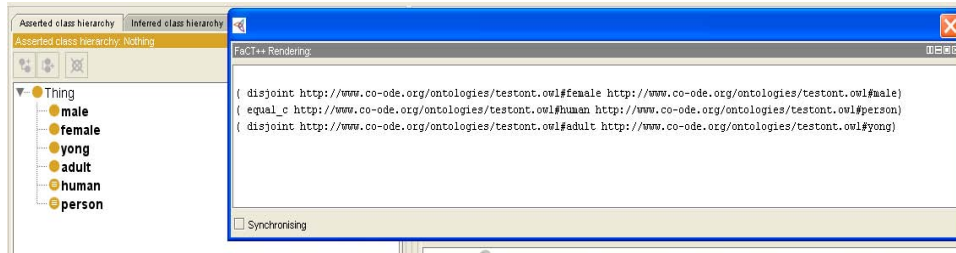


Fig. 4: Concepts hierarchy of the ontology OA with Protégé 2000

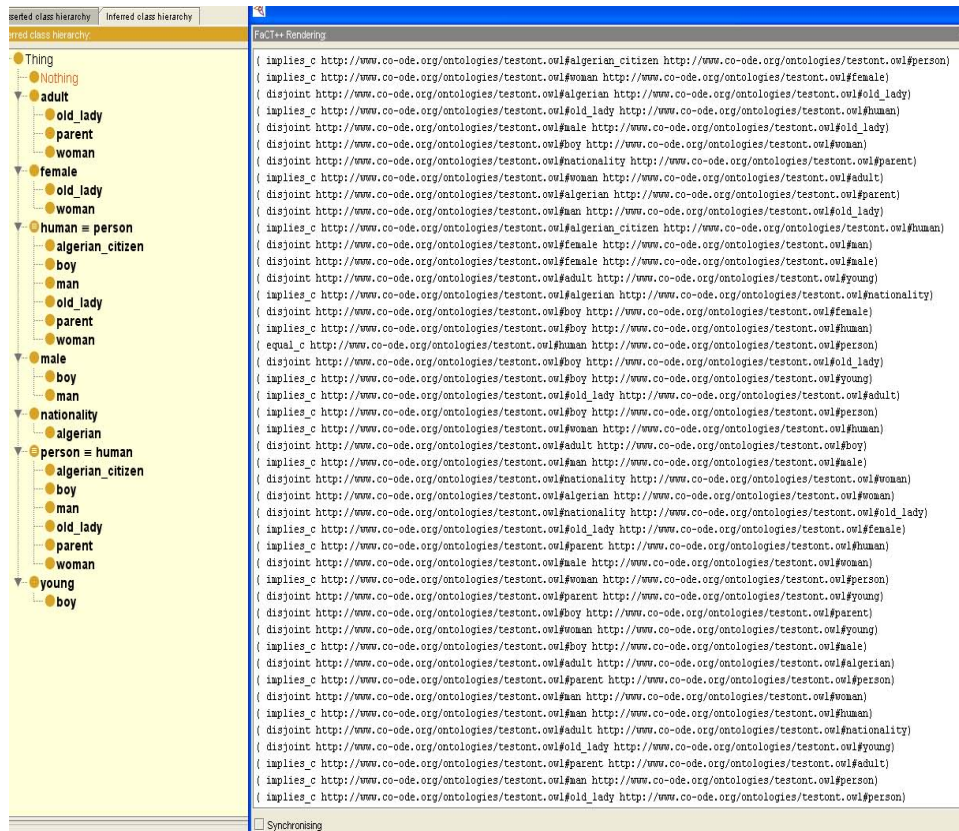


Fig. 5: Concepts hierarchy of the ontology OA resulting after the execution of the composition program

In the previous example, the concepts: Man, the concept person and the concepts subsume boy and Algerian-citizen: Kid, parent, old-lady and woman are subsumed by the concept human. According to the ontology  $O_A$  (Fig. 4), the concepts person and human are equivalents. Indeed, the subsumes of the concept person are the subsumes of the concept human and vice versa (rule 3). Thus, the concept person subsumes the concepts: Kid, parent, old-lady and woman and the

concept human subsumes the concepts: Man, boy and Algerian-citizen.

The results of the composition program are memorized in the ontology of alignment  $O_A$ . Thus, after executing the composition program in present example, the ontology  $O_A$  will be as follows (Fig. 5).

**Verification:** For what concerns the testing approach, to provide a ground for evaluating the quality of match

results, all the pairs of semantic entities have been manually matched to produce expert mappings.

The left window in Fig. 5 shows the cases of mappings deducted after the execution of present composition program. These results have been compared with the expert mappings. We can say that present ontology is valid since all the cases are identical to the expert mappings.

The verification of the consistency of our alignment ontology  $O_A$  was carried out using the reasoner FACT<sup>++</sup> available on protégé 2000. The application of this reasoner allows detecting the smallest anomalies and inconsistencies in the hierarchy of the ontology.

FACT<sup>++</sup> is a classifier, it allows the classification of concepts according to their definition based on description logics. Indeed, it also allows the classification test. As shown in the right window in Fig. 5, the classes hierarchy of  $O_A$  inferred by FACT<sup>++</sup> is identical to the hierarchy of the concepts of  $O_A$  after the execution of our program. This comparison, though preliminary, shows the efficiency and effectiveness of our approach.

## CONCLUSION

In this study, we have presented in detail, composing the semantic relations between the subsumers of the semantic entities in order to detect the semantic relations among them. In order to experiment and verify the effectiveness of this process we have implemented it by the OWL API language that allows us to loaded, manipulated and queried the ontologies.

We have tested our current implementation on two state real-life ontologies and the results look very promising, comparing with expert mappings. This is only a first step and many issues still need to be dealt with.

Notice that in some cases, the composition operation generates a fuzzy results ( $\approx$ ). We can take for example the case where we have:  $(C_1 \supset C_2)$  and  $(C_2 \subset C_3)$ . This implies that  $(C_1^1 \supset C_2^1)$  and  $(C_2^1 \subset C_3^1)$ . So, we can deduce that:  $(C_1^1 = C_3^1)$  or  $(C_1^1 \supset C_3^1)$  or  $(C_1^1 \subset C_3^1)$ . In this case, the relation between  $C_1$  and  $C_3$  can be one of the relations from the set:  $\{=, \supset, \subset\}$ . Thus, the application of the composition operation for these cases cannot solves the problem of relations' deduction. For the purpose of detecting all the possible relations between all concepts, we have developed another operation which is the similarity aggregation (Kolli and Boufaida, 2010). So, our future investigation aims certainly at implementing this operation.

## REFERENCES

- Allen, J.F., 1983. Maintaining knowledge about temporal intervals. *ACM Commun.*, 26: 832-843.
- Caracciolo, C., J. Euzenat, L. Hollink, R. Ichise and A. Isaac *et al.*, 2008. Results of the ontology alignment evaluation initiative 2008. *Proceedings of the Ontology Matching Workshop*, Oct. 26-30, Karlsruhe, Germany, pp: 73-119.
- Cullot, N. and F. Jouanot, 2003. Context comparison for object fusion. *Proceedings of the 15th International Conference on Advanced Information Systems Engineering*, June 16-20, University of Klagenfurt, Klagenfurt, Austria, pp: 536-551.
- D'Aquin, M., F. Badra, S. Lafrogne, J. Lieber, A. Napoli and L. Szathmary, 2006. Adaptation knowledge discovery from a case base. *Proceedings of the 17th European Conference on Artificial Intelligence*, Aug. 28-Sept. 1, Trento, Italy, pp: 795-796.
- David, J. and J. Euzenat, 2008. On fixing semantic alignment evaluation measures. *Proceedings of the 7th International Semantic Web Conference*, Oct. 26-30, Springer, Karlsruhe, pp: 25-36.
- Djaghloul, Y. and Z. Boufaida, 2005. Toward peer to peer platform integration based on OWL ontology and roaming service. *Int. Rev. Comput. Software*, 1: 31-42.
- Djaghloul, Y. and Z. Boufaida, 2007. Building OWL ontology to enhance the peer profile in platforms integration case. *Int. J. Soft Comput.*, 2: 48-60.
- Fagin, R., P.G. Kolaitis, L. Popa and W.C. Tan, 2005. Composing schema mappings: Second-order dependencies to the rescue. *ACM Trans. Database Syst.*, 4: 994-1005.
- Kolli, M. and Z. Boufaida, 2004. A method for the integration of ontologies in the semantic web. *Proceedings of the International Arab Conference on Information Technology*, Dec. 12-15, Constantine, Algeria, pp: 665-665.
- Kolli, M. and Z. Boufaida, 2009. Detecting semantic relations among ontologies with a CBR system. *Proceedings of the IADIS International Conference on Applied Computing*, Nov. 19-21, Rome, Italy, pp: 361-368.
- Kolli, M. and Z. Boufaida, 2010. A description logics formalization for the ontology matching. *Proc. Comput. Sci.*, 3: 29-35.
- Madhavan, J. and A. Halevy, 2003. Composing mappings among data sources. *Proceedings of the 29th International Conference on Very Large Data Bases*, Oct. 2, Berlin, Germany, pp: 572-583.

- Noy, N.F. and M.A. Musen, 2003. The PROMPT suite: Interactive tools for ontology merging and mapping. *Int. J. Human-Comput. Stud.*, 59: 983-1024.
- Schorlemmer, M., Y. Kalfoglou and M. Atencia, 2007. A formal foundation for ontology-alignment interaction models. *Int. J. Semantic Web Inform. Syst.*, 3: 47-65.
- Smiljanic, M., M. van Keulen and W. Jonker, 2005. Formalizing the XML schema matching problem as a constraint optimization problem. *Proceedings of the 16th International Conference on Database and Expert Systems Applications*, Aug. 22-26, Copenhagen, Denmark, pp: 333-342.
- Sotnykova, A., C. Vangenot, N. Cullot, N. Bennacer and M.A. Aufaure, 2005. Semantic mappings in description logics for spatio-temporal database schema integration. *J. Data Semantics*, 3: 143-167.
- Tang, J., J. Li, B. Liang, X. Huang, Y. Li and K. Wang, 2006. Using bayesian decision for ontology mapping. *J. Web Semantics*, 4: 243-262.