Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 83 (2016) 648 – 652

Procedia Computer Science

The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016)

Inferred ontology concepts alignment using instances and an external dictionary

Saida Gherbi[a], Mohamed Tarek Khadir[b]

[a]LabGed, EPST, Annaba and 23000, Algeria
[b]LabGed,University Badji Mokhatar Annaba, Annaba and 23000, Algeria

Abstract

In this paper, a novel approach using Description Logic (DL) based inference rules, for ontology matching is presented. Alignment concerns ontology concepts, with the application of similarity measures to perform concepts and instances relationship alignments. Moreover, external knowledge, in the form of WordNet dictionary is then used to solve usual matching problems encountered with synonyms, polesemy, homonyms, etc. Illustrative examples are then presented to support the developed approach.

Peer-review under responsibility of the Conference Program Chairs

Keywords: reasoner; ontology matching; description logics; OWL; WordNet.

1. Introduction

The representation of Ontology is used by many applications to represent a given domain knowledge, such as: semantic web services, database integration, peer-to-peer systems, social networks, etc.[1]. However, in evolving systems such as the semantic web, different parties would, in general, adopt diverse ontologies[2]. Before being able to combine similar ontologies, a semantic and structural mapping between them has to be established. The process of establishing such a mapping is called ontology alignment[3].

Matching ontologies will become a cornerstone in the realisation of the semantic web vision, and several automatic or semi-automatic ontology alignment tools have been proposed e.g.[2 4]. In the literature there are several ontology matching methods, and most of them are established on similarity measures between the entities to assess the alignment sets for the ontology matching system[5], for instance Coma++[6]. The value of these measures, often determines the similar/dissimilar entities of the matched ontologies. In other words, these measures define just the equivalence and disjunction relations, which do not address on ontology matching issues, such as interoperability or data inte-

* Saida Gherbi. Tel.: +213662955545.
E-mail address: gharbi@labged.net

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.04.145

gration. In the present paper, the focus is put on the discovery of the equivalence/disjunction relations as well as the subsumption relations between the concepts of ontologies to align.

This paper is organized as follows: Section 2 defines the ontology and Description logic (DL) used in this paper. Section 3 discusses the matching process, where Section 4 describes the first level of the proposed approach. The paper will then be ended by a conclusion and some perspectives describing future work directions.

2. Premilinaries

Definition 1. (Ontology) The ontology is defined here as the tuple: $O_i := (D, A, KB, Lex)$ where D represents the core ontology, A the L-axiom system, KB the knowledge base, and Lex is the lexicon used. Other otology descriptions may be found in[7] but in order to describe the different techniques handled for the matching task, the definition of ontology given by the Karlsruhe Ontology Model[8] seems to be the most advised.

Definition 2. (Description Logic DL) The Description Logic languages is considered as the core of knowledge representation systems, viewing both the structure of DL knowledge base and its associated reasoning services[9]. The knowledge base of DL is expressed by a pair $< T, A >$, where T is a terminological box (TBox), a finite set containing the definition of concepts and roles. The concepts definition is expressed by a terminological axioms of the forms $C_1 \subseteq C_2, C_1 \supseteq C_2, C_1 \equiv C_2$, or $C_1 \perp C_2$, where C_1, C_2 are atomic concepts. Furthermore, A the assertional knowledge (ABox) describes individuals by naming and specifying them to its concepts and roles. Several ABoxes may be associated with a same Tbox, as well as the association function used in this paper. DL is characterized also by an interpretation, consisting in a non-empty set $\triangle$ called the interpretation domain, composed of individuals set, expressed here as instances sets I_i and an interpretation function assigning to each atomic concept A, a set of individuals $A^I \subseteq \triangle$, as well as to each atomic binary relation B, a sets of pairs of individuals $B^I \subseteq \triangle x \triangle$[9].

3. Ontology Matching

The matching process expresses an alignment of two ontologies[10] O_1 and O_2 . The Alignment methods require the assessment of the similarity and/or the relation among the concepts and between the relations of ontologies to align. These concepts C and the relations B can be presented as a structure $D := (C, <_C, F, B, <_B)$ of O, where the concept hierarchy or taxonomy is represented by a partial order $<_c$ on C, correspond to set-theoretic relations $Rel = \{\equiv, \subseteq, \supseteq, \perp\}$. The function signature $F : B \longrightarrow CxC$ restricts the model to binary relations, where $F(B_1) = \{dom(B_1), rang(B_1)\}$, for $B_1 \in B$, $dom(B_1)$ symbolized the domain and range $ran(B_1)$, which is treated as an instances of the concepts in the first level and a concept in the second one. The relation hierarchy defined by a partial order $<_B$ on B as:

$$B_1 <_B B_2 \quad Iff \quad (dom(B_1) <_C dom(B_2) \quad and \quad rang(B_1) <_C rang(B_2)) \tag{1}$$

In order to discover the relation between these concepts and binary relations, we start by comparing their instances, for allowing grounding during this operation[2]. These instances (if they exist) are expressed as a structure $KB := (C, B, I, i_C, i_B)$; where the sets C and B as presented before; I is a set of instances, $i_C : C \longrightarrow 2^T$[11] is the association function, associate every concept such as C_1 and $C_1^{'}$ to its instances in I; as well as $i_{B_1} : B_1 \subseteq 2^T$, with $B_1 \subseteq i_{C_1^{'}} x i_{C_1}$ for all $B_1 \in B$. Afterwards, the terminological methods are used to compare the names of instances (relationships and entities in below sections) presented with $Lex := (S_C, S_B, S_I)$. Where, the identifier A_i denotes the three sets S_C, S_B, S_I, express respectively the names of instances, relations, and concepts. This identifier is associated to an axiom by an associate function named x in L-axiom $A := (A_i, x)$.

To illustrate the proposed alignment process two ontologies O_1 and O_2 describing *Human* and *Person*, shown as graphical hierarchies in Fig. 1. are presented. Rectangular boxes indicate concepts, the octagons design properties, the instances are depicted as ellipse and the hierarchy relations as solid arrows. The incoming arrow of relation comes from its domain and an outgoing arrow to its range. Alignments are represented by dotted angle connectors.

4. Level 1 of the proposed alignment algorithm

In this level, we first compare the instances of concepts to deduce the relations among them. After, from these relations, we will infer other relations and align the binary relations.