SEMANTICS BASED ON DISTRIBUTED INTERPRETATION FOR ONTOLOGY INTEGRATION

1KUN LIU, 1SHENGQUN TANG, 2WEIFENG PAN
1State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, Hubei, China
2School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hanzhou,310018, Zhejiang, China

ABSTRACT

Under the background of information integration, the semantics of ontology integration is still an open issue. In this paper, we propose a revised distributed interpretation which is adapted from distributed description logics. In our proposal, ontology integration is taken as global ontology and local ontologies connected by ontology mapping. They are respectively interpreted with description logic semantics and semantic import semantics. Due to the latter, consistency checking can be focused on global ontology and the mapped relations from local ontologies to global ontology. In this way, our method can facilitate understanding and maintenance of ontology integration.

Keywords: Ontology Integration, Distributed Interpretation, Description Logic, Ontology mapping

1. INTRODUCTION

Ontology has played a great role in resolving the heterogeneous of information[1][2][3]. Generally, ontologies are mapped to an upper ontology in order to communication among information sources. These works mainly pay attention to how to find more exactly matched concepts or roles and handle queries on heterogeneous information and seldom discuss its semantics. So how to interpret ontology integration is still an open issue.

Naturally, ontology integration which comprises a set of ontologies is viewed as a whole large one and given a global interpretation. This method makes it easy to reuse classical description logic reasoner to check satisfiability and make further diagnoses[3]. But it will encounter the sufficiency of reasoning existing in large ontology. Fahad[4] talked about semantics of merged ontologies. However, according to [5], there are some differences between ontology merging and ontology integration so that it is not suitable to directly use semantics of ontology merging.

In this paper, we propose a method constructing semantics based on distributed interpretation for ontology integration. We differentiate ontology integration as global ontology and local ontology. A set of interpretations explain semantics of each local ontology and global ontology respectively and use semantic import to interpret mapped relations between them. Our approach makes a distinction between ontologies integrated and mapping relations among them. It can make understanding and maintenance of ontology integration more convenient. Due to semantic import, we can check consistency and satisfiability on global ontology and mapped relations. This decreases the amount and avoids reasoning on large ontology. Our approach can apply to the situation in which each local ontology is only mapped to global ontology and no relations exist between local ontologies. If there is any, our approach is invalid.

The rest of paper is organized as follows: Section 2 explains the motivation of our method. Section 3 talks about some preliminary knowledge of description logics. Section 4 introduces our proposed method. In section 5, we make some discussion with our method. Section 6 introduces some related work. We make conclusions in section 7.

2. MOTIVATION

In [7][8], Borgida and Serafini proposed DDL to describe integration information system(IS). Each IS is described with DL, and then relations between each IS are turned into bridge rules between DL ontologies. Flow of information from one IS to another one can also be described by bridge rules. This situation is applied to information integration. But bridge rules employ correspondence to describe relations of concepts or roles between two DL ontologies. It is not DL tradition which uses subsumption expressions or something else. And more, limited to bridge rules, DDL takes one
ontology as another one’s context, so its reasoning uses context reasoning method\[9\]. In the case of information integration, query answering is usually handled at global side and queries are translated into ones over local ends. Context reasoning is not well suited to query decomposition.

In fact, global ontology and local ontologies in information integration have relations and should not be separated. This means some concepts or roles have direct relations between global ontology and local ontologies. It is similar with some cases of ontology reuse, but ontologies on each side do not include any syntactic symbols from the other side. This inspires us to introduce semantic import\[10\] to improve this situation.

3. PRELIMINARIES

In this section, we will introduce some contents about description logics(DL) and its model semantics.

Generally, ontology can be viewed as a set of DL axioms. According to the expressivity, DL is divided into several types, such as $ALC$, $SHIQ$, $SHOIN(D)$ etc. The difference between them is amount of the constructors each of them adopts.

Basically, $ALC$ provides the minimal set of constructors including $C \odot D$ (conjunction), $C \odot \neg D$ (disjunction), $\neg C$ (negation), $\forall r.C$ (universal qualification) and $\exists r.C$ (existential qualification).

For $ALC$, interpretation $I$ is given on a pair $(\mathcal{A}, \mathcal{J})$. $\mathcal{A}$ is a set which is not empty. $\mathcal{J}$ is a function which maps each class or role to a subset of $\mathcal{A}$ as following:

\begin{enumerate}
  \item $\bullet \mathcal{J} = C'$,
  \item $D' = \emptyset$,
  \item $C \odot D = C' \cap D'$,
  \item $(C \odot D)' = C' \cup D'$,
  \item $\neg C' = \mathcal{A} \setminus C'$,
  \item $\exists r.C' = \{x \in \mathcal{A} | (x, y) \in \mathcal{J} \land y \in C'\}$,
  \item $\forall r.C' = \{x \in \mathcal{A} | (x, y) \in \mathcal{J} \rightarrow y \in C'\}$.
\end{enumerate}

A DL knowledge base comprises TBox and ABox which include asserted axioms. Here we only talk about TBox. If a given interpretation $I$ can satisfy $C' \subseteq D'$, then $C \odot D$ is satisfied by $I$, also marked as $I \models C \odot D$. If $I$ can satisfy $C' = D'$, then $C \equiv D$ is satisfied. It means $I \models C \equiv D$. If every axiom like $C \odot D$ or $C \equiv D$ is satisfied, this DL TBox $T$ is consistent. Sometimes $I$ is called a model of $T$.

4. OUR PROPOSED APPROACH

4.1 Ontology Mapping And Ontology Integration

In our approach, we adopt hybrid ontologies to describe information integration \[1\], but we make a little modification. This method uses local ontologies to represent information sources and global ontology to integrate all local ontologies, as showed in figure 1.

![Hybrid Ontologies](image)

There are mappings between global ontology and local ontologies which relate concept or roles. But different from Wache\[11\], there no relations among local ontologies. Then global ontology can serve as bridges or connections between local ontologies.

Hybrid ontologies can also be denoted by $T = < T_b, \{T_i\}_{i \in \mathcal{I}}, \{B_{ij}\}_{i \in \mathcal{I}} >$. $T_b$ means global ontology and $T_i$ represents each local ontology. $B_{ij}$ shows that a local ontology $T_i$ has mapping relations with global ontology $T_b$.

Global ontology and local ontologies have concepts or roles relations. In DDL, their relations are represented by bridge rules. We choose bridge rules to represent mapping relations from local ontologies to global ontology.

From a syntactic point of view, mappings do not appear on global or local side. It lists concepts or roles names and their relation type.

**Definition 1** (concept mapping): A concept mapping from ontology $T_i$ to $T_j$ is denoted by $h_{ij}$.

It includes three types of relations:

\begin{enumerate}
  \item equation: $i:C \rightarrow j : D$
  \item into: $i:C \rightarrow j : D$
  \item disjointness: $i:C \rightarrow j : D$
\end{enumerate}
Similarly, the definition of roles mapping is
given in definition 2, but for roles mapping, there is
no disjointness relations.

Definition 2 (role mapping): A role mapping
from ontology $R_i$ to $R_j$ is denoted by $b_{ij}$ which
includes two types of relations:

(i) equation: $i:R \rightarrow j:S$

(ii) into: $i:R \rightarrow j:S$

The notation $\rightarrow$ is borrowed from DDL. Some
ontology matching tools can help to find equation
or into relation between concepts.

Example 1: there are two ontologies. One is
University and the other is UNIV. They have the
following concept and role mappings which are
expressed using bridge rule as following.

$b_1 : University : Professor \rightarrow UNIV : Professor$

$b_2 : University : Course \rightarrow UNIV : Course$

$b_3 : University : teach \rightarrow UNIV : teacherOf$

In this example, ontology University is mapped
to UNIV. Professor and Course in University has a
corresponding concept in UNIV and role teach is
mapped into teacherOf in UNIV. Then,

$T = \{ T_{UNIV} , \{ T_{University} \} , \{ b_1 , b_2 , b_3 \} \}$

In the situation of information integration, global
ontology is the center and its semantics is the most
important. All other local ontologies’ semantics
should conform to it.

For ontology integration, mapping relations point
from local ontologies to global ontology. Under this
situation, the semantics of global ontology should
maintain and its consistency should not violate.
Three types of relations can be turned into the
following axioms:

(i) $i: C = j: D$ or $i: R = j: D$

(ii) $i: C \equ D$ or $i: R \equ D$

We call these expressions cross concepts or roles
inclusions. In this situation, concepts or roles from
different ontologies can make inclusions directly.
For example, $i: C$ is directly used in axioms (i), just
like $C = D$ in an ontology, but the prefix $i$ or $j$
should be listed showing which ontology this class
or role belongs to.

4.2 Interpretation

Global interpretation is a usual way to expain the
semantics of DL ontology. But it is not suitable for
ontology integration.

1) All related ontologies will be united as one
ontology. If there are many ontologies to be
integrated, the efficiency of reasoning will be an
issue.

2) Resolving inconsistency among so many
ontologies will be another problem.

3) It is not easy to maintain. Under global
interpretation, all ontologies should be stored in one
file or one physics place, it is difficult to read or
maintenance so large ontology.

When ontology $T_i$ is mapped to another
ontology $T_j$, from the point view of ontology $T_j$
, concepts or roles defined in ontology $T_i$ which are
mapped to $T_j$ can be seen ontology reuse by $T_j$.
It means that ontology $T_j$ reuse some concepts or
roles from ontology $T_i$. And more, under the
background of ontology integration, these reused
concepts or roles should change their semantics to
conform to ontology $T_j$.

Concretely, a distributed interpretation which is
denoted by $I$ comprises a set of interpretation $I =<\{d_i\} | i \in I_i, I_j>$. For each $T_i$, an interpretation $I_i$ should consider
two aspects[10]:

(1) if $i: C$ is a class name in $T_i$, then $i: C\equiv (i: C) \cap \Delta^\theta$.

This case is showed in figure 2. It means that a
class defined in an ontology should conform to the
semantics of the ontology which it is mapped to. Thus it may differ from its original semantics.

![Figure 2: Interpretation For A Mapped Class](image)

(2) if $i: R$ is a role name in $T_i$, then for all $d \in \Delta^\theta \cap \Delta^\theta$, for all $d' \in \Delta^\theta$, $<d,d'> \in R^\theta$, iff $<d,d'> \in (i: R)\cap\theta$.

This case is showed in figure 3. It means that a
role defined in an ontology should conform to the
semantics of the ontology which it is mapped to. Thus it may differ from its original semantics.

![Figure 3: Interpretation For A Mapped Role](image)
Figure 3 shows how the semantics of a role changes when it is mapped to another one.

4.3 Satisfiability And Consistency

Similar to classical description logics, the definition of satisfiability is based on interpretation. But for ontology integration, satisfiability focuses on mapping between global ontology and local ontologies.

**Definition 3 (Satisfiability):** A distributed interpretation \( I = \langle \{ I_i \} i \in I, I_g \rangle \) satisfies \( T \), when the following conditions are satisfied:

1. \( I_i \cup \{ B_i \}_{i \in I} \), if \( I_i \cup T_g \) and \( I_i \cup \{ B_i \}_{i \in I} \);
2. \( I_i \cup T_g \), if \( I_i \cup T_g \);
3. for every concept axiom in \( B_{ig} \),
   \[ (i : C) \equiv \exists (g : D) \]
   \[ (i : C) \sqcap \exists (g : D) = \emptyset \]
4. for every role axiom in \( B_{ig} \),
   \[ (i : R) \equiv \exists (g : S) \]
   \[ (i : R) \sqcap \exists (g : S) = \emptyset \]

This interpretation is also called a model of \( T \).

\( I_i \) in this definition is not same as the \( I_i \) in \( I \).

Based on the definition of satisfiability, consistency is given in definition 4.

**Definition 4 (Consistency):** for an ontology integration \( T \), if a model can be found to satisfy \( T \), then \( T \) is consistent, or else it is inconsistent.

**Example 2:** In ontology \( T_1 \), there are an axiom which states that penguin is a bird. There are another axiom which says fling is disjunction with notFlying in \( T_2 \). And there exists two bridge rules from \( T_1 \) to \( T_2 \) as following.

\( T_1: \ Penguin \delta \ Bird \)
\( T_2: \ Flying \delta \lnot \NotFlying \)
\( B : 1: Penguin \delta 2: \NotFlying \)
\( 1: Bird \delta 2: Flying \)

If \( T_1, T_2 \) and \( B \) are put into one ontology, we can draw an conclusion \( 1: Penguin \delta 2: \NotFlying \) which conflicts with \( 1: Penguin \delta 2: NotFlying \). But according to our proposed distributed interpretation, although class Penguin and Bird are mapped into \( B \) and has semantics given by \( T_2 \), the axiom \( Penguin \delta Bird \) is not admitted by \( T_2 \) and will not affect the semantics of \( T_2 \). So it is consistent.

4.4 Consistency Checking Algorithm

Based on the definition of consistency of ontology integration, we give the following consistency checking algorithm.

**Table 1: Consistency Checking Algorithm**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.</td>
<td>Get each ontology including global ontology from ( T ) and record as ( O = { T_i } )</td>
</tr>
<tr>
<td>02.</td>
<td>for each ( T_i ) in ( O )</td>
</tr>
<tr>
<td>03.</td>
<td>{</td>
</tr>
<tr>
<td>04.</td>
<td>Check the consistency of ( T_i )</td>
</tr>
<tr>
<td>05.</td>
<td>if (( T_i ) is not consistent) return false;</td>
</tr>
<tr>
<td>06.</td>
<td>Get all mappings from all local ontologies to global ontology and record as ( B = { B_{ig} } );</td>
</tr>
<tr>
<td>07.</td>
<td>Combine ( B ) and global ontology ( T_g );</td>
</tr>
<tr>
<td>08.</td>
<td>Check the consistency of ( T_g \cup B );</td>
</tr>
<tr>
<td>09.</td>
<td>if(( T_g \cup B ) is not consistent) return false;</td>
</tr>
<tr>
<td></td>
<td>return true;</td>
</tr>
</tbody>
</table>

Firstly, it checks consistency of global ontology and each local ontology separately. We should separate them from \( T \) (1) and then, it will invoke the classical DL reasoner, such as Pellet, Racer Pro etc to perform this work (2-5).

Secondly, we retrieve mapping relations from global ontology and each local ontology (6). Generally, these relations are translated from the results of ontology matching tools, such as Alignment API etc. Whatever the format of matching result is, they all will be expressed with OWL axioms which is stated aforementioned. Then, these mapping relations can be handled by DL reasoners.

Thirdly, these mapping relations are combined with global ontology (7) and tested whether they are consistent or not (8). This results from the semantics proposed in our approach. According to
the definition, when a class or role is mapped into another ontology, its meaning is subjected to that ontology. Its original relations with other concepts can not be transferred to that ontology. For example, in example 2, the axiom $\text{Penguin} \subseteq \text{Bird}$ defined in $T_1$ is not admitted by $T_2$. So from the point view of this algorithm, ontology integration is consistent.

5. EXAMPLE AND DISCUSSION

We show ontology integration with an example adapted from [11]. In this example which is showed in figure 4, there are four ontologies. One is UNIV which serves as global ontology and the others are local ontologies which named as University, College and Publication. Some of their concepts or roles are mapped to ontology UNIV. It states that In figure 4, only mapped roles are showed. For example, from University to UNIV, there are some relations like the following:

1. mapped concepts:
   - $\text{University} : \text{Person} \rightarrow \text{UNIV} : \text{Person}$
   - $\text{University} : \text{Professor} \rightarrow \text{UNIV} : \text{Professor}$
   - $\text{University} : \text{Student} \rightarrow \text{UNIV} : \text{Student}$
   - $\text{University} : \text{Course} \rightarrow \text{UNIV} : \text{Course}$

2. mapped roles:
   - $\text{University} : \text{teach} \rightarrow \text{UNIV} : \text{teacherOf}$
   - $\text{University} : \text{degreeFrom} \rightarrow \text{UNIV} : \text{degreeFrom}$

Global interpretation is a nature way to explain ontology integration which means that integrated ontologies are seen as a whole ontology. In this way, many existing methods and tools can be reused. But, if there are so many heterogeneous ontologies and so much dissimilarity, it is hard work to check consistency and repair inconsistency, especially when the amount of ontologies reaches a degree.

Compared to global interpretation, distributed interpretation has some advantages. It conforms to the status of scattered ontologies. For example, as to the case of ontology integration in figure 4, four ontologies will be checked to test whether they are consistent or not. But depending on distributed interpretation, there is no need to check all these four ontologies at the same time. It first check the consistency of ontology UNIV, University, College and Publication respectively and then check the combination of UNIV and those concepts and roles which are mapped from University, College and Publication to UNIV.

Our method receive some inspirations from DDL which first propose distributed interpretation, but DDL limits ontology mapping only between two ontologies and cross concept or role inclusions are not allowed, such as $\text{University} : \text{Person} \subseteq \text{UNIV} : \text{Person}$. It doesn’t consider the need of combination. In our method, ontology mapping connects distributed ontologies and class or roles subsumptions across ontologies are permitted.

DDL explains the semantics of ontology mapping with domain relation, but it is not available for ontology mapping expressed with cross inclusions in our approach. We take cross concepts inclusions as ontology reuse and use semantic import to interpret its semantics which differs from DDL much.

Our method still operates when ontology integration overlaps which means global ontology acts as local ontology in another ontology integration. But DDL can not interpret it.

6. RELATED WORK

Jimenez-Ruiz and Grau[4] base their work on global interpretation. They propose a framework named ContentMap to check and repair consistency which fully makes use of existing ontology debugging technology.

Similar to ontology integration, Fahad[5] talks about semantics of ontology merge and algorithms of check consistency. As Flouris[6] have talked
about, there are some differences between ontology integration and ontology merge. So ontology integration cannot fully adopt methods proposed in [5].

In [7][8], DDL is discussed. Borgida and Serafini introduce distributed description logics to express ontologies connected with ontology mappings which called bridge rules. They use domain relation to interpret bridge rules. Serafini [9][12] continues to design reasoning algorithms of DDL.

Connected ontology servers as a bridge between local ontology and shard ontology. Grau[13] talks about linking ontologies with $\varepsilon$-connections which uses a set of new OWL constructs.

Bao[14][15] discusses modular ontologies and its semantics. At some extent, we can conceive of ontology integration as some modular ontologies linked together. But some of its features are not suitable for ontology integration.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we talk about the semantics for ontology integration under the background of information integration. Distributed interpretation is used to explain its semantics. In our method, a set of interpretations is used to explain the semantics of global ontology and local ontologies respectively. The mappings between them which are expressed by bridge rules are interpreted with semantic import. Based on distributed interpretation, consistency checking algorithm is given on global ontology and mapped relation. As long as these bridge rules can be satisfied by distributed interpretation, ontology integration is consistent. We use an example to compare our method with method of global interpretation and prove that distributed interpretation is more suitable for ontology integration. But it is limited when local ontologies have relations each other.

Nowadays, our approach mainly focuses on constructing the proper theory of distributed interpretation for ontology integration. In the future, we will make some research on how to link the theory with OWL and put our approach into practice.

ACKNOWLEDGEMENTS

This work is supported by Zhejiang Provincial Natural Science Foundation of China under grant No.LQ12F02011.

REFERENCES:


