Towards Distributed Ontologies with Description Logics

Martin Homola

Comenius University, Bratislava, Slovakia
homola@fmph.uniba.sk

1 Introduction

Mainstream research on Description Logics (DLs) usually treats DL knowledge bases as monolithic structures (cf. [1]). It has been pointed out, however, that in environments such as Semantic Web, distribution of ontological knowledge across various sources is expected and accepted [2]. Several use-cases can be provided for the envisioned distributed ontology environment:

- Two distinct applications may use two different ontologies to refer to the same concept. Thanks to mapping between these ontologies, association of these concepts can be derived.
- One particular application (e.g., a semantic annotation of a document) can use different ontologies to refer to two distinct concepts. These concepts, may be related by mapping, and so further semantic consequences can be derived.
- Developers may choose to map from a foreign ontology instead of repeating a complex description of some concept. Reuse of concepts can be facilitated.

As Kalfoglou et al. conclude in [3], "…ontology mapping nowadays faces some of the challenges we were facing ten years ago when the ontology field was at its infancy. We still do not understand completely the issues involved." This suggests that accomplishing these scenarios is a long-term and incremental task. In our research, we focus on DLs, a formalism with precise, logical semantics, with encouraging computational properties and practical reasoner implementations [1]. Also, the most prominent ontology language suggested by W3C, OWL Web Ontology Language [4], is derived form DLs.

From the DL point of view, the actual line of research suggests itself:

1. Describe useful syntactic constructs and intuitions behind them.
2. Provide formal model-theoretic semantics for these constructs, yielding a logic with reasoning tasks such as satisfiability of concepts and entailment of concept subsumption.
3. Provide reasoning algorithms for the decision tasks.
4. Develop and optimize implementations of reasoners.

Moreover, each time different set of constructs is put together in Step 1, the following steps need to be repeated, in order to investigate properties and practical usability of thus constructed DL.

In Step 1, for a start, the existing research on ontology mapping (see [3] for a survey) can provide us with intuitions and suggest syntactic constructs for ontology combination. In Step 2, we shall craft formal frameworks that allow combinations of DL knowledge bases using the syntactic constructs selected in Step 1. Reasoning algorithms and implementations follow in Steps 3 and 4.

There are several existing approaches. Distributed Description Logics (DDLs) of Borgida, Serafini and Tamilin [5–7], a framework in which concepts of DL knowledge bases are associated by so called bridge rules, thus allowing for inter-ontology subsumption. Grau et. al in [8] combine DL knowledge bases with $\mathcal{E}$-connections [9], thus allowing for links – inter-ontology role relationships.

It is also noted in [8] that DDLs expose unintuitive behaviour in some situations as demonstrated therein (see below). We find the idea of inter-ontology subsumption appealing. We have analyzed the problem mentioned in [8] and found that it can be "fixed". We summarize our results below in this paper.

2 Distributed Description Logics

A DDL knowledge bases consist of a distributed TBox – a set of local TBoxes, each over its own DL language $\mathcal{L}_i$ – and a set of bridge-rules $\mathfrak{B}$ between these local TBoxes. Each of the local TBoxes $\mathcal{T}_i$ is a collection of GCIs of the form: $i : C \sqsubseteq D$, where the prefix i identifies the TBox the GCI belongs to. Bridge-rules are of two forms, *into* bridge-rules and *onto* bridge-rules:

$$i : A \xrightarrow{\sqsubseteq} j : G \ , \qquad\qquad i : B \xrightarrow{\sqsupseteq} j : H \ .$$

For precise formal semantics of DDLs, please refer to [7]. A particularly attractive feature of DDLs is that they capture the reuse of concepts between several ontologies. This combines well with the basic assumption of Semantic web that no central ontology but many ontologies with redundant knowledge will exist.

In [8], it is noted that certain properties of subsumption relations are not modeled properly by DDL. This problem is demonstrated by the following example that we borrow from [8]. Consider the ontology O:

$$\mathrm{NonFlying} \equiv \neg\mathrm{Flying} \ , \qquad \mathrm{Penguin} \sqsubseteq \mathrm{Bird} \ ,$$
$$\mathrm{Bird} \sqsubseteq \mathrm{Flying} \ , \qquad \mathrm{Penguin} \sqsubseteq \mathrm{NonFlying} \ .$$

And the distributed counterpart of O, divided into two ontologies A (on the left) and B (on the right):

$$\mathrm{NonFlying_A} \equiv \neg\mathrm{Flying_A} \ , \qquad A : \mathrm{Bird_A} \xrightarrow{\sqsupseteq} B : \mathrm{Penguin_B} \ ,$$
$$\mathrm{Bird_A} \sqsubseteq \mathrm{Flying_A} \ . \qquad A : \mathrm{NonFlying_A} \xrightarrow{\sqsupseteq} B : \mathrm{Penguin_B} \ .$$

While Penguin in O is not satisfiable, the corresponding $\mathrm{Penguin_B}$ in B is. The problem is that the DDL framework allows an interpretation to associate each instance x of $\mathrm{Penguin_B}$ with two distinct elements, say y_1 and y_2, one instance of $\mathrm{Bird}_{\mathrm{A}}$ and the other one of $\mathrm{NonFlying}_{\mathrm{A}}$, even if $\mathrm{Bird}_{\mathrm{A}}$ and $\mathrm{NonFlying}_{\mathrm{A}}$ are disjoint. We agree with [8] that it is intuitive to expect that certain relations among concepts of one ontology propagate along bridge rules. So, we would expect $\mathrm{Penguin}_{\mathrm{B}}$ to be unsatisfiable, as it is a "subconcept of two imported concepts" $\mathrm{Bird}_{\mathrm{A}}$ and $\mathrm{NonFlying}_{\mathrm{A}}$ which in their original ontology are disjoint.

3 Our Contribution So Far

We address the problem outlined above by introducing so called conjunctive bridge-rules. We use the following syntax (into and onto form respectively):

$$i:C \overset{\sqsubseteq}{\twoheadrightarrow} j:G \ , \qquad\qquad i:D \overset{\sqsupseteq}{\twoheadrightarrow} j:H \ .$$

Recall from [7], that distributed interpretation is a tuple $\mathfrak{I} = ((\Delta^{\mathcal{I}_1}, \cdot^{\mathcal{I}_1}), \ldots, (\Delta^{\mathcal{I}_n}, \cdot^{\mathcal{I}_n}), r)$ comprising of local interpretations $(\Delta^{\mathcal{I}_i}, \cdot^{\mathcal{I}_i})$ for each local TBox $\mathcal{T}_i$ and $r(\cdot) = \bigcup_{i \neq j} r_{ij}(\cdot)$ interprets the mapping. In addition to the clauses that formally define the semantics of DDL (please refer to [7]), the semantics of conjunctive bridge-rules is given by the following two clauses:

1. $\mathfrak{I} \models_{\mathrm{d}} i:C \overset{\sqsubseteq}{\twoheadrightarrow} j:G$ if for each $i:D \overset{\sqsubseteq}{\twoheadrightarrow} j:H$, $r_{ij}\big(C^{\mathcal{I}_i} \cap D^{\mathcal{I}_i}\big) \subseteq G^{\mathcal{I}_j} \cap H^{\mathcal{I}_j}$,
2. $\mathfrak{I} \models_{\mathrm{d}} i:C \overset{\sqsupseteq}{\twoheadrightarrow} j:G$ if for each $i:D \overset{\sqsupseteq}{\twoheadrightarrow} j:H$, $r_{ij}\big(C^{\mathcal{I}_i} \cap D^{\mathcal{I}_i}\big) \supseteq G^{\mathcal{I}_j} \cap H^{\mathcal{I}_j}$,

where $\mathfrak{I} \models_{\mathrm{d}} R$ is to be read $\mathfrak{I}$ satisfies the bridge rule R.

Our choice of adding new kind of bridge-rules instead of replacing the old semantics by the new one is to underline that both kinds can co-exist and be used according to the intentions of the ontology editor. We continue with characterization of conjunctive bridge-rules. First, they are stronger than the original form in a sense: the semantic condition imposed by a bridge-rule of the original form is also imposed by the corresponding conjunctive form.

Theorem 1. *Given a distributed TBox $\mathfrak{T}$ with a set of bridge-rules $\mathfrak{B}$ and some local TBoxes $\mathcal{T}_i$ and $\mathcal{T}_j$ such that $i \neq j$ and $i:C \overset{\sqsubseteq}{\twoheadrightarrow} j:G \in \mathfrak{B}$ ($i:C \overset{\sqsupseteq}{\twoheadrightarrow} j:G \in \mathfrak{B}$), for each distributed interpretation $\mathfrak{I}$ such that $\mathfrak{I} \models_{\mathrm{d}} \mathfrak{T}$ it holds that $r_{ij}\big(C^{\mathcal{I}_i}\big) \subseteq G^{\mathcal{I}_j}$ ($r_{ij}\big(C^{\mathcal{I}_i}\big) \supseteq G^{\mathcal{I}_j}$) respectively.*

Next theorem shows that choosing conjunctive bridge-rules solves the problem outlined by the example above.

Theorem 2. *Given a distributed TBox $\mathfrak{T}$ with a set of bridge-rules $\mathfrak{B}$ and some local TBoxes $\mathcal{T}_i$ and $\mathcal{T}_j$ such that $i \neq j$, if for some $n > 0$ the bridge-rules $i:C_1 \overset{\sqsubseteq}{\twoheadrightarrow} j:G_1, \ldots, i:C_n \overset{\sqsubseteq}{\twoheadrightarrow} j:G_n$ are all part of $\mathfrak{B}$ then for every distributed interpretation $\mathfrak{I}$ such that $\mathfrak{I} \models_{\mathrm{d}} \mathfrak{T}$ it holds that $r_{ij}\Big((C_1 \sqcap \cdots \sqcap C_n)^{\mathcal{I}_i}\Big) \subseteq (G_1 \sqcap \cdots \sqcap G_n)^{\mathcal{I}_j}$.*

Likewise, if for some $n > 0$ the bridge-rules $i:C_1 \overset{\sqsupseteq}{\twoheadrightarrow} j:G_1, \ldots, i:C_n \overset{\sqsupseteq}{\twoheadrightarrow} j:G_n$ are all part of $\mathfrak{B}$ then for every distributed interpretation $\mathfrak{I}$ such that $\mathfrak{I} \models_{\mathrm{d}} \mathfrak{T}$ it holds that $r_{ij}\Big((C_1 \sqcap \cdots \sqcap C_n)^{\mathcal{I}_i}\Big) \supseteq (G_1 \sqcap \cdots \sqcap G_n)^{\mathcal{I}_j}$.

And so, for concepts involved in conjunctive bridge-rules, the intersection is always "properly related" (subset/superset, w.r.t. the kind of the bridge-rules) to the image of the intersection of their mapped counterparts. In the example above, the concept $\mathrm{Penguin_B}$ is mapped to both $\mathrm{Bird_A}$ and $\mathrm{NonFlying_A}$ which are disjoint. If we would use conjunctive bridge rules in this example, we would have $\mathrm{Penguin_B}^{\mathcal{I}_\mathrm{B}} \cap \mathrm{Penguin_B}^{\mathcal{I}_\mathrm{B}} \subseteq r_{ij}\left(\mathrm{Bird_A}^{\mathcal{I}_\mathrm{A}} \cap \mathrm{NonFlying_A}^{\mathcal{I}_\mathrm{A}}\right)$, which yields $\mathrm{Penguin_B}^{\mathcal{I}_\mathrm{B}} \subseteq r_{ij}(\emptyset)$ and so $\mathrm{Penguin_B}^{\mathcal{I}_\mathrm{B}} \subseteq \emptyset$. That is, in the distributed knowledge base, the concept $\mathrm{Penguin_B}$ is unsatisfiable, according to our intuition.

In [5–7] various desiderata for DDLs are stated. These include:

Monotonicity. Bridge-rules do not delete local subsumptions.

Simple subsumption propagation. Combination of into and onto bridge-rules allows for propagation of subsumption across ontologies. Formally, if $i : C \xrightarrow{\sqsupseteq} j : G \in \mathfrak{B}$ and $i : D \xrightarrow{\sqsubseteq} j : H \in \mathfrak{B}$ then $\mathfrak{T} \models_\mathrm{d} i : C \sqsubseteq D \implies \mathfrak{T} \models_\mathrm{d} j : G \sqsubseteq H$.

Generalized subsumption propagation. If $i : C \xrightarrow{\sqsupseteq} j : G \in \mathfrak{B}$ and $i : D_k \xrightarrow{\sqsubseteq} j : H_k \in \mathfrak{B}$, for $1 \leq k \leq n$ then $\mathfrak{T} \models_\mathrm{d} i : C \sqsubseteq \bigsqcup_{k=1}^{n} D_k \implies \mathfrak{T} \models_\mathrm{d} j : G \sqsubseteq \bigsqcup_{k=1}^{n} H_k$.

These desiderata are satisfied even in the presence of conjunctive bridge-rules. Moreover, subsumption propagates over conjunctive bridge-rules even for intersection of concepts, as follows.

Theorem 3. *If* $i : C \xrightarrow{\sqsupseteq} j : G \in \mathfrak{B}$ *and* $i : D_k \xrightarrow{\sqsubseteq} j : H_k \in \mathfrak{B}$, $1 \leq k \leq n$ *then* $\mathfrak{T} \models_\mathrm{d} i : C \sqsubseteq \bigsqcap_{k=1}^{n} D_k \implies \mathfrak{T} \models_\mathrm{d} j : G \sqsubseteq \bigsqcap_{k=1}^{n} H_k$.

These results have not been published yet. We have submitted a technical paper to DL-2007. Other interesting desiderata for DDLs published in [7] include requirements that local inconsistency does not pollute the entire distributed system and that the information flow along the bridge-rules respects the direction of these bridge-rules (i.e., no information flows the other way). Evaluating conjunctive bridge-rules with respect to these desiderata is subject to our ongoing research.

4 Conclusion and Future Work

We decided to pursue research for our Ph.D. dissertation in the area of Distributed Ontologies based on Description Logics. Several approaches exist in this field, most notably those of [8], exploiting the $\mathcal{E}$-connections framework, and those of [5–7] introducing Distributed Description Logics (DDLs). We find the latter one interesting, since it allows for inter-ontology subsumption – a notion that combines well with the vision of Semantic Web. However, an unintuitive behaviour of DDLs has been demonstrated in [8]. We have found it natural to start our research, concentrating on this problem. We have figured out that the semantics of DDLs can be amended to solve this problem. There are several possibilities to continue with our research:

– Continue with the evaluation of conjunctive bridge-rules (e.g., with respect to the remaining desiderata for DDLs.)
– Devising a reasoning algorithm, studying its computational complexity.
– Implementation and practical evaluation. (Depends on the previous item.)
– Exploiting the possibility to combine DDLs with links used in $\mathcal{E}$-connections.

We intend to address these issues in our Ph.D. research. We will continue with the evaluation of conjunctive bridge-rules first. Also, there is a tableaux algorithm known for the original DDL framework (see [6, 7]), we will start by checking whether it can be adopted for conjunctive bridge-rules effectively. We would like to continue with the complexity analysis next. The last item of the list are very interesting as well.

Acknowledgement. The author is a Ph.D. student in third year, however, he only switched to Distributed Ontologies a year ago. Support from Slovak Agency for Promotion of Research and Development projects, contract no. APVV-20-P04805 and APVV-99-PO5005; and project VEGA of Slovak Ministry of Education and Slovak Academy of Sciences no. 1/0173/03 is acknowledged.

References

1. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., eds.: The Description Logic Handbook. Cambridge University Press (2003)
2. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American **284**(5) (2001) 34–43
3. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: The state of the art. In: Semantic Interoperability and Integration. Number 04391 in Dagstuhl Seminar Proceedings (2005)
4. : OWL web ontology language overview. A W3C Recommendation. Available online, at `http://www.w3.org/TR/owl-features/`
5. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. Journal of Data Semantics **1** (2003) 153–184
6. Serafini, L., Tamilin, A.: Local tableaux for reasoning in distributed description logics. In: Procs. of DL'04. CEUR-WS (2004)
7. Serafini, L., Borgida, A., Tamilin, A.: Aspects of distributed and modular ontology reasoning. In: Procs. of IJCAI'05. (2005) 570–575
8. Cuenca Grau, B., Parsia, B., Sirin, E.: Combining OWL ontologies using $\mathcal{E}$-connections. Journal of Web Semantics **4**(1) (2006) 40–59
9. Kutz, O., Lutz, C., Wolter, F., Zakharyaschev, M.: $\mathcal{E}$-connections of abstract description systems. Artificial Intelligence **156**(1) (2004) 1–73