Three Semantics for the Core of the Distributed Ontology Language (Extended Abstract)*

Till Mossakowski1, Christoph Lange2, Oliver Kutz3

1. Research Centre on Spatial Cognition, University of Bremen and DFKI GmbH Bremen, Germany
2. University of Birmingham, UK
3. Research Centre on Spatial Cognition, University of Bremen, Germany
ontoioop-forum@ontolog.cim3.net

Abstract

The Distributed Ontology Language DOL, currently being standardized as ISO WD 17347 within the OntoIOp (Ontology Integration and Interoperability) activity of ISO/TC 37, provides a unified framework for (1) ontologies formalized in heterogeneous logics, (2) modular ontologies, (3) links between ontologies, and (4) ontology annotation.

A DOL ontology consists of modules formalized in languages such as OWL or Common Logic, serialized in the existing syntaxes of these languages. On top, DOL’s meta level allows for expressing heterogeneous ontologies and links between ontologies, including (heterogeneous) imports and alignments, conservative extensions, and theory interpretations.

We present the abstract syntax of these meta-level constructs, with three alternative semantics: direct, translational, and collapsed semantics.

1 Introduction

A variety of languages is used for formalizing ontologies.1 OWL is widely used but its restriction to a decidable description logic hinders ontology designers from expressing richer knowledge structures. In biomedical ontologies, for example, mereological relations such as parthood are of great importance, but OWL cannot fully express them. Instead of continuing the earlier practice of embedding first-order mereology axioms as informal annotations into OWL ontologies, we aim at a formal semantics for heterogeneous ontologies. Acknowledging the pluralism in ontology languages, we do not aim at yet another “Esperanto” ontology language, but provide a formal semantics to integrate ontologies written in coexisting different formalisms. We particularly support the widely used OWL, the first-order Common Logic (CL) language, but our framework specifies general conformance criteria and is thus open for plugging in any other ontology language with a formal semantics.

2 The Distributed Ontology Language DOL

The Distributed Ontology Language DOL allows users to use their preferred ontology formalism while becoming interoperable with other formalisms. DOL is the core of the OntoIOp ISO standardization effort on ontology interoperability [OntoIOp, 2012]. It builds on a graph of ontology languages and translations, which will enable users to distribute ontologies by (i) relating ontologies written in different formalisms (e.g. stating that the DOLCE Lite OWL ontology is entailed by the reference first-order DOLCE ontology [Masolo et al., 2003]), (ii) re-using ontology modules given in a different formalism, and (iii) re-using tools such as theorem provers and module extractors along translations between formalisms. (iv) DOL uses IRIs for globally unique identification, which allows for distributing ontologies over the Web. On a meta level on top of basic ontology languages, DOL allows for modeling logically heterogeneous ontologies, comprising of modules written in ontology languages with different underlying logics, and for links between ontologies, such as alignments, relative interpretations or conservative extensions.

We are particularly interested in establishing conformance of the following widely used ontology languages and logics with DOL (ordered by increasing complexity): propositional logic, the W3C-standardized OWL 2 [W3C, 2009] (with its profiles, i.e. sublanguages, EL, RL and QL), classical first-order logic with equality (FOL=) and the ISO/IEC standard CL [ISO/IEC, 2007], a variant of first-order logic with an impredicative Lisp-like wild-west syntax, coming in the full variant CL (with sequence markers denoting lists of individuals) and the unofficial sublanguage CL− (without these). Translations between these languages have been developed in previous research [Mossakowski and Kutz, 2011]. Among these languages, CL is most expressive, and therefore can serve as a common translation target. Reasoning about heterogeneous distributed ontologies is possible on demand by translating all modules to first-order logic. This is different from translating them to FOL in the first place, which would no longer give access to optimized language-specific tools decidable/tractable OWL reasoning procedures.

3 An Introductory Example

While mereological relations such as parthood are frequently used in ontologies (including large biomedical ontologies im-
implemented in OWL EL for efficiency), their complete definitions require at least first-order logic and thus exceed the expressiveness of many ontology languages. The following listing shows a heterogeneous mereological ontology, based on an example from [Kutz et al., 2010]:

```plaintext
\prefix\ % IRI prefix for this distributed ontology

.owl: <http://www.example.org/mereology#>

.owl: % OWL

.log: % DOL-conforming logics

.trans: <http://purl.net/dol/logics/>

% translations

.ser: <http://purl.net/dol/serializations/>

% serializations, i.e. concrete syntaxes


distributed-ontology Mereology

% some basics in propositional logic, so we can use SAT solvers for
% efficient and early detection of modelling errors

logic log:Propositional syntax

ontology Taxonomy =

% basic taxonic information about mereology reused from DOLCE

.props PT \,% Particular \%, PD \,% Perdurant \%,

T \,% TimeInterval \%, S \,% SpaceRegion \%, AR \,% AbstractRegion \%

.% S \,\lor\, T \,\lor\, AR \,\lor\, PD \,\rightarrow\, PT

.% S \,\land\, T \,\rightarrow\, \bot

.% PD, S, T, AR are pairwise disjoint

.% T \,\land\, AR \,\rightarrow\, \bot


logic log:SR0IQ syntax

.ontology BasicParthood =

% as far as easily expressible

.Class: Particular Category
.SubClassOf: Particular

.% similar declarations of the other classes omitted

.DisjointUnionOf: SpaceRegion, TimeInterval, AbstractRegion, Perdurant

.% pairwise disjointness more compact thanks to an OWL built-in

.ObjectProperty: isProperPartOf

.Characteristics: Asymmetric

.Class: Atom EquivalenceTo: inverse isProperPartOf only

.owl: Nothing

% The OWL ontology interprets the propositional logic as follows:
% InterpretationToParthood : Taxonomy to BasicParthood =
% with translation trans:PropositionalToSR0IQ,
% % I. translate logic

.T \,\rightarrow\, TimeInterval

.% A \,\rightarrow\, AbstractRegion

.% and so on %

.logic log:CommonLogic syntax

.ontology ClassicalExtensionalParthood =

% import OWL ontology above,

.BasicParthood with translation trans:SR0IQtoCL

.% translate if ...

.then ( % % ... to CL, then extend it there, using second-order
% features to quantify over the taxonomic category predicates:
% forall (Cat) (if (or (= Cat S) (= Cat T) (= Cat AR) (= Cat PD))

.(forall (x y z) (if (and (Cat x) (Cat y) (Cat z))

.and (and now list all the axioms

.(if (and (isProperPartOf x y) \,(isProperPartOf y x))\,(x y)) % antisymmetry

.(if (and (isProperPartOf x y) \,(isProperPartOf y x) \,(isProperPartOf y z)) % gether with asymmetry

.(iff \,(overlaps x y) \,(exists (pt))

.(and (isProperPartOf pt x) \,(isProperPartOf pt y)))

.(iff \,(isProperPartOf x y) \,(isProperPartOf y x) \,(Atom x)))

.(iff \,(sum x y) \,(forall w) \,(iff \,(overlaps w z) \,(exists (s) \,(sum x s))))

.(forall (Set a) \,(iff \,(fusion Set a)) \,(forall b) \,(iff \,(overlaps b a)))

.(exists (c) \,(and (Set c) \,(overlaps c a)))) }

.% 2nd-order fusion

Figure 1: Core of the logic translation graph

---

4 Syntax of DOL

The following exposition of DOL’s syntax and semantics only covers the core meta-logical constructs. A distributed ontology consists of at least one (possibly heterogeneous) ontology, plus, optionally, interpretations between its participating ontologies. More specifically, a distributed ontology consists of a name, followed by a list of sections, parsed and interpreted subject to the ontology language, logic and syntax

5 The Logic Graph

An ontology language translation ρ is either specified by its name, or it is inferred as the default translation between a given source and target ontology language. A link connects two ontologies; possible links include interpretations and alignments. DOL uses IRIs (Internationalized Resource Identifiers, RFC 3987) for names of distributed ontologies, ontologies, logics, translations, links and other objects.

An ontology language translation ρ is either specified by its name, or it is inferred as the default translation between a given source and target ontology language. A link connects two ontologies; possible links include interpretations and alignments. DOL uses IRIs (Internationalized Resource Identifiers, RFC 3987) for names of distributed ontologies, ontologies, logics, translations, links and other objects.

OWL calls its different syntaxes serializations; CL has different dialects, which also have different syntaxes.

---

2OWL calls its different syntaxes serializations; CL has different dialects, which also have different syntaxes.
a $\Sigma_1$-model ($M_2$ is then called an expansion of $M_2|_{\Sigma_1}$). For each logic, it is easy to show that $M_2 \models \varphi$ iff $M_2|_{\Sigma_1} \models \varphi$, i.e., satisfaction is invariant under reduct. Finally, we assume a signature union operation.

Fig. 1 also shows some logic translations, with some of their properties. Such a translation consists of three components: a signature translation $\Phi$, which is expected to map signature extensions to signature extensions, a sentence translation $\alpha$ and a model translation $\beta$ (where models are translated in the reverse direction). Each translation enjoys the following representation condition:

$$\beta(M) \models \varphi \text{ iff } M \models \alpha(\varphi),$$

where $\varphi$ is a sentence in the source of the translation, and $M$ is a model in the target of the translation. Moreover, we require that each model translation is compatible with reduct, i.e., $\beta(M'|_{\Phi(\Sigma_2)}) = \beta(M'|_{\Sigma_2})$ for $\Sigma \subseteq \Sigma'$ and $M' \in \text{Mod}(\Phi(\Sigma'))$. In most cases, model translations will also be surjective. For the translational semantics below, we will need the following notion: We call a translation weakly exact, if for each $M' \in \text{Mod}(\Sigma')$ and $M_1 \in \text{Mod}(\Phi(\Sigma))$ with $\beta(M_1) = M'|_{\Phi(\Sigma)}$, there is a model $M'_1 \in \text{Mod}(\Phi(\Sigma'))$ with $M'_1|_{\Phi(\Sigma)} = M_1$ and $\beta(M'_1) = M'$.

Some of the translations in Fig. 1 are sublogic inclusions (namely the inclusions of the $\text{QL}$ class of $\Sigma$, to $\text{CL}$). The abstract syntax of all basic ontology languages is translated into that of $\text{CL}$, which is then interpreted w.r.t. the CL semantics as in (i). The semantics of the meta level remains semi-formal. (iii) The collapsed semantics is fully specified in CL. It translates the abstract syntax of the meta level to CL and then re-uses the CL semantics. Thus, the meta and object levels collapse into CL but may still be distinguished by a closer look into the CL theory.

The model-theoretic nature of the semantics ensures a better representation of the model theory than a theory-level semantics would do. In particular, Theorem 1 ensures that models classes of logical theories represented in CL can be recovered through a model translation. This is of particular importance when studying model-theoretic properties like finite model or tree model properties.

In the context of a global environment $\Gamma$, which maps IRIs to (semantics of) ontologies, and the currently selected logic $L$, the direct semantics interprets an ontology $O$ as a signature $\Sigma = \text{sig}(\Gamma, L, O)$ in some logic $L' = \text{logic}(\Gamma, L, O)$ and a class of models $M = \text{Mod}(\Gamma, L, O)$ over that signature. We combine this into

$$\text{sem}(\Gamma, L, O) = (\text{logic}(\Gamma, L, O), \text{sig}(\Gamma, L, O), \text{Mod}(\Gamma, L, O)).$$

The direct semantics of a basic ontology is given by its signature and the class of model satisfying its axioms. The semantics of the translation of an ontology along a logic translation uses the signature and model translation components of the logic translation$^3$. A reference to a named ontology is looked up in the global environment and, if needed, translated to the current logic (this kind of implicit coercion is used in the listing in section 3 when including an OWL ontology into a CL ontology). Finally, a logic qualification replaces the current logic with a new one.

<table>
<thead>
<tr>
<th>$O'$</th>
<th>$\text{sem}(\Gamma, L, O')$ = ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \Sigma, \Delta \rangle$</td>
<td>$(\langle \Sigma, \Delta \rangle, \text{Mod}(\Gamma, L, O) = \ldots)$</td>
</tr>
<tr>
<td>$O$</td>
<td>$\text{translate}$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$\rho = (\Phi, \alpha, \beta) : L_1 \rightarrow L_2$. Then $\text{logic}(\Gamma, L, O') = \text{LogicRef}$ and $\text{Mod}(\Gamma, L, O') = {M \in \text{Mod}(\Phi(\Sigma))</td>
</tr>
<tr>
<td>$O$ then $\text{OntoRef}$</td>
<td>$\text{sig}(\Gamma, L, O') = \Sigma \cup \Sigma'$ and $\text{Mod}(\Gamma, L, O') = {M' \in \text{Mod}(\Sigma \cup \Sigma')</td>
</tr>
<tr>
<td>$\text{LogicRef}$</td>
<td>$\text{sem}(\Gamma, \text{LogicRef}, O)$</td>
</tr>
</tbody>
</table>

The semantics of relative interpretations is formulated in terms of model class inclusion:

$$\text{sem}(\Gamma, L, \text{interpretation IntpName} : O_1 \rightarrow O_2)) = \Gamma[\text{IntpName} \rightarrow (\Sigma_1, \Sigma_2)]$$

where $\{L_1, \Sigma_1, M_1\} = \text{sem}(\Gamma, L, O_1)$,

$$(L_2, \Sigma_2, M_2) = \text{sem}(\Gamma, L, O_2),$$

and the semantics is defined only if $\beta(M_2|_{\Sigma_1}) \subseteq M_1$ where $\beta(\Sigma_2|_{\Sigma_1})$ is the default translation from $L_1$ to $L_2$.

With the direct semantics, one can define many standard logical notions in a straightforward way. For example, a heterogeneous ontology is satisfiable if it has a nonempty model class. A sentence $\varphi$ is a logical consequence of a heterogeneous ontology $O$ in context $\Gamma$ and $L$, written $\Gamma, L, O \models \varphi$, if $\varphi$ is a sentence in the logic $\text{logic}(\Gamma, L, O)$, and each model in $\text{Mod}(\Gamma, L, O)$ satisfies $\varphi$.

The translational semantics uses CL as a foundational framework for the distributed ontology language DOL, similar to what set theory provides for general mathematical theories. This semantics assumes that each involved ontology

$^3$A theory-level semantics would use the sentence translation instead of the model translation.
language is mapped to CL by a weakly exact translation. The semantics is defined by first translating a heterogeneous ontology to CL, and then using the direct semantics for the result.

The \textbf{collapsed semantics} requires the representation of the meta level within CL. For this purpose, the direct model-level semantics should be complemented by a theory-level semantics: a distributed ontology then denotes a basic theory in some logic (which amounts to flattening out all structure), plus some conditions for conservativity and relative interpretations. For each logic, one needs to axiomatize a specific partial order of signatures in CL, plus a set of sentences equipped with a logical consequence relation. In order to avoid the formalization of models and the satisfaction relation (which would require the inclusion of a set theory like ZFC), a sound and complete calculus is axiomatized for each logic. For each logic translation, the signature and sentence translations need to be axiomatized. We require that this axiomatization is done in such a way that the resulting semantics is compatible with the translational semantics. Although this formalization is doable in principle, we refrain from providing the (massive) details.

7 Relations Among the Different Semantics

\textbf{Theorem 1 (Compatibility of semantics)} Given an ontology $O$ written in logic $L$, if $O$ does not involve conservative extensions, then the direct semantics and the translational semantics coincide:

\[ \Phi(\Sigma) = \Sigma_{CL} \quad \text{and} \quad M = \beta(M_{CL}) \]

where $(L, \Sigma, M)$ is the direct semantics of $O$, $(CL, \Sigma_{CL}, M_{CL})$ is the translational semantics of $O$, and $\rho = (\Phi, \alpha, \beta)$ is the default translation from $L$ to $CL$.\textsuperscript{4}

While Theorem 1 provides a good compatibility of the first two semantics, there are some differences:

\textbf{Example 2 (Conservative extensions)} Conservative extensions, both in the consequence-theoretic and the model-theoretic variant, play an important role, in particular concerning ontology module extraction, see e.g. [Lutz and Wolter, 2010]. Consider a large ontology like SNOMED CT [Schulz et al., 2009]. For a particular application, typically only a small portion, i.e. a subtheory, of the ontology is needed. However, this subtheory should contain all logical information about its signature that is implied by the whole ontology. This requirement is precisely that of conservative extension. Hence, a module is a subtheory of an ontology such that the ontology is a conservative extension of the module. Consider the following EL theory about lectures and their subjects [Lutz et al., 2007]:

Lecture $\sqsubseteq$ has\_subject.\ Subject $\sqcap$ has\_subject.\ Lecturer

Intro\_AI $\sqsubseteq$ Lecture

\[ \text{This theory is extended as follows:} \]

Intro\_AI $\sqsubseteq$ 2 has\_subject. Logic

Intro\_AI $\sqsubseteq$ 2 has\_subject. NeuralNetworks

Logic $\sqcap$ NeuralNetworks $\sqsubseteq$ $\bot$.

Now this extended theory logically implies that Intro\_AI $\sqsubseteq$ 2 has\_subject; this follows since Logic and NeuralNetworks are disjoint and both related via has\_subject to Intro\_AI. Hence, in OWL, the larger theory is not a consequence-theoretic conservative extension of the smaller one, because Intro\_AI $\sqsubseteq$ 2 has\_subject is a sentence in the signature of the smaller theory that follows from the larger theory, but not from the smaller one. But in EL, such a sentence does not exist. In particular, the number restriction $\geq 2$ has\_subject cannot be expressed in EL.

This example shows that the translational semantics differs from the direct semantics w.r.t reasoning tasks that are syntax sensitive. While the direct semantics would recognize the above extension as consequence-theoretically conservative (because it directly works with EL), the translational semantics would not, because after translating to CL, the conservative extension property is lost. A similar difference arises when computing the least common subsumer of a set of concept descriptions, because this also depends on the expressiveness of the used language.

8 Conclusion

We have presented three different semantics for the Distributed Ontology Language DOL, which are compatible with each other, with some exceptions like conservative extensions. The direct semantics stays close to the semantics of the individual ontology languages such as OWL and CL. The translational semantics is based on a mapping to CL, such that only knowledge of CL is required to understand this semantics. The collapsed semantics formalizes also the meta theory in CL, and thus makes the meta level itself amenable to computer-assisted theorem proving and verification.

A first application of DOL is the (indeed homogeneous) COLORE repository [Grüninger, ], containing more than 400 theories written in CL. So far, the relations between these theories have been stated informally and proved manually. With DOL, these relations can be stated formally. This allows for using theorem provers and model finders for (dis-)proving logical facts about these relations [Lange et al., 2012a].

Future work will include specification of further logics and translations. The direct semantics has a clear advantage here, because the logics can be directly included in the graph, while the translational semantics first requires translation to CL, which may be involved or impossible.

\textbf{Acknowledgments:} We would like to thank the OntoIOp working group within ISO/TC 37/SC 3 for providing valuable feedback, in particular Michael Grüninger, Pat Hayes, Maria Keet, Chris Menzel, and John Sowa. This work has been supported by the DFG-funded Research Centre on Spatial Cognition (SFB/TR 8), project I1-[OntoSpace], and EPSRC grant EP/J007498/1.

\textsuperscript{4}If we take global environments into account, these need to be compatible in an obvious way, see [Mossakowski et al., 2012] for details.

\textsuperscript{5}Note that the $\bot$ concept is part of the $\mathcal{EL}$++ extension of the logic $\mathcal{EL}$ and so available in the OWL profile EL.
References


