Soundness Proof of Z Semantics of OWL Using Institutions

Dorel Lucanu
Faculty of Computer Science
"A.I.Cuza" University
Iaşi, Romania
dlucanu@info.uaic.ro

Yuan Fang Li and Jin Song Dong
School of Computing
National University of Singapore
Singapore
[liyf,dongjs]@comp.nus.edu.sg

ABSTRACT

The correctness of the Z semantics of OWL is the theoretical foundation of using software engineering techniques to verify Web ontologies. As OWL and Z are based on different logical systems, we use institutions to represent their underlying logical systems and use institution morphisms to prove the correctness of the Z semantics for OWL DL.

Categories and Subject Descriptors

F.4 [**MATHEMATICAL LOGIC AND FORMAL LANGUAGES**]: Miscellaneous; I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Representation languages*

General Terms

Languages, Theory, Verification

Keywords

OWL, Z, institution, comorphism of institutions

1. INTRODUCTION

In our previous works [2], we proposed to use software engineering techniques in a combined approach to verify the correctness of Web ontologies. The validity of the combined approach relies on the correctness of the Z semantics of the ontology language. As OWL and Z are based on different logical systems (description logics and first-order logic), the proof of the correctness requires a high-level device that is able to represent and relate different logical systems.

The notion of institutions [4] was introduced to formalize the concept of "logical systems". Institution morphisms [3] captures the migration between logical systems. In this paper, we prove the correctness of the Z semantics [1] for OWL DL using institutions and institution morphisms, by representing the underlying logical systems of OWL DL and Z as institutions and applying institution comorphisms.

2. THE OWL INSTITUTION $\mathfrak{O}$

We recall from [5] the definition of the institution formalizing the logic OWL DL. The OWL institution $\mathfrak{O}$ is given by $\mathfrak{O} = (\mathsf{Sign}(\mathfrak{O}), \mathsf{sen}(\mathfrak{O}), \mathsf{Mod}(\mathfrak{O}), \models_{\mathfrak{O}})$. The definition of $\mathfrak{O}$ follows mainly the lines described in [6]. The use of the institution theory offers several significant advantages: ability to work with structured ontologies, use of constraints to distinguish between OWL DL and OWL Full ontologies, and a solid foundation for tools extending, linking OWL languages with other formalisms similar to those presented in [2].

Briefly, an *OWL signature* consists of a quadruple $\mathcal{O} = (\mathbb{C}, \mathbb{R}, \mathbb{U}, \mathbb{I})$, where $\mathbb{C}$ is the set of *concept (class) names*, $\mathbb{R}$ and $\mathbb{U}$ are the sets of *individual-valued* and *data-valued property names*, respectively, and $\mathbb{I}$ is the set of *individual names*.

Given an OWL signature, an $\mathcal{O}$*-structure (model)* is a tuple $A = (\Delta_A, [\![\_]\!]_A, Res_A, res_A)$ consisting of a set of *resources* Res_A, a subset $\Delta_A \subseteq Res_A$ called *domain*, a function $res_A : \mathcal{N}(\mathcal{O}) \cup \mathbb{D} \to Res_A$ associating a resource with each name in $\mathcal{O}$ or $\mathbb{D}$, and an interpretation function $[\![\_]\!]_A : \mathbb{C} \cup \mathbb{R} \cup \mathbb{U} \to \mathcal{P}(Res) \cup P(Res) \times P(Res)$.

The set of $\mathcal{O}$*-sentences* is defined by:

$$
\begin{aligned}
F ::= &\, \mathcal{C} \sqsubseteq \mathcal{C} \mid \mathcal{C} \equiv \mathcal{C} \mid \mathsf{Disjoint}(\mathcal{C}, \ldots, \mathcal{C}) \\
&\mid \mathsf{Tr}(R) \mid \mathcal{R} \sqsubseteq \mathcal{R} \mid \mathcal{R} \equiv \mathcal{R} \\
&\mid U \sqsubseteq U \mid U \equiv U \\
&\mid o : \mathcal{C} \mid (o, o') : \mathcal{R} \mid (o, v) : U \mid o \equiv o' \mid o \not\equiv o'
\end{aligned}
$$

where o and o' range over individuals names, v ranges over data values, $\mathcal{C}$ ranges over OWL class descriptions and restrictions and U and $\mathcal{R}$ range over datatype- and object-properties, respectively.

The details of the satisfaction relation can be found in [5].

3. THE INSTITUTION $\mathfrak{Z}$

We briefly recall from [1] the institution $\mathfrak{Z}$, formalizing the logic underlying the specification language Z.

A *Z signature* $\mathcal{Z}$ is a triple (G, Op, τ) where G is the set of the *given-set names*, Op is a set of the *identifiers*, and τ is a function mapping the names in Op into types $\mathcal{T}(G)$.

Given a Z signature $\mathcal{Z} = (G, Op, \tau)$, a $\mathcal{Z}$*-structure (model)* is a pair (A_G, A_{Op}) where A_G is a functor from G, viewed as a discrete category, to Set, and A_{Op} is a set $\{(o, v) \mid o \in Op\}$ where $v \in \overline{A}_G(\tau(o))$. The functor $\overline{A}_G$ is the standard extension of A_G to $\overline{A}_G : \mathcal{T}(G) \to \mathsf{Set}$.

Given a Z signature $\mathcal{Z}$, the set of $\mathcal{Z}$*-sentences* P are defined by:

$$
\begin{aligned}
P ::= &\, \mathtt{true} \mid \mathtt{false} \mid E \in E \mid E = E \mid \neg P \mid P \vee P \mid P \wedge P \\
&\mid P \Rightarrow P \mid \forall S.P \mid \exists S.P
\end{aligned}
$$

where E and S represent the sets of $\mathcal{Z}$*-expressions* and $\mathcal{Z}$*-schema-expressions*, respectively.

[1] The semantics can be found at `http://www.comp.nus.edu.sg/~liyf/OWL2Z.tex`

WWW 2005, May 10–14, 2005, Chiba, Japan.
ACM 1-59593-051-5/05/0005.

The details of the Z *environment*, the satisfaction relation and the use of *mathematical toolkit* can be found in [1].

4. ENCODING $\mathfrak{O}$ IN $\mathfrak{Z}$

The main idea is to associate a Z specification $\Phi(\mathcal{O}, F)$ with each OWL specification $(\mathcal{O}, F)$ such that an $(\mathcal{O}, F)$-model can be extracted from each $\Phi(\mathcal{O}, F)$-model. The construction of $\Phi(\mathcal{O}, F)$ is given in two steps: we first associate a Z specification $\Phi(\mathcal{O})$ with each OWL signature $\mathcal{O}$ and then we add to it the sentences F translated via a natural transformation.

Since $\Phi(\mathcal{O}, F)$ can be seen as a Z semantics of $(\mathcal{O}, F)$, it includes a distinct subspecification $(\mathcal{Z}^{\emptyset}, P^{\emptyset})$ defining the main OWL concepts and the operations over sets. More precisely, we consider $(\mathcal{Z}^{\emptyset}, P^{\emptyset})$ as being the vertex of the colimit having as base the standard library, the specification of the data types, together with the Z specification about OWL DL [5].

We define $\Phi^{\diamond} : \mathsf{Sign}(\mathfrak{O}) \to \mathsf{Sign}(\mathfrak{Z})$ as follows. Let $\mathcal{O} = (\mathbb{C}, \mathbb{R}, \mathbb{U}, \mathbb{I})$ be an OWL signature. Then $\Phi^{\diamond}(\mathcal{O}) = (G, Op, \tau)$ is defined as follows:

$$
\begin{aligned}
&G = G^{\emptyset};\\
&Op = Op^{\emptyset} \cup \mathbb{C} \cup \mathbb{R} \cup \mathbb{U} \cup \mathbb{I};\\
&\tau(C) = \texttt{Resource} \text{ for each } C \in \mathbb{C},\\
&\tau(R) = \texttt{Resource} \text{ for each } R \in \mathbb{R},\\
&\tau(U) = \texttt{Resource} \text{ for each } U \in \mathbb{U},\\
&\tau(o) = \texttt{Resource} \text{ for each } o \in \mathbb{I}.
\end{aligned}
$$

If $\varphi : \mathcal{O} \to \mathcal{O}'$ is an OWL signature morphism and $\Phi^{\diamond}(\mathcal{O}) = (G^{\emptyset}, Op, \tau)$ and $\Phi^{\diamond}(\mathcal{O}') = (G^{\emptyset}, Op', \tau')$, then $\Phi^{\diamond}(\varphi) : \Phi(\mathcal{O}) \to \Phi(\mathcal{O}')$ is the Z signature morphism $(\mathrm{id} : G^{\emptyset} \to G^{\emptyset}, \Phi^{\diamond}(\varphi)_{op} : Op \to Op')$ such that $\Phi^{\diamond}(\varphi)_{Op}$ is the identity over the subset $Op^{\emptyset}$ and $\Phi^{\diamond}(\varphi)_{op}(N) = \varphi(N)$ for each name N in $\mathcal{O}$.

We extend $\Phi^{\diamond}$ to $\Phi : \mathsf{Sign}(\mathfrak{O}) \to \mathsf{Th}(\mathfrak{Z})$ by defining $\Phi(\mathcal{O}) = (\Phi^{\diamond}(\mathcal{O}), P)$, where P is $P^{\emptyset}$ together with the following sentences:

$$
\begin{aligned}
&\{C \in \texttt{Class}) \mid C \in \mathbb{C}\} \cup\\
&\{R \in \texttt{ObjectProperty} \mid R \in \mathbb{R}\} \cup\\
&\{U \in \texttt{DatatypeProperty} \mid U \in \mathbb{U}\} \cup\\
&\{o \in \texttt{Individual} \mid o \in \mathbb{I}\}.
\end{aligned}
$$

If $\mathcal{O}$ is an OWL signature, then

$$\alpha_{\mathcal{O}} : \mathsf{sen}(\mathfrak{O})(\mathcal{O}) \to \mathsf{sen}(\mathfrak{Z})(\Phi(\mathcal{O}))$$

is defined by:

$$
\begin{aligned}
&\alpha_{\mathcal{O}}(\bot) = \texttt{Nothing},\ \alpha_{\mathcal{O}}(\top) = \texttt{Thing},\\
&\alpha_{\mathcal{O}}(N) = N \text{ for each name } N \text{ in } \mathcal{O}\\
&\alpha_{\mathcal{O}}(C_1 \sqcap C_2) = \texttt{intersectionOf}(\alpha_{\mathcal{O}}(C_1), \alpha_{\mathcal{O}}(C_2)),\\
&\ldots\\
&\alpha_{\mathcal{O}}(\forall R.C) = \texttt{allValuesFrom}(\alpha_{\mathcal{O}}(R), \alpha_{\mathcal{O}}(C)),\\
&\ldots\\
&\alpha_{\mathcal{O}}(E) = \{\alpha_{\mathcal{O}}(e) \mid e \in E\}.
\end{aligned}
$$

LEMMA 1. *$\alpha = \{\alpha_{\mathcal{O}} \mid \mathcal{O} \in \mathsf{Sign}(\mathfrak{O})\}$ is a natural transformation $\alpha : \mathsf{sen}(\mathfrak{O}) \Rightarrow \Phi^{\diamond};\ \mathsf{sen}(\mathfrak{Z})$.* [2]

If $\mathcal{O} = (\mathbb{C}, \mathbb{R}, \mathbb{U}, \mathbb{I})$ is an OWL signature and $A' = (A'_G, A'_{Op})$ a $\Phi^{\diamond}(\mathcal{O})$-model, then $\beta_{\mathcal{O}}(A')$ is the $\mathcal{O}$-model $A = (\Delta_A, [\![\_]\!]_A, Res_A, res_A)$ defined as follows:

[2] The details of the proofs of this and following lemmas and theorem can be found in [5].

$Res_A = A'_G(\texttt{Resource})$,
$res_A(N) = v$ where $(N, v) \in A'_{Op}$ for each name $N \in \mathcal{O}$,
$\Delta_A = v$ where $(\texttt{Thing}, v) \in A'_{Op}$,
if $C \in \mathbb{C}$, then $[\![C]\!]_A = v_C$ where $(\texttt{instances}, v) \in A'_{Op}$ and $(C, v_C) \in v$,
if $R \in \mathbb{R}$, then $[\![R]\!]_A = v_R$ where $(\texttt{subVal}, v) \in A'_{Op}$ and $(R, v_R) \in v$,
if $U \in \mathbb{U}$, then $[\![U]\!]_A = v_U$ where $(\texttt{subDVal}, v) \in A'_{Op}$ and $(U, v_U) \in v$.

where `instances` and `subVal` are the corresponding expansion functions. A is indeed an $\mathcal{O}$-model. For instance, if $(\texttt{instances}, v) \in A'_{Op}$, then v is the graph of the function defined in A' by `instances` and v_C is just the value of this function for the argument C. Since $\tau^{\emptyset}(\texttt{instances}) = \mathcal{P}(\texttt{Resource} \times \mathcal{P}(\texttt{Resource}))$, it follows that $v_C \subseteq A'_G(\texttt{Resource})$. We obtain $[\![C]\!]_A \subseteq \Delta_A$ applying the sentences in $P^{\emptyset}$. We extend $\beta_{\mathcal{O}}$ to a functor $\beta_{\mathcal{O}} : \mathsf{Mod}'(\Phi^{\diamond}(\mathcal{O})) \to \mathsf{Mod}(\mathcal{O})$ as follows: if $h : A' \to B'$ is a $\Phi^{\diamond}(\mathcal{O})$-homomorphism, then $\beta_{\mathcal{O}}(h)$ is the $\mathcal{O}$-homomorphism $\beta_{\mathcal{O}}(h) : \beta_{\mathcal{O}}(A') \to \beta_{\mathcal{O}}(B')$ given by $\beta_{\mathcal{O}}(h) = h_{\texttt{Resource}}$.

LEMMA 2. *$\beta = \{\beta_{\mathcal{O}} \mid \mathcal{O} \in \mathsf{Sign}(\mathfrak{O})\}$ is a natural transformation $\beta : \Phi^{\diamond\, op};\ \mathsf{Mod}(\mathfrak{Z}) \Rightarrow \mathsf{Mod}(\mathfrak{O})$.*

THEOREM 1. *$(\Phi, \alpha, \beta) : \mathfrak{O} \to \mathfrak{Z}$ is a simple theoroidal comorphism.*

5. CONCLUSION

In this paper, we demonstrated the soundness of the Z semantics for OWL through the use of institution morphisms. This allows us to use Z reasoners for proving properties of OWL ontologies. If e is a property of the OWL ontology $(\mathcal{O}, F)$ and we prove that the Z-encoding of $(\mathcal{O}, F)$ satisfies the translation of e, $\alpha_{\mathcal{O}}(e)$, then $(\mathcal{O}, F)$ satisfies e by the satisfaction condition from the definition of the comorphism.

The data type is a parameter for the OWL institution. This allows to use the built-in Z data types for proving properties of the OWL ontologies. The correctness of the properties is then preserved by the institution morphism induced by the translation of the data type.

The paper exhibits also a practical way to put at work the theoretical results concerning the migration between logical systems.

6. REFERENCES

[1] H. Baumeister. Relating abstract datatypes and Z-schemata. In *Recent Trends in Algebraic Development Techniques - Selected Papers*, volume 1827 of *Lect. Notes in Comput. Sci.*, pages 366–382, Bonas, France, 2000. Springer-Verlag.

[2] J. S. Dong, C. H. Lee, Y. F. Li, and H. Wang. A Combined Approach to Checking Web Ontologies. In *Proceedings of 13th World Wide Web Conference (WWW'04)*, pages 714–722, New York, USA, May 2004.

[3] J. Goguen and G. Roşu. Institution morphisms. Formal Aspects of Computing, 2002.

[4] J. A. Goguen and R. M. Burstall. Introducing institutions. In *Proc. Logics of Programming Workshop*, number 164 in Lect. Notes in Comput. Sci., pages 221–256. Springer-Verlag, 1984.

[5] D. Lucanu, Y. F. Li, and J. S. Dong. Web Ontology Verification and Analysis in the Z Framework. Technical Report TR 05-01, University "Alexandru Ioan Cuza" of Iaşi, Romania, January 2005. `http://thor.info.uaic.ro/~tr/tr05-01.ps`.

[6] P. F. Patel-Schneider and I. Horrocks (editors). OWL: Direct Model-Theoretic Semantics. `http://www.w3.org/TR/owl-semantics/direct.html`.