ORM2 Encoding into Description Logic
(Extended Abstract)

Enrico Franconi[1], Alessandro Mosca[1], and Dmitry Solomakhin[1]

Free University of Bozen-Bolzano, Italy
[1] `lastname@inf.unibz.it`

Abstract. The *Object Role Modelling* (ORM2) is a conceptual modelling approach combining both textual specifications and graphical language, similar to UML and ER, and adopted by Visual Studio, the integrated development environment designed by Microsoft. This paper introduces a new linear syntax and corresponding complete set-theoretic semantics for a generalization of ORM2 language. A core fragment of ORM2 is defined, for which a provably correct encoding into $\mathcal{ALCQI}$ description logic is presented. Based on these results, an extensive and systematic critique of alternative approaches to the formalisation of ORM2 in (description) logics published so far is provided. A first prototype has been implemented, which offers a back-end for the automated support of consistency and entailment checks for ORM2 conceptual schemas along with its translation into $\mathcal{ALCQI}$ knowledge bases.

1 Introduction

Automated support to enterprise modelling has increasingly become a subject of interest for organisations seeking solutions for storage, distribution and analysis of knowledge about business processes [6], and the main expectation from automated solutions built upon these approaches is the ability to automatically determine consistency of a business model. Despite existence of reasoning tools for Unified Modelling Language (UML) [2], its known weakness with regard to verbalisation of facts and constraints restricts its usage by domain experts [10]. Recently becoming popular ORM2 ('Object Role Modelling 2') is a graphical fact-oriented approach for modelling, transforming, and querying business domain information, which allows for a verbalisation in language readily understandable by non-technical users. Being domain expert oriented, the semantics of ORM2 differs from that of UML (e.g. permitting optionality of cardinality constraints) and thus makes ORM2 richer in its capacity to express business constraints [10].

The NIAM language ('Natural-language Information Analysis Method'), ancestor ORM, has been equipped with an FOL-based semantics for the first time in 1989 [9]. Since then, despite the remarkable evolution in terms of expressivity and graphical notation that ORM2 has experienced, much less attention has been paid in the consequent development of appropriate formal foundations for the modelling language.

This paper addresses the main problem of providing a logic formalism, equipped with sound and complete reasoning services, that captures the expressiveness of ORM2. The first contribution of the paper is thus the introduction of a completely new linear syntax and a set-theoretic semantics for ORM2 matching the usage patterns in the community. The new syntax can be used to express the full set of ORM2 graphical symbols introduced in [10]. The second contribution of the paper is driven by a practical objective. On the basis of well known results developed in the Description Logics (DLs) community, we identified a ‘core’ fragment of ORM2 that can be translated in a sound and complete way into the ExpTime-complete logic $\mathcal{ALCQI}$ [1], through *n*-ary relations *reification*. On the basis of the results presented in the paper, a first prototype, built on top of available DL reasoners, has been implemented, which provides an automated support for schema consistency, entity/relations consistency check, and entailment verification for user-defined ORM2 statements.

The rest of the paper is organised as follows: Section 2 is about the introduction, through examples, of the ORM2 graphical notation and intended semantics in the framework of the fact modelling approach; Section 3 introduces the new linear syntax by means of expressing the example from the previous section in this syntax. The encoding of the corresponding set-theoretic semantics into the DL logic $\mathcal{ALCQI}$ is the main topic of Section 4. Finally, Section 5 gives an overview of the implemented reasoning support prototype, its interface and gives an example of its usage.

2 Fact-oriented Modelling in ORM2

Basic ORM2 objects are: **entities** (e.g. a house or a car) and **values** (e.g. *character string* or *number*). Moreover, entities and values are described in terms of the **types** they belong to: a type (e.g. House, Car) is a set of possible instances. In order to avoid ambiguity among the possible instances of a given type, entities are identified also by means of a particular *reference mode* and a value. The roles played by the entities in a given domain are introduced by means of logical **predicates**; each predicate (or relation) has a given set of roles according to its arity. Each role is connected to exactly one object type, indicating that the role is played only by possible instances of that type.

The first step of the ORM2 design procedure thus concerns the specification of the relevant **object types** (i.e. entity and value types), predicates and reference modes. All the subsequent steps in the procedure mostly deal with the specification of static *constraints*. Let us consider the example in Fig. 1:

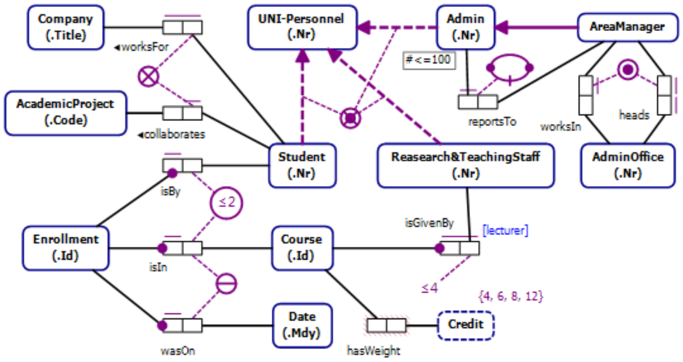


Fig. 1. A conceptual schema including an instantiation of most of the ORM2 constraints

The schema includes:

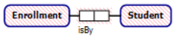
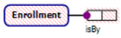
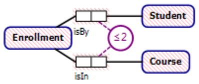
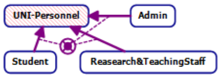
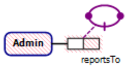
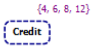
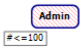
1. **Entity types**: Enrollment, Student, Date, ... ;
2. **Binary predicates**: isBy, wasOn, worksFor, ... ;
3. A user-defined **role name** [lecturer], for the role played by Research&TeachingStaff;
4. **Reference modes** for each entity: .Id, .Nr, .Mdy, ... ;
5. **Subtyping** links (depicted as thick arrows) indicating 'isa' relationships among types, and a constraint combination, called **partition**, made of an **exclusive** constraint (a circled 'X' for *mutual disjointness*), and a **total** constraint (a circled dot for complete coverage of the common super-type).
6. **Internal frequency occurrence** constraint indicating that *if* an instance of Research&TeachingStaff plays the role of being lecturer in the relation isGivenBy, it plays the role at most 4 times.
7. An **external frequency occurrence** that applies to roles played by Student and Course, meaning that 'Students are allowed to enroll in the same course *at most twice*'.
8. An **external uniqueness** constraint between the role played by Course in isIn and the role played by Date in wasOn, saying that 'For each combination of Course and Date, *at most one* Enrollment isIn that Course and wasOn that Date'.
9. A disjunctive mandatory 'circled dot', called **inclusive-or**, linking the roles played by AreaManager indicating that 'Each area manager *either* works in *or* heads (or *both*)'.
10. An **object cardinality** constraint forcing the number of the Admin instances to be less or equal to 100.
11. An **object type value** constraint indicating which values are allowed in Credit.
12. An **exclusion** constraint (depicted as circled 'X') between the roles played by Student in the relations worksFor and collaborates, expressing the fact that no student can play *both* these roles.
13. A **ring** constraint expressing that the relation reportsTo is *asymmetric*.

3 Proposed Formalisation of ORM2

As mentioned before, the modelling activity in ORM2 is supported by several tools that provide user friendly graphical interfaces to build complex conceptual schemas. However, none of the available design tools offers automated reasoning support on specific combinations of ORM2 constraints. The automated verification of *schema consistency* and *consistency of an object type* over a conceptual schema strictly depends on the possibility to perform reasoning and make inferences on it by means of a semantic-based logic representation of the schema itself.

With this goal in mind, we propose a linear syntax that fully covers the set of graphical symbols of ORM2. Table 1, using the example from previous section, shows how a new introduced syntax can be used to encode conceptual schemas that have been originally specified in graphical terms. For each construct ϕ in the syntax, its corresponding set-theoretic semantics expressed in relational algebra is also introduced in table 2 (where O denotes an object type).

Table 1. Constraints C1-C7 below represent a fragment of the schema from Fig. 1.

		ENTITYTYPES:{Enrollement,Student,...} VALUETYPES:{Credit,Student-Nr,...} RELATIONS:{isIn,isBy,collaborates,student-Nr,...}
C1.		TYPE(isBy.enrollment,Enrollment) TYPE(isBy.student,Student)
C2.		MAND({isBy.enrollment},Enrollment)
C3.		FREQ({isBy.student,isIn.course}, {isBy.enrollment,isIn.enrollment},⟨1, 2⟩)
C4.		O-SET$_{\text{Tot}}$({R&TStaff,Student,Admin},UNI-Personnel) O-SET$_{\text{Ex}}$({R&TStaff,Student,Admin},UNI-Personnel)
C5.		RING$_{\text{Asym}}$(reportsTo.sub, reportsTo.obj)
C6.		V-VAL(Credit)={[4,6,8,12]}
C7.		O-CARD(Admin)=(0,100)

The signature $\mathcal{S}$ of the linear ORM2 syntax is made of:

- Disjoint sets $\mathcal{E}$ and $\mathcal{V}$ of *entity type* and *value type* symbols, respectively;
- a set $\mathcal{R}$ of *relation* symbols and a set $\mathcal{A}$ of corresponding *role* symbols;
- a set $\mathcal{D}$ of *domain* symbols, and a set Λ of pairwise disjoint sets of values;
- for each $D \in \mathcal{D}$, an injective extension function $\Lambda_{(\cdot)} : \mathcal{D} \rightarrow \Lambda$ associating each domain symbol D to an extension Λ_D;
- a binary relation $\varrho \subseteq \mathcal{R} \times \mathcal{A}$ linking role symbols to relation symbols. We take the pair $R.a$ as the atomic elements of the syntax, and we call it *localised role*. Given a relation symbol R, $\varrho_R = \{R.a | R.a \in \varrho\}$ is the set of localised roles with respect to R;
- for each relation symbol R, a bijection $\tau_R : \varrho_R \rightarrow [1..|\varrho_R|]$ mapping each element in ϱ_R to an element in the finite sequence of natural numbers $[1..|\varrho_R|]$. The mapping τ_R guarantees a correspondence between role components and argument positions in a relation.

Given the signature $\mathcal{S}$, an **ORM2** ***conceptual schema*** Σ over $\mathcal{S}$ includes a finite combination of the constructs in table 2.

4 Encoding in $\mathcal{ALCQI}$

With the main aim of relying on available reasoning tools to reason in an effective way on ORM2 schemas, we present here the encoding in the logic $\mathcal{ALCQI}$, for which tableaux-based reasoning algorithms with a tractable computational complexity have been developed [8,12]. The $\mathcal{ALCQI}$ encoding has been devised through an intermediate translation step in the logic $\mathcal{DLR}$, where arbitrary n-ary relations are allowed [3].

Table 2. Linear Syntax and Semantics table

Syntax	Semantics
$\mathsf{TYPE} \subseteq \varrho \times (\mathcal{E} \cup \mathcal{V})$	If $\mathsf{TYPE}(R.a, O) \in \Sigma$ then $\Pi_{R.a} R^{\mathcal{I}} \subseteq O^{\mathcal{I}}$
$\mathsf{FREQ} \subseteq \wp(\varrho) \times \wp(\varrho) \times (\mathbb{N} \times (\mathbb{N} \cup \{\infty\}))$	If $\mathsf{FREQ}(\{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}, \{R^1.a, \ldots, R^k.a\}, \langle \mathsf{min}, \mathsf{max} \rangle) \in \Sigma$ then $\Pi_{\varrho^{\mathsf{C}}}(R^{1^{\mathcal{I}}} \underset{R^1.a=R^2.a}{\bowtie} \cdots \underset{R^{k-1}.a=R^k.a}{\bowtie} R^{k^{\mathcal{I}}}) \subseteq \{\overline{x} \mid \mathsf{min} \leq \sharp\{\sigma_{\overline{x}=\varrho^{\mathsf{C}}}(R^{1^{\mathcal{I}}} \underset{R^1.a=R^2.a}{\bowtie} \cdots \underset{R^{k-1}.a=R^k.a}{\bowtie} R^{k^{\mathcal{I}}})\} \leq \mathsf{max}\}$ where $\varrho^{\mathsf{C}} = \{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}$
$\mathsf{MAND} \subseteq \wp(\varrho) \times (\mathcal{E} \cup \mathcal{V})$	If $\mathsf{MAND}(\{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}, O) \in \Sigma$ then $O^{\mathcal{I}} \subseteq \Pi_{R^1.a_1} R^{1^{\mathcal{I}}} \cup \cdots \cup \Pi_{R^1.a_n} R^{1^{\mathcal{I}}} \cup \cdots \cup \Pi_{R^k.a_1} R^{k^{\mathcal{I}}} \cup \cdots \cup \Pi_{R^k.a_m} R^{k^{\mathcal{I}}}$
$\mathsf{R\text{-}SET_H} \subseteq ((\wp(\varrho) \times \wp(\varrho)) \times (\wp(\varrho) \times \wp(\varrho)) \times (\mu\colon \varrho \to \varrho)$ where μ is a partial bijection s.t. for any $\langle \varrho^{\mathsf{C_A}}, \varrho^{\mathsf{C_B}}, \mu \rangle \in \mathsf{R\text{-}SET_H}$ $\varrho^{\mathsf{C}_A} = \{R.a \mid \mu(R.a) \in \varrho^{\mathsf{C}_B}\}$, and $\mathsf{H} = \{\mathsf{Sub}, \mathsf{Exc}\}$	• If $\mathsf{R\text{-}SET_{Sub}}((\{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}, \{R^1.a, \ldots, R^k.a\}),$ $(\{S^1.b_1, \ldots, S^1.b_n, \ldots, S^k.b_1, \ldots, S^k.b_m\}, \{S^1.b, \ldots, S^k.b\}), \mu) \in \Sigma$ then $\Pi_{\varrho^{\mathsf{C_A}}}(R^{1^{\mathcal{I}}} \underset{R^1.a=R^2.a}{\bowtie} \cdots \underset{R^{k-1}.a=R^k.a}{\bowtie} R^{k^{\mathcal{I}}}) \subseteq \Pi_{\varrho^{\mathsf{C_B}}}(S^{1^{\mathcal{I}}} \underset{S^1.b=S^2.b}{\bowtie} \cdots \underset{S^{k-1}.b=S^k.b}{\bowtie} S^{k^{\mathcal{I}}})$ • If $\mathsf{R\text{-}SET_{Exc}}((\{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}, \{R^1.a, \ldots, R^k.a\}),$ $(\{S^1.b_1, \ldots, S^1.b_n, \ldots, S^k.b_1, \ldots, S^k.b_m\}, \{S^1.b, \ldots, S^k.b\}), \mu) \in \Sigma$ then $\Pi_{\varrho^{\mathsf{C_A}}}(R^{1^{\mathcal{I}}} \underset{R^1.a=R^2.a}{\bowtie} \cdots \underset{R^{k-1}.a=R^k.a}{\bowtie} R^{k^{\mathcal{I}}}) \cap \Pi_{\varrho^{\mathsf{C_B}}}(S^{1^{\mathcal{I}}} \underset{S^1.b=S^2.b}{\bowtie} \cdots \underset{S^{k-1}.b=S^k.b}{\bowtie} S^{k^{\mathcal{I}}}) = \varnothing$
$\mathsf{O\text{-}SET_H} \subseteq \wp(\mathcal{E} \cup \mathcal{V}) \times \mathcal{E} \cup \mathcal{V}$ where $\mathsf{H} = \{\mathsf{Isa}, \mathsf{Tot}, \mathsf{Ex}\}$	• If $\mathsf{O\text{-}SET_{Isa}}(\{O_1, \ldots, O_n\}, O) \in \Sigma$ then $O_i^{\mathcal{I}} \subseteq O^{\mathcal{I}}$ for $1 \leq i \leq n$ • If $\mathsf{O\text{-}SET_{Tot}}(\{O_1, \ldots, O_n\}, O) \in \Sigma$ then $O^{\mathcal{I}} \subseteq \bigcup_{i=1}^{n} O_i^{\mathcal{I}}$ • If $\mathsf{O\text{-}SET_{Ex}}(\{O_1, \ldots, O_n\}, O) \in \Sigma$ then $\mathsf{O\text{-}SET_{Isa}}(\{O_1, \ldots, O_n\}, O) \in \Sigma$ and $O_i^{\mathcal{I}} \cap O_j^{\mathcal{I}} = \varnothing$ for any $1 \leq i < j \leq n$
$\mathsf{O\text{-}CARD} \subseteq (\mathcal{E} \cup \mathcal{V}) \times (\mathbb{N} \times (\mathbb{N} \cup \{\infty\}))$	If $\mathsf{O\text{-}CARD}(O) = (\mathsf{min}, \mathsf{max}) \in \Sigma$ then $\mathsf{min} \leq \sharp\{o \mid o \in O^{\mathcal{I}}\} \leq \mathsf{max}$
$\mathsf{R\text{-}CARD} \subseteq \mathcal{R} \times (\mathbb{N} \times (\mathbb{N} \cup \{\infty\}))$	If $\mathsf{R\text{-}CARD}(R) = (\mathsf{min}, \mathsf{max}) \in \Sigma$ then $\mathsf{min} \leq \sharp\{\langle o_1, \ldots, o_n \rangle \mid \langle o_1, \ldots, o_n \rangle \in R^{\mathcal{I}}\} \leq \mathsf{max}$
$\mathsf{OBJ} \subseteq \mathcal{R} \times (\mathcal{E} \cup \mathcal{V})$	If $\mathsf{OBJ}(R, O) \in \Sigma$ then $O^{\mathcal{I}} = \mathrm{ID}^{\mathcal{I}}(R^{\mathcal{I}})$
$\mathsf{RING_J} \subseteq \wp(\varrho \times \varrho)$ where $\mathsf{J} = \{\mathsf{Irr}, \mathsf{Asym}, \mathsf{Trans}, \mathsf{Intr}, \mathsf{Antisym}, \mathsf{Acyclic}, \mathsf{Sym}, \mathsf{Ref}\}$	If $\mathsf{RING_J}(R.a, R.b) \in \Sigma$ then $\Pi_{(R.a,R.b)} R^{\mathcal{I}}$ is *irreflexive, asymmetric, transitive, intransitive, antisymmetric, acyclic, symmetric, reflexive*
$\mathsf{V\text{-}VAL}\colon \mathcal{V} \to \wp(\Lambda_D)$ for some $\Lambda_D \in \Lambda$	If $\mathsf{V\text{-}VAL}(V) = \{v_1^D, \ldots, v_n^D\} \in \Sigma$ then $V^{\mathcal{I}} = \{v_1^D, \ldots, v_n^D\}$ for some D

$\mathcal{ALCQI}$ corresponds to the basic DL $\mathcal{ALC}$ equipped with *qualified cardinality restrictions* and *inverse roles*, and it can also be viewed as a fragment of $\mathcal{DLR}$ [4] where relations are restricted to be binary. The difficulty implied by the absence of n-ary relations has been overcome by means of *reification*. Unfortunately, apart from the necessity of introducing reified relations, the restricted expressive power of $\mathcal{ALCQI}$ does not allow to fully capture the semantics of the ORM2 constraints. The analysis of corresponding limitations thus led to identification of a fragment of ORM2, called ORM2$^{\text{zero}}$, that is maximal with respect to the expressiveness of $\mathcal{ALCQI}$, and that is still expressive enough to capture the most frequent usage pattern of the modelling community [5].

The ORM2$^{\text{zero}}$ fragment considers the following constraints:
ORM2$^{\text{zero}}$ $= \{\text{TYPE}, \text{FREQ}^-, \text{MAND}, \text{R-SET}^-, \text{O-SET}_{\text{Isa}}, \text{O-SET}_{\text{Tot}}, \text{O-SET}_{\text{Ex}}, \text{OBJ}\}$, where: (i) FREQ$^-$ can be applied to only one role at time, and (ii) R-SET$^-$ applies either to a pair of relations of the same arity or to two single roles. The encoding of the semantics of ORM2$^{\text{zero}}$ is shown in table 3.

Table 3. $\mathcal{ALCQI}$ encoding

Background domain axioms:	$E_i \sqsubseteq \neg(D_1 \sqcup \cdots \sqcup D_l)$ for each $i = 1, \ldots, n$ $V_i \sqsubseteq D_j$ for each $i = 1, \ldots, m$, and some j, with $1 \leq j \leq l$ $D_i \sqsubseteq \bigsqcap_{j=i+1}^{l} \neg D_j$ for each $i = 1, \ldots, l$
TYPE$(R.a, O)$	$\exists\tau(R.a)^-.A_R \sqsubseteq O$
FREQ$^-(R.a, \langle\text{min}, \text{max}\rangle)$	$\exists\tau(R.a)^-.A_R \sqsubseteq\, \geq \text{min}\ \tau(R.a)^-.A_R \sqcap \leq \text{max}\ \tau(R.a)^-.A_R$
MAND$(\{R^1.a_1, \ldots, R^1.a_n, \ldots, R^k.a_1, \ldots, R^k.a_m\}, O)$	$O \sqsubseteq \exists\tau(R^1.a_1)^-.A_{R^1} \sqcup \cdots \sqcup \exists\tau(R^1.a_n)^-.A_{R^1} \sqcup \cdots \sqcup$ $\exists\tau(R^k.a_1)^-.A_{R^k} \sqcup \cdots \sqcup \exists\tau(R^k.a_m)^-.A_{R^k}$
• If $A = \{R.a_1, \ldots, R.a_n\}$, $B = \{S.b_1, \ldots, S.b_n\}$ and $n = \lvert\varrho_R\rvert = \lvert\varrho_S\rvert$:	R-SET$^-_{\text{Sub}}(A, B)$ $A_R \sqsubseteq A_S$ R-SET$^-_{\text{Exc}}(A, B)$ $A_R \sqsubseteq A_{\top_n} \sqcap \neg A_S$
• If $A = \{R.a_i\}$, $B = \{S.b_j\}$:	R-SET$^-_{\text{Sub}}(A, B)$ $\exists\tau(R.a_i)^-.A_R \sqsubseteq \exists\tau(S.b_j)^-.A_S$ R-SET$^-_{\text{Exc}}(A, B)$ $\exists\tau(R.a_i)^-.A_R \sqsubseteq A_{\top_n} \sqcap \neg\exists\tau(S.b_j).A_S$
O-SET$_{\text{Isa}}(\{O_1, \ldots, O_n\}, O)$	$O_1 \sqcup \cdots \sqcup O_n \sqsubseteq O$
O-SET$_{\text{Tot}}(\{O_1, \ldots, O_n\}, O)$	$O \sqsubseteq O_1 \sqcup \cdots \sqcup O_n$
O-SET$_{\text{Ex}}(\{O_1, \ldots, O_n\}, O)$	$O_1 \sqcup \cdots \sqcup O_n \sqsubseteq O$ $O_i \sqsubseteq \bigsqcap_{j=i+1}^{n} \neg O_j$ for each $i = 1, \ldots, n$
OBJ(R, O)	$O \equiv A_R$

Given the encoding above, a fragment of $\mathcal{ALCQI}$ KB corresponding to the schema from the Fig. 1 is the following (where reified relations have been prefixed with 'R-'):

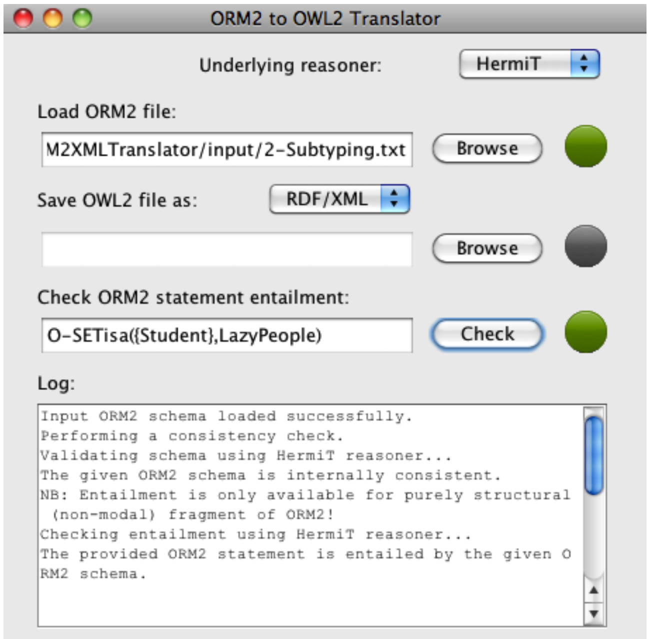
Example 1. $\exists\tau(\text{reportsTo.sub})^-.\text{R-reportsTo} \sqsubseteq \text{Admin}$
$\exists\tau(\text{reportsTo.obj})^-.\text{R-reportsTo} \sqsubseteq \text{AreaManager}$
$\text{Admin} \sqsubseteq \exists\tau(\text{reportsTo.sub})^-.\text{R-reportsTo}$

The correctness of the introduced encoding is guaranteed by the following theorem:

Theorem 1. *Let Σ^{zero} be an* ORM2$^{\text{zero}}$ *conceptual schema and $\Sigma^{\mathcal{ALCQI}}$ the $\mathcal{ALCQI}$ knowledge base constructed as described above. Then an entity/value type O is consistent in Σ^{zero} if and only if the concept O is satisfiable w.r.t. $\Sigma^{\mathcal{ALCQI}}$.*

5 Prototype of Automated Reasoning Support Tool

The ORM2 automated reasoning support tool is implemented in Java and includes a parser for ORM2 linear syntax, a set of Java classes representing the ORM2 knowledge database, a translator into an OWL2 ontology and a modal reasoning engine using HermiT or FaCT++ as an underlying reasoner. The graphical user interface of a tool is introduced on Figure 2 and contains controls which allow to select an input file, underlying reasoner, output file and output format for resulting ontology (if needed). The consistency check for an input ORM2 schema is performed after loading the input file and the result of the check is communicated by visual flag as well as by a detailed log in the corresponding window.



Underlying reasoner: HermiT

Load ORM2 file: M2XMLTranslator/input/2-Subtyping.txt

Save OWL2 file as: RDF/XML

Check ORM2 statement entailment: O-SETisa({Student},LazyPeople)

Log:

```
Input ORM2 schema loaded successfully.
Performing a consistency check.
Validating schema using HermiT reasoner...
The given ORM2 schema is internally consistent.
NB: Entailment is only available for purely structural
 (non-modal) fragment of ORM2!
Checking entailment using HermiT reasoner...
The provided ORM2 statement is entailed by the given O
RM2 schema.
```

Fig. 2. A conceptual schema including an instantiation of most of the ORM2 constraints

Let us now consider the following example in ORM2 linear syntax. This model describes a domain of university personnel, which is partitioned into three mutually exclusive categories: research staff, administration and some lazy people. We then introduce an entity type Student as a part of university personnel, which is claimed to be neither researchers nor administration staff.

ENTITYTYPES: {UNI-Personnel,LazyPeople,Student,Admin,RTStaff}
O-SETtot({LazyPeople,RTStaff,Admin},UNI-Personnel)
O-SETex({LazyPeople,RTStaff,Admin},UNI-Personnel)
O-SETisa({Student},UNI-Personnel)
O-SETex({Student,RTStaff},UNI-Personnel)
O-SETex({Student,Admin},UNI-Personnel)

The resulting encoding in $\mathcal{ALCQI}$ will look like the following:

$$\begin{aligned}
\mathsf{UNI\text{-}Personnel} &\sqsubseteq \mathsf{RTStaff} \sqcup \mathsf{Admin} \sqcup \mathsf{LazyPeople}\\
\mathsf{RTStaff} &\sqsubseteq \neg\mathsf{Admin} \sqcup \neg\mathsf{LazyPeople}\\
\mathsf{Admin} &\sqsubseteq \neg\mathsf{LazyPeople}\\
\mathsf{Student} &\sqsubseteq \neg\mathsf{RTStaff} \sqcup \neg\mathsf{Admin}\\
\mathsf{Student} &\sqsubseteq \mathsf{UNI\text{-}Personnel}
\end{aligned}$$

After loading the linear syntax input file, the corresponding conceptual schema is automatically checked for consistency (see Figure 2). As a result of the consistency check we obtain the message 'The given ORM2 schema is internally consistent', which confirms the correctness of the schema.

Implemented entailment check functionality allows user to analyze the conceptual schema and to discover interesting inferences. For example, let us analyze the entity type Student, which is claimed to be part of the university personnel. However, since the partition of UNI-Personnel described above is total, the only possibility for Student to be non-empty is to be equivalent to the entity type LazyPeople. Which is indeed confirmed by the inference engine of the implemented prototype when performing entailment check for the following ORM2 statement (see Figure 2):

O-SETisa({Student},LazyPeople)

6 Related Works

In the last few years, several papers addressed the issue of encoding ORM2 conceptual schema into DL knowledge bases [15,14,13,11]. Among those proposals, [15] can be taken as the only one going through the encoding with a formal perspective. In particular, [15] pretends to start from the Halpin's FOL semantics, and introduces an encoding of a fragment of ORM2 into the logic $\mathcal{DLR}_{ifd}$ [3]. In general, the paper suffers from the presence of several imprecisions, redundancy, and syntactical mistakes that makes the proposed mapping solutions not always clearly understandable. Moreover, the bottom-up approach that avoids the specification of a complete theoretical framework for the mapping of the ORM2 semantics into $\mathcal{DLR}_{ifd}$, makes some of these solutions extremely questionable, such as in the case where 'objectification' is simply treated as 'relation reification' in DL.

As regards to [14], we mostly rely here on the extensive review already made by Keet in [15]. Starting from this, it should be also noticed that subsequent attempts, focused on the possibilities of encoding ORM2 into the the web ontology language OWL2 [13,11], suffer from the same formal inconsistencies and limitations of [14]. In particular, [14] is misleading with respect to the underlying DL formalism: distinct DL languages (e.g. $\mathcal{DLR}$, plus $\mathcal{DLR}$-*Lite*, plus $\mathcal{SROIQ}$, plus 'role composition' operator) are there arbitrary mixed together. No special semantics is provided by [14] in correspondence with these combinations, nor theorems showing the complexity of reasoning with them.

Another paper focused on the encoding of ORM2 in OWL has also been recently published [16]. The paper introduces a set of informal 'rules' devoted to the mapping of a subset of ORM2 constructs into OWL Manchester Syntax [7]. Unfortunately, the paper is misleading in several respects (for instance: (i) the OWL EquivalentTo, instead of the SubClassOf, is erroneously introduced several times; (ii) optionality of uniqueness constraints is definitively lost). In general, the paper covers a fragment that is smaller than ORM2$^{\text{zero}}$, and the proposed mapping mostly remains formally unjustified.

7 Concluding Remarks

In this paper we introduced a linear syntax and a complete set-theoretic semantics for the ORM2 conceptual modelling language. A decidable, and computationally tractable, fragment of ORM2 has been clearly identified and mapped into the DL logic $\mathcal{ALCQI}$. Finally, a first reasoning support prototype for ORM2 has been implemented, which enables consistency and entailment checks for the defined fragment of ORM2. Future theoretic works will be mainly focused on the extension of the ORM2$^{\text{zero}}$ towards the identification of a more expressive, still decidable, 'object role' modelling language. The practical objectives of the research will be directed towards full integration of the prototype into third-party solutions providing graphical user interface for designing ORM2 conceptual schemas (e.g. NORMA plugin for Microsoft Visual Studio). In particular, such integration will benefit from reasoning capabilities of the tool by providing the user with a list of all meaningful inferences entailed by the original schema.

References

1. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The description logic handbook: theory, implementation, and applications. Cambridge University Press, New York, NY, USA (2003)
2. Berardi, D., Cali, A., Calvanese, D., Giacomo, G.D.: Reasoning on UML class diagrams. Artificial intelligence 168 (2003)
3. Calvanese, D., De Giacomo, G., Lenzerini, M.: Identification constraints and functional dependencies in description logics. In: Proceedings of the 17th international joint conference on Artificial intelligence - Volume 1. pp. 155–160. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2001), `http://dl.acm.org/citation.cfm?id=1642090.1642111`
4. Calvanese, D., De Giacomo, G., Lenzerini, M., Nardi, D., Rosati, R.: Description logic framework for information integration. In: Proc. of the 6th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR'98). pp. 2–13 (1998)
5. Calvanese, D., Lenzerini, M., Nardi, D.: Description logics for conceptual data modeling. In: Chomicki, J., Saake, G. (eds.) Logics for Databases and Information Systems. pp. 229–263. Kluwer (1998)
6. Ceravolo, P., Fugazza, C., Leida, M.: Modeling semantics of business rules. In: Proceedings of the Inaugural IEEE International Conference On Digital Ecosystems and Technologies (IEEE-DEST) (February 2007)
7. Grau, B.C., Hitzler, P., Shankey, C., Wallace, E. (eds.): Proceedings of the OWLED*06 Workshop on OWL: Experiences and Directions, Athens, Georgia, USA, November 10-11, 2006, CEUR Workshop Proceedings, vol. 216. CEUR-WS.org (2006)
8. Haarslev, V., Mller, R.: Expressive ABox reasoning with number restrictions, role hierarchies, and transitively closed roles. In: KR'00. pp. 273–284 (2000)
9. Halpin, T.: A logical analysis of information systems: static aspects of the data-oriented perspective. PhD thesis, Department of Computer Science, University of Queensland (1989)

10. Halpin, T., Morgan, T.: Information modeling and relational databases: from conceptual analysis to logical design. Morgan Kaufmann, 2nd edn. (2001)
11. Hodrob, R., Jarrar, M.: ORM to OWL2 DL mapping. In: International conference on intelligent semantic web: applications and services. ACM (2010)
12. Horrocks, I., Sattler, U., Tobies, S.: Practical reasoning for expressive description logics. In: Proceedings of the 6th International Conference on Logic Programming and Automated Reasoning. pp. 161–180. LPAR '99, Springer-Verlag, London, UK (1999), `http://dl.acm.org/citation.cfm?id=645709.664314`
13. Jarrar, M.: Mapping ORM into the SHOIN/OWL description logic – towards a methodological and expressive graphical notation for ontology engineering. In: OTM 2007 workshops: Proceedings of the International Workshop on Object-Role Modeling (ORM'07). LNCS, vol. 4805, pp. 729–741. Springer (November 2007)
14. Jarrar, M.: Towards automated reasoning on ORM schemes. mapping ORM into the $\mathcal{DLR}_{ifd}$ description logic. In: Proceedings of the 26th International Conference on Conceptual Modeling (ER 2007). LNCS, vol. 4801, pp. 181–197. Springer (November 2007)
15. Keet, M.: Mapping the Object-Role Modeling language ORM2 into description logic language $\mathcal{DLR}_{ifd}$. Tech. Rep. KRDB07-2, KRDB Research Centre, Faculty of Computer Science, Free University of Bozen-Bolzano (2007)
16. Wagih, H.M., ElZanfaly, D.S., Kouta, M.M.: Mapping Object Role Modeling 2 schemes to OWL2 ontologies. In: Ting, Z. (ed.) Proceedings of the 3rd IEEE International Conference on Computer Research and Development (ICCRD), Shanghai, China, March 11-13, 2011. vol. 4, pp. 126–132. IEEE Press (2011)