Meta-Modeling Extension of Horn-SROIQ and Query Answering

Zhenzhen Gu

Academy of Mathematics and Systems Science, Chinese Academy of Sciences,
Beijing, China

Abstract. We investigate reasoning and query answering in expressive domain knowledge bases (KBs) that are in OWL 2 Full and contain large individual data sets. For this, we introduce Hi(Horn-SROIQ) and meta-queries based on HiLog semantics. Hi(Horn-SROIQ) is extended from Horn-SROIQ, the horn fragment of the most expressive description logic (DL) for OWL 2 with well known low data complexity, by allowing classes and roles to be used as individuals. Accordingly, meta-queries are obtained from conjunctive queries by allowing variables to appear in the class and role positions. For reasoning, we first provide a method of reducing satisfiability checking and conjunctive query answering in Hi(Horn-SROIQ) to the corresponding reasoning tasks in Horn-SROIQ soundly and completely, then we show that meta-query answering in Hi(Horn-SROIQ) can be captured by conjunctive query answering. Based on this, we obtain that adding meta-modeling capability to Hi(Horn-SROIQ) has no impact on the complexity of the considered reasoning tasks. These results make Hi(Horn-SROIQ) appealing for practical usage.

1 Introduction

With the successful application of the Semantic Web technologies, more and more domain KBs have been developed over the past decades for reusing and sharing. Most of them are in OWL 2 DL so that efficient and mature DL Reasoners, such as HermiT [8], Pellet [33] and RacerPro [10], can take effect to facilitate reasoning and query answering. However, caused by meta-modeling, there exist certain complex KBs that fall into OWL 2 Full. This can be exemplified by the Foundational Model of Anatomy (FMA) [27], common sense KB OpenCyc [23] and Suggested Upper Merged Ontology (SUMO) [31]. All of these widely used KBs not only contain large-scale individual data but also most classes and roles are also used as individuals at the same time. OWL 2 Full is undecidable. Compared with OWL 2 DL, reasoning in OWL 2 Full has largely been unexplored. Up to now, no Reasoner has been especially developed for it. Thus, efficiently reasoning and query answering in such KBs raise new challenges.

For meta-modeling, OWL 2 DL provides a technique called punning by syntactically allowing names to have multiple uses while semantically treating the different uses of the same name as completely separate [9]. Such way of processing meta-modeling dooms to no new entailments can be obtained. Moreover, in order

to reuse the techniques as well as systems for DLs, [3,7,11,13,21,24–26,30] investigate adding meta-modeling capability to decidable DLs by allowing one name to act as multiple roles in various ways. The referred DLs contain SHOIQ [15], SHIQ [16], SROIQ [14], and ALCQ [5]. For these works, excepting high reasoning complexity, query answering which plays an important role in Semantic Web for realizing knowledge sharing and reusing has rarely been discussed, nor the ability of handing scalability. Extension of tractable language DL-Lite$_R$ is discussed in [3] and [22]. Although having low reasoning complexity, the expressivity is every restricted.

On the other hand, in DL based knowledge representation systems, individual assertions describe the concrete world and usually account for a large proportion of domain KBs, such as, 94.6% in FMA and 78.5% in OpenCyc. As more and more applications require scalability in terms of individual assertions, Horn-DLs [1,4,18–20,29] have been introduced and attract more and more attention, since they promise to make a suitable trade-off between scalability and expressivity. In [28] and [29], Ortiz has shown that data complexities (measured in the size of individual assertions) of satisfiability checking and conjunctive query answering in Horn-SROIQ, the horn fragment of the most expressive DL for OWL 2, are PTime and PTime-complete. Lower data complexity makes Horn DLs a natural and efficient choice for reasoning with large number of individuals. Moreover, [20] has pointed out that many OWL 2 ontologies are Horn. When omitting meta-modeling, many actual OWL 2 Full ontologies, such as OpenCyc, SUMO and Yago, fall into Horn-SROIQ.

For reasoning and query answering in OWL 2 Full ontologies containing large number of individuals, attracted by the low data complexity as well as suitable expressivity of Horn-SROIQ, in this paper, we discuss extending Horn-SROIQ with meta-modeling capability. Our contributions can be summarized as follows:

(1) We formalize Hi(Horn-SROIQ) and meta-queries based on HiLog semantics. Hi(Horn-SROIQ) is defined by unifying the sets of names for classes, roles and individuals. Accordingly, meta-queries are defined by allowing variables to appear in the class and role positions of conjunctive queries.

(2) We provide a way of reducing satisfiability checking and conjunctive query answering in a Hi(Horn-SROIQ) KB to the corresponding reasoning tasks in a Horn-SROIQ KB with size no more than the original KB soundly and completely.

(3) We show that meta-query answering in Hi(Horn-SROIQ) can be reduced to conjunctive query answering through materializing the variables appearing in the class and role positions soundly and completely.

(4) We show that satisfiability checking and meta-query answering in Hi(Horn-SROIQ) still have PTime and PTime-complete data complexities, respectively.

From the practical point of view, our work has the following two advantages. Firstly, by allowing all the names to have different uses without any restriction, Hi(Horn-SROIQ) has the flexibility to capture meta-modeling in the actual KBs. And secondly, by reasoning reduction, the systems tailored for DLs or Horn Logic can be used to realize reasoning with Hi(Horn-SROIQ).

2 Hi(Horn-SROIQ) and meta-query

In this section, we formalize the syntax and semantics of Hi(Horn-SROIQ) and meta-queries. And we start with defining Hi(Horn-SROIQ) based on Horn-SROIQ [20, 29]

r_1	$R \circ R \sqsubseteq_r R$	B, C and D are inductively defined as follows:
r_2	$\mathsf{Inv}(R) \sqsubseteq_r R$	
r_3	$\omega \sqsubseteq_r R$	$B ::= A\|\{a\}\|\exists S.\mathsf{Self}$
r_4	$\omega \circ R \sqsubseteq_r R$	$C ::= B\|\exists R.C\|C \sqcup C\|C \sqcap C$
r_5	$R \circ \omega \sqsubseteq_r R$	$D ::= B\|\neg B\|\exists R.D\| \geq mS.D\| \leq 1S.D$
r_6	$\mathsf{Disj}(S, S')$	$\|D \sqcap D$
r_7	$C \sqsubseteq_c D$	
r_8	$B(a), \quad R(a,b), \quad a \approx b, \quad a \not\approx b$	

Fig. 1. Syntax of Hi(Horn-SROIQ) where R is a role, $\omega = R_1 \circ \cdots \circ R_n$ and $R_i \prec R$ for each $1 \leq i \leq n$, S and S' are simple roles, and $A, P, a, b \in N$

Hi(Horn-SROIQ). Let N be a countably infinite set of names such that $\{\top, \bot\} \subseteq N$. For each $P \in N$, P and P^- are roles, and their respective inverses are $\mathsf{Inv}(P) = P^-$ and $\mathsf{Inv}(P^-) = P$. A role hierarchy $\mathcal{R}_h$ is a set of role inclusion axioms which take the form $R_1 \circ \cdots \circ R_n \sqsubseteq_r R$ where $R_1, \cdots, R_n$ and R are roles. $\mathcal{R}_h$ is called regular, if there exists an irreflexive and transitive binary relation $\prec$ on the set $N \cup \{P^- | P \in N\}$ of roles such that $S \prec R$ iff $\mathsf{Inv}(S) \prec R$ for all roles S and R, and all the role inclusion axioms in $\mathcal{R}_h$ have the forms $r_1 - r_5$ in figure 1.

Given a role hierarchy $\mathcal{R}_h$, the set of roles that are non-simple in $\mathcal{R}_h$ is inductively defined as follows: (1) R is non-simple if $\mathcal{R}_h$ contains a RIA $R_1 \circ \cdots \circ R_n \sqsubseteq_r R$ such that $n > 1$, or $n = 1$ and R_1 is non-simple; (2) an inverse role $\mathsf{Inv}(R)$ is non-simple if R is. A role R is simple in $\mathcal{R}_h$ if it is not non-simple in $\mathcal{R}_h$. Given $\mathcal{R}_h$, a role disjoint axiom and a class inclusion axiom take the form r_6 and form r_7 in figure 1, respectively, and an individual assertion has one of the forms in r_8 in figure 1.

A Hi(Horn-SROIQ) KB $\mathcal{K}$ is a tuple $(\mathcal{R}_h \cup \mathcal{R}_d \cup \mathcal{C}, \mathcal{A})$ where $\mathcal{R}_h$ is a regular role hierarchy, $\mathcal{R}_d$, $\mathcal{C}$ and $\mathcal{A}$ are finite sets of role disjoint axioms, class inclusion axioms, and individual assertions, respectively. And we use $|\mathcal{K}|$, $|\mathcal{R}_h \cup \mathcal{R}_d \cup \mathcal{C}|$, $|\mathcal{A}|$ to denote the size of $\mathcal{K}$, $\mathcal{R}_h \cup \mathcal{R}_d \cup \mathcal{C}$ and $\mathcal{A}$, respectively.

Hi(Horn-SROIQ) does not separate names for classes, roles and individuals. Thus, symbols $\sqsubseteq_c$ and $\sqsubseteq_r$ are used to distinguish between class inclusion axioms and role inclusion axioms. Moreover, for simplification, we use $a =_l b$ as an abbreviation of the two axioms $a \sqsubseteq_l b$ and $b \sqsubseteq_l a$, where $l \in \{c, r\}$. The following example illustrates a simple Hi(Horn-SROIQ) KB which models some knowledge about football described in OpenCyc.

Example 1. Let $\mathcal{K}$ be a Hi(Horn-SROIQ) KB consisting of the following axioms $((a),\ (b)$ and $(c))$ and individual assertions $((d),\ (e),$ and $(f))$:

`SportsTeam` $\sqsubseteq_c$ `¬AllStarTeam`	(a)
`Football_team` $\sqsubseteq_c$ `SportsTeam`	(b)
$\top \sqsubseteq_c\ \leq$ `1rewriteOf.` $\top \sqcap$ `∃rewriteOf.Self`	(c)
`rewriteOf(FootballTeam, Football_team)`	(d)
`SportsTeamTypeBySport(Football_team)`	(e)
`FootballTeam(BarcelonaDragons)`	(f)

In this Hi(Horn-SROIQ) KB, names `FootballTeam` and `Football_team` have multiple uses, i.e, as classes and individuals.

The semantics of Hi(Horn-SROIQ) is captured by v-interpretation [26] which is based on HiLog [2] and takes a similar way of OWL 2 RDF-Based Semantics [32] to interpret the multiple uses of names in a KB.

Syntax	Semantics
P	$\mathfrak{R}^{\mathcal{V}}(P^{\mathcal{V}})$
P^-	$\{(y,x)\|(x,y)\in\mathfrak{R}^{\mathcal{V}}(P^{\mathcal{V}})\}$
A	$\mathfrak{C}^{\mathcal{V}}(A^{\mathcal{V}})$
$\{a\}$	$\{a^{\mathcal{V}}\}$
$\exists S.\mathsf{Self}$	$\{x\|(x,x)\in\mathfrak{R}^{\mathcal{V}}(R)\}$
$\neg B$	$\Delta^{\mathcal{V}}-\mathfrak{C}^{\mathcal{V}}(B)$
$C\sqcap D$	$\mathfrak{C}^{\mathcal{V}}(C)\cap\mathfrak{C}^{\mathcal{V}}(D)$
$C\sqcup D$	$\mathfrak{C}^{\mathcal{V}}(C)\cup\mathfrak{C}^{\mathcal{V}}(D)$
$\exists R.C$	$\{x\|\exists y.(x,y)\in\mathfrak{R}^{\mathcal{V}}(R)\wedge y\in\mathfrak{C}^{\mathcal{V}}(C)\}$
$\forall R.C$	$\{x\|\forall y.(x,y)\in\mathfrak{R}^{\mathcal{V}}(R)\rightarrow y\in\mathfrak{C}^{\mathcal{V}}(C)\}$
$\leq 1S.D$	$\{x\|\#\{y\|(x,y)\in\mathfrak{R}^{\mathcal{V}}(S)\wedge y\in\mathfrak{C}^{\mathcal{V}}(D)\}\leq 1\}$
$\geq mS.D$	$\{x\|\#\{y\|(x,y)\in\mathfrak{R}^{\mathcal{V}}(S)\wedge y\in\mathfrak{C}^{\mathcal{V}}(D)\}\geq m\}$
$B\sqsubseteq_c C$	$\mathfrak{C}^{\mathcal{V}}(B)\subseteq\mathfrak{C}^{\mathcal{V}}(C)$
$\omega\sqsubseteq_r R$	$\mathfrak{R}^{\mathcal{V}}(R_1)\circ\cdots\circ\mathfrak{R}^{\mathcal{V}}(R_n)\subseteq\mathfrak{R}^{\mathcal{V}}(R)$
$\mathsf{Disj}(S,R)$	$\mathfrak{R}^{\mathcal{V}}(S)\cap\mathfrak{R}^{\mathcal{V}}(R)=\emptyset$
$B(a)$	$a\in\mathfrak{C}^{\mathcal{V}}(B)$
$R(a,b)$	$(a^{\mathcal{V}},\ b^{\mathcal{V}})\in\mathfrak{R}^{\mathcal{V}}(R)$
$a\approx b$	$a^{\mathcal{V}}=b^{\mathcal{V}}$
$a\not\approx b$	$a^{\mathcal{V}}\neq b^{\mathcal{V}}$

Fig. 2. Semantics of Hi(Horn-SROIQ) refers to a v-interpretation $\mathcal{V}$ where P, A, a and b are names in N, $m\geq 0$ and $\omega=R_1\circ\cdots\circ R_n$.

v-semantics. A v-interpretation $\mathcal{V}=(\Delta^{\mathcal{V}},\cdot^{\mathcal{V}},\mathfrak{C}^{\mathcal{V}},\mathfrak{R}^{\mathcal{V}})$ is a tuple where $\Delta^{\mathcal{V}}$ is a non-empty domain set, $\cdot^{\mathcal{V}}$, $\mathfrak{C}^{\mathcal{V}}$ and $\mathfrak{R}^{\mathcal{V}}$ are functions satisfying: (a) $\cdot^{\mathcal{V}}$ maps each name in N to an element in $\Delta^{\mathcal{V}}$; (b) $\mathfrak{C}^{\mathcal{V}}$ maps each element in $\Delta^{\mathcal{V}}$ to a subset of $\Delta^{\mathcal{V}}$ such that $\mathfrak{C}^{\mathcal{V}}(\top^{\mathcal{V}})=\Delta^{\mathcal{V}}$ and $\mathfrak{C}^{\mathcal{V}}(\bot^{\mathcal{V}})=\emptyset$; (c) $\mathfrak{R}^{\mathcal{V}}$ maps each element in $\Delta^{\mathcal{V}}$ to a subset of $\Delta^{\mathcal{V}}\times\Delta^{\mathcal{V}}$. The interpretation of class and role constructors as well as axioms and assertions are specified in figure 2. We say $\mathcal{V}$ is a v-model of a Hi(Horn-SROIQ) KB $\mathcal{K}$ if $\mathcal{V}$ satisfies each axiom and assertion in $\mathcal{K}$. We say

$\mathcal{K}$ is v-satisfiable iff it has a v-model. Moreover, the v-entailment ($\models_v$) is defined as usual.

In a v-interpretation, the class and role extensions of names are obtained from the domain elements they mapped into. This indicates that under v-semantics, a Hi(Horn-SROIQ) KB may imply that a non-simple role and a simple role are equivalent. Such equivalences between roles may cause (a) transitive roles ($R \circ R \sqsubseteq_r R$) are used in number restrictions ($\leq 1R.D$ and $\geq mS.D$) or (b) role hierarchies contain cyclic dependencies. (a) and (b) are well known factors that lead to undecidability [16, 17]. Consider the following axioms and assertion:

$$P_5 \circ P_4 \sqsubseteq_r P_4, \ P_3 \circ P_4 \sqsubseteq_r P_3, \ P_3 \circ P_2 \sqsubseteq_r P_2, \ P_1 \circ P_2 \sqsubseteq_r P_1 \ (1)$$
$$P_1 \circ P_1 \sqsubseteq_r P_1, \top \sqsubseteq_c \geq 5P_5.\top \qquad\qquad (2)$$
$$P_5 \approx P_1 \qquad\qquad\qquad (3)$$

Under v-semantics, the assertion in (3) implies the axiom $P_1 =_r P_5$. Such equivalence between non-simple role P_1 and simple role P_5 leads to the fact that transitive role P_1 is used in the number restriction in (2) and the roles in (1) contain cyclic dependencies. For decidability, motivated by [26], we assume Hi(Horn-SROIQ) adopts **unique non-simple role assumption** (UNRA), i.e., for each Hi(Horn-SROIQ) KB $\mathcal{K}$ and each two different names a and b in $\mathcal{K}$, if a or b is a non-simple role in $\mathcal{K}$ then we assume $\mathcal{K}$ contains the assertion $a \not\approx b$.

Meta-queries. Let V be a set of variables such that $V \cap N = \emptyset$. A query atom has the form $x(y)$, $x(y, z)$ or $x \approx y$ where $x, y, z \in N \cup V$. A meta-query Q is an expression of the form:

$$\alpha_1 \wedge \cdots \wedge \alpha_n \rightarrow q(\boldsymbol{x})$$

where $\alpha_1, \cdots, \alpha_n$ are query atoms and $\boldsymbol{x}$ called the head of Q is a tuple of elements in $N \cup V$ such that each variable in $\boldsymbol{x}$ appears in some α_i. Variable x is called a class (role) variable of Q if Q contains an atom $x(y)$ ($x(y, z)$). Without class and role variables, Q is a conjunctive query.

For a tuple $\boldsymbol{x}$, we use $|\boldsymbol{x}|$ to denote the length of $\boldsymbol{x}$ and use $\boldsymbol{x}[i]$ to denote the i-th element of $\boldsymbol{x}$ for $1 \leq i \leq |\boldsymbol{x}|$. For another tuple $\boldsymbol{a}$ such that $|\boldsymbol{a}| = |\boldsymbol{x}|$, we use $[\boldsymbol{x}/\boldsymbol{a}]$ to denote a substitution. And for an object O (a KB, query or tuple), we use $O[\boldsymbol{x}/\boldsymbol{a}]$ to denote the result of replacing each appearance of $\boldsymbol{x}[i]$ in O with $\boldsymbol{a}[i]$ for each $1 \leq i \leq |\boldsymbol{x}|$.

Meta-query answering. For v-interpretation $\mathcal{V}$ and meta-query Q, a binding π of Q over $\mathcal{V}$ is a function that maps each variable in Q to an element in $\Delta^{\mathcal{V}}$ and each name a in Q to $a^{\mathcal{V}}$. The satisfaction of query atoms is defined as follows:

$$\mathcal{V}, \pi \models_v x(y) \text{ if } \pi(y) \in \mathfrak{C}^{\mathcal{V}}(\pi(x))$$
$$\mathcal{V}, \pi \models_v x(y, z) \text{ if } (\pi(y), \pi(z)) \in \mathfrak{R}^{\mathcal{V}}(\pi(x))$$
$$\mathcal{V}, \pi \models_v x \approx y \text{ if } \pi(x) = \pi(y)$$

We write $\mathcal{V}, \pi \models_v Q$ if $\mathcal{V}$ and π satisfy all the query atoms in Q. We write $\mathcal{V} \models_v Q$ if there exists a binding π of Q over $\mathcal{V}$ such that $\mathcal{V}, \pi \models_v Q$. Let $\boldsymbol{x}$ be the head of Q. For Hi(Horn-SROIQ) KB $\mathcal{K}$, a tuple $\boldsymbol{u}$ of names appearing in $\mathcal{K}$, such that

$|\boldsymbol{u}| = |\boldsymbol{x}|$ and if $|\boldsymbol{x}| \geq 1$, $\boldsymbol{u}[i] = \boldsymbol{x}[i]$ for each $1 \leq i \leq |\boldsymbol{x}|$ with $\boldsymbol{x}[i] \in N$, is called a certain answer of Q over $\mathcal{K}$ if $\mathcal{V} \models_v Q[\boldsymbol{x}/\boldsymbol{u}]$ for each v-model $\mathcal{V}$ of $\mathcal{K}$. We use $\mathsf{ans}_v(Q, \mathcal{K})$ to denote the set of all the certain answers of Q over $\mathcal{K}$.

In a meta-query, the variables that do not appear in the head of this query are treated as existential variables. This is the main difference between meta-queries and SPARQL OWL 2 RDF-Based Semantics Entailment Regime [6] where all the variables in a query are treated as free variables. Such difference indicates that under v-semantics, more certain answers can be obtained. For the reasoning tasks of Hi(Horn-SROIQ), we just consider v-satisfiability checking and meta-query answering, since the other reasoning problems, such as individual checking, can be reduced to v-satisfiability checking.

3   Reasoning with Hi(Horn-SROIQ)

In this section, we study reasoning in Hi(Horn-SROIQ). For this, we first present a way of reducing v-satisfiability checking and conjunctive query answering in Hi(Horn-SROIQ) to the corresponding reasoning tasks in Horn-SROIQ soundly and completely, then we show that meta-query answering in Hi(Horn-SROIQ) can be captured by conjunctive query answering.

3.1 v-satisfiability checking and conjunctive query answering

For the reasoning reduction, a naive solution is to use OWL 2 DL punning where a Hi(Horn-SROIQ) KB can be considered as a Horn-SROIQ KB by treating the multiple uses of a same name as completely separate. Although simple, completeness can not be guaranteed. Next, we explain the reason and provide a technique to obtain completeness. Before that, we first show a way of translating a Hi(Horn-SROIQ) KB into a Horn-SROIQ KB by renaming the names appearing in the class and role positions in the original KB.

role R	$\tau_r(R)$
P	$v_r(P)$
P^-	$v_r(P)^-$

axiom α	$\tau(\alpha)$
$C \sqsubseteq_c D$	$\tau_c(C) \sqsubseteq \tau_c(D)$
$\omega \sqsubseteq_r R$	$\omega' \sqsubseteq \tau_r(R)$
$\mathsf{Disj}(S, R)$	$\mathsf{Disj}(\tau_r(S), \tau_r(R))$
$B(a)$	$\tau_c(B)(a)$
$R(a, b)$	$\tau_r(R)(a, b)$
$a \approx b$	$a \approx b$
$a \not\approx b$	$a \not\approx b$

class C	$\tau_c(C)$
A	$v_c(A)$
$\{a\}$	$\{a\}$
$\exists S.\mathsf{Self}$	$\exists \tau_r(S).\mathsf{Self}$
$\neg B$	$\neg \tau_r(B)$
$C \sqcap D$	$\tau_c(C) \sqcap \tau_c(D)$
$C \sqcup D$	$\tau_c(C) \sqcup \tau_c(D)$
$\exists R.C$	$\exists \tau_r(R).\tau_c(C)$
$\forall R.C$	$\forall \tau_r(R).\tau_c(C)$
$\leq 1 S.D$	$\leq 1 \tau_r(S).\tau_c(D)$
$\geq m S.D$	$\geq m \tau_r(S).\tau_c(D)$

atom α	$\tau(\alpha)$
$A(x)$	$v_c(A)(x)$
$P(x, y)$	$v_r(P)(x, y)$
$x \approx y$	$x \approx y$

Fig. 3. Definition of functions τ_r, τ_c and τ where P, A, a and b are names in N, $m \geq 0$, $\omega = R_1 \circ \cdots \circ R_n$, $\omega' = \tau_r(R_1) \circ \cdots \circ \tau_r(R_n)$, and x and y are elements in $N \cup V$.

Let N_C and N_R be two disjoint sets of names for Horn-SROIQ classes and roles such that N_C and N_R are disjoint with N and have the same size with N. For simplicity, we specify N to be the set of names for Horn-SROIQ individuals. Let v_c and v_r be two injective functions that map each name in N to a unique Horn-SROIQ class name and role name, respectively. The translation of Hi(Horn-SROIQ) classes and roles as well as axioms, assertions and query atoms is realized by functions τ_c, τ_r and τ defined in figure 3. For a Hi(Horn-SROIQ) KB $\mathcal{K}$, we use $\tau(\mathcal{K})$ to denote the KB obtained by replacing each axiom (assertion) α in $\mathcal{K}$ with $\tau(\alpha)$. And for a conjunctive query q, we use $\tau(q)$ to denote the query obtained by replacing each query atom α in q with $\tau(\alpha)$. The soundness of punning in terms of KB non-satisfiability and conjunctive query answering is shown in the following proposition.

Proposition 1 *For Hi(Horn-SROIQ) KB $\mathcal{K}$, we can get that:*

- *if $\mathcal{K}$ is v-satisfiable then $\tau(\mathcal{K})$ is satisfiable;*
- $\mathsf{ans}(\tau(q), \tau(\mathcal{K})) \subseteq \mathsf{ans}_v(q, \mathcal{K})$[1] *for each conjunctive query q.*

Next, we show the incompleteness of punning by an example. Consider the Hi(Horn-SROIQ) KB $\mathcal{K}$ in example 1. Obviously, $\mathcal{K}$ is v-satisfiable. However, when we add the axiom (g): $\mathtt{FootballTeam} \sqsubseteq_c \mathtt{AllStarTeam}$ to it, the resultant KB $\mathcal{K}'$ is not v-satisfiable anymore. Since, under v-semantics, by (d), (c), (b) and (a), we can get that $\mathcal{K} \models_v \mathtt{FootballTeam} \sqsubseteq_c \neg \mathtt{AllStarTeam}$. This conclusion contradicts with (g). However $\mathcal{K}'$ is still satisfiable in terms of punning. Such incompleteness is caused by the fact that under v-semantics, the behaviors of names used as individuals effect the same names used as classes and roles.

For completeness, an intuitive way is to use materialization, i.e, materializing the entailments that can be entailed by v-semantics but not punning to the original Hi(Horn-SROIQ) KBs. The next Lemma, which illustrates a sufficient and necessary condition under which v-semantics and punning coincide in terms of KB satisfiability and conjunctive query answering, indicates a way of doing such materialization soundly and completely.

Lemma 1 *For a Hi(Horn-SROIQ) KB $\mathcal{K}$, if $\mathcal{K}$ contains the axiom $a =_c b$ and $a =_r b$ for each two different names a and b in N such that $\tau(\mathcal{K}) \models a \approx b$, then:*

- $\mathcal{K}$ *is v-satisfiable iff $\tau(\mathcal{K})$ is satisfiable;*
- $\mathsf{ans}_v(q, \mathcal{K}) = \mathsf{ans}(\tau(q), \tau(\mathcal{K}))$ *for each conjunctive query q;*

For a Hi(Horn-SROIQ) KB $\mathcal{K}$, by lemma 1, the materialization can be easily realized by the following two steps. (1) If $\tau(\mathcal{K})$ is satisfiable then for each assertion $a \approx b$ entailed by $\tau(\mathcal{K})$, add the axioms $a =_c b$ and $a =_r b$ to $\mathcal{K}$; (2) Iterate (1) until $\mathcal{K}$ does not change anymore. Then by proposition 1 and lemma 1, we can easily obtain that reasoning with the original KB under v-semantics can

---

[1] $\mathsf{ans}(\tau(q), \tau(\mathcal{K}))$ denotes the set of certain answers of $\tau(q)$ over $\tau(\mathcal{K})$

be captured by reasoning with the resultant KB using punning soundly and completely. However, such naive way of materialization will lead to the size of the original KB increased quadratically.

Note that, lemma 1 also indicates that if a Hi(Horn-SROIQ) KB cannot entail any equivalence between different names by punning then v-semantics and punning are also coincident in terms of KB satisfiability and conjunctive query answering. This motivates us that in order not to increase the size of the original Hi(Horn-SROIQ) KB, for each detected name equivalence, we use one name to replace the other name in the original KB rather than add new axioms. For a Hi(Horn-SROIQ) KB $\mathcal{K}$, let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two tuples such that $|\boldsymbol{a}|=|\boldsymbol{b}|=0$. Then such materialization procedure can be realized by the following two steps:

Step 1: If $\tau(\mathcal{K})$ is satisfiable then for each two different names a and b such that $\tau(\mathcal{K})\models a\approx b$, let $\mathcal{K}=\mathcal{K}[b/a]$, $\boldsymbol{b}=(\boldsymbol{b},b)$ and $\boldsymbol{a}=(\boldsymbol{a}[b/a],a)$;
Step 2: Iterate step 1, until $\mathcal{K}$ does not change anymore.

For $\mathcal{K}$, we use $\widetilde{\mathcal{K}}$ to denote the KB obtained by step 1 and step 2. And we use $\varepsilon_{\mathcal{K}}$ to denote the corresponding substitution $[\boldsymbol{b}/\boldsymbol{a}]$. The correctness of this materialization procedure can be guaranteed by the following two lemmas which respectively indicate that after a sires of name replacement, the resultant KB is still a Hi(Horn-SROIQ) KB, and reasoning with the original KB can be captured by the resultant KB soundly and completely.

Lemma 2 *Let $\mathcal{K}$ be a Hi(Horn-SROIQ) KB and a and b be two names in N. If $\tau(\mathcal{K})$ is satisfiable and entails $a\approx b$ then $\mathcal{K}[b/a]$ is still a Hi(Horn-SROIQ) KB.*

Lemma 3 *Let $\mathcal{K}$ be a Hi(Horn-SROIQ) KB and $\delta=[\boldsymbol{b}/\boldsymbol{a}]$ a substitution such that $\mathcal{K}\models_v \boldsymbol{a}[i]\approx\boldsymbol{b}[i]$ for each $1\leq i\leq|\boldsymbol{b}|$. If $\mathcal{K}\delta$ is a Hi(Horn-SROIQ) KB, then:*

- *$\mathcal{K}$ is v-satisfiable iff $\mathcal{K}\delta$ is v-satisfiable;*
- *$\boldsymbol{u}\in\mathrm{ans}_v(q,\mathcal{K})$ iff $\boldsymbol{u}\delta\in\mathrm{ans}_v(q\delta,\mathcal{K}\delta)$ for each conjunctive query q.*

Combing the above three lemmas, we can finally obtain the way of reducing v-satisfiability checking and conjunctive query answering in a Hi(Horn-SROIQ) to the corresponding reasoning tasks in a Horn-SROIQ KB with size no more than the original KB soundly and completely, showing in the following theorem.

Theorem 1 *For Hi(Horn-SROIQ) KB $\mathcal{K}$, the following three conclusions hold:*

- *$\widetilde{\mathcal{K}}$ is a Hi(Horn-SROIQ) KB;*
- *$\mathcal{K}$ is v-satisfiable iff $\tau(\widetilde{\mathcal{K}})$ is satisfiable;*
- *$\boldsymbol{u}\in\mathrm{ans}_v(q,\mathcal{K})$ iff $\boldsymbol{u}\varepsilon_{\mathcal{K}}\in\mathrm{ans}(\tau(q\varepsilon_{\mathcal{K}}),\tau(\widetilde{\mathcal{K}}))$ for each conjunctive query q.*

Example 2. For the KB $\mathcal{K}$ in example 1, by step 1 and step 2, we can obtain a Hi(Horn-SROIQ) KB $\widetilde{\mathcal{K}}=\mathcal{K}\varepsilon_{\mathcal{K}}$ where $\varepsilon_{\mathcal{K}}=[\texttt{Football\_team}/\texttt{FootballTeam}]$. By theorem 1, we can get that v-satisfiability checking and conjunctive query answering in $\mathcal{K}$ can be captured by $\widetilde{\mathcal{K}}$ soundly and completely.

3.2 Meta-query answering in Hi(Horn-SROIQ)

For a Hi(Horn-SROIQ) KB $\mathcal{K}$, analogous to lemma 3, we can get that meta-query answering in $\mathcal{K}$ can be captured by meta-query answering in the materialized KB $\widetilde{\mathcal{K}}$, showing in the following lemma.

Lemma 4 *For a Hi(Horn-SROIQ) KB $\mathcal{K}$ and meta-query Q, then we can get that $\boldsymbol{u} \in \mathrm{ans}_v(Q, \mathcal{K})$ iff $\boldsymbol{u}\varepsilon_{\mathcal{K}} \in \mathrm{ans}_v(Q\varepsilon_{\mathcal{K}}, \widetilde{\mathcal{K}})$.*

For a Hi(Horn-SROIQ) KB $\mathcal{K}$ and meta-query Q, a CP-binding ξ of Q over $\mathcal{K}$ is a function that maps each class variable of Q to $\top$ or a name used as class in $\mathcal{K}$ and maps each role variable of Q to a name used as role in $\mathcal{K}$. We use $Q\xi$ to denote the conjunctive query obtained by replacing each appearance of x in Q with $\xi(x)$ for each class (role) variable x of Q. The following theorem shows that meta-query answering over the materialized KB can be captured by conjunctive query answering through CP-bindings.

Theorem 2 *Let $\mathcal{K}$ be a v-satisfiable Hi(Horn-SROIQ) KB. Then for each meta-query Q, we can get that $\boldsymbol{u} \in \mathrm{ans}_v(Q, \widetilde{\mathcal{K}})$ iff there exists a CP-binding ξ of Q over $\widetilde{\mathcal{K}}$ such that $\boldsymbol{u} \in \mathrm{ans}_v(Q\xi, \widetilde{\mathcal{K}})$.*

Example 3. Consider the Hi(Horn-SROIQ) KB $\mathcal{K}$ in example 1 again. Suppose we want to know whether there exists a role that relates `BarcelonaDragous` to `FootballTeam`. This task can be realized by answering the query Q over $\mathcal{K}$:

$$\boxed{?p(\texttt{BarcelonaDragous}, \texttt{FootballTeam}) \rightarrow q()}$$

By theorem 2, lemma 4 and theorem 1, we can get that answering Q over $\mathcal{K}$, can be realized by evaluating the conjunctive query:

$$\boxed{\texttt{rewriteOf}(\texttt{BarcelonaDragous}, \texttt{FootballTeam}) \rightarrow q()}$$

over $\widetilde{\mathcal{K}}$ in terms of punning, since $Q\varepsilon_{\mathcal{K}}$ has just one CP-binding over $\widetilde{\mathcal{K}}$.

Furthermore, by theorem 2 and 1 as well as the results in [29], we can get that meta-query answering in Hi(Horn-SROIQ) has universal model feature [29].

Theorem 3 *Let $\mathcal{K}$ be a v-satisfiable Hi(Horn-SROIQ) KB. Then $\mathcal{K}$ has v-model $\mathcal{V}$ such that $\boldsymbol{u} \in \mathrm{ans}_v(Q, \mathcal{K})$ iff $\mathcal{V} \models_v Q[\boldsymbol{x}/\boldsymbol{u}]$ for each meta-query Q with head $\boldsymbol{x}$.*

4 Complexity of reasoning with Hi(Horn-SROIQ)

In this section, we investigate the complexity of reasoning with Hi(Horn-SROIQ). Before that, we first show the complexity of materializing Hi(Horn-SROIQ) KBs.

Theorem 4 *For Hi(Horn-SROIQ) KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, the procedure of computing $\widetilde{\mathcal{K}}$ terminates. Moreover $\widetilde{\mathcal{K}}$ can be obtained in $2\mathrm{ExpTime}$ w.r.t. $|\mathcal{K}|$ and PTime w.r.t. $|\mathcal{A}|$.*

Based on theorem 1 and theorem 4 as well as the results in [28], we can get the complexities of v-satisfiability checking in Hi(Horn-SROIQ).

Theorem 5 *For Hi(Horn-SROIQ) KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, v-satisfiability checking of $\mathcal{K}$ can be done in 2ExpTime w.r.t. $|\mathcal{K}|$ and PTime w.r.t. $|\mathcal{A}|$.*

For a Hi(Horn-SROIQ) KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and meta-query Q, theorem 2 indicates that answering Q over $\mathcal{K}$ needs to evaluate as many as $|\mathcal{K}|^{|Q|}$ conjunctive queries over a Horn-SROIQ KB with size no more than $\mathcal{K}$. Actually, in the appendix of this paper, we show that the number of queries need to be answered can be measured by the size of $\mathcal{T}$. Then based on the results in [29], we can obtain the complexity of answering Q over $\mathcal{K}$.

Theorem 6 *For Hi(Horn-SROIQ) KB $\mathcal{K}$ and meta-query Q, $\mathtt{ans}_v(Q, \mathcal{K})$ can be computed in 2ExpTime-complete w.r.t. $|\mathcal{K}|$ and PTime-complete w.r.t. $|\mathcal{A}|$.*

The above two theorems indicate that adding meta-modeling capability to Horn-SROIQ has no impact on the complexity of the considered reasoning tasks.

5 Related work

Here, we discuss other works on extending decidable DLs with meta-modeling.

[26] investigates extending SHOIQ with meta-modeling by unifying the sets of names for classes, roles and individuals. In [26], v-satisfiability checking of a SHOIQ KB can be realized by checking the satisfiability of as many as 2^{n^2} SHOIQ KBs where n denotes the number of different names in the original K-B. Thus satisfiability checking in the extended SHOIQ under v-semantics has NExpTime combined complexity. Excepting high reasoning complexity, query answering has not been discussed in [26]. On the other hand, [3] addresses extending SHIQ with meta-modeling under a stronger variant of the HiLog-style semantics where extensions are also assigned to complex classes and roles. However, punning is adopted to reason with the extended SHIQ with coNP-complete data complexity of both satisfiability checking and query answering. Compared with these two works, although some class axioms are not allowed in Hi(Horn-SROIQ), such as $A \sqcup B \sqsubseteq C$, Hi(Horn-SROIQ) supports richer role axioms, such as role chains ($\circ$) and role disjoint axioms, which are useful in medical terminologies. Most importantly, Hi(Horn-SROIQ) has much lower data complexities of satisfiability checking and meta-query answering. Extension of tractable language DL-Lite$_R$ is discussed in [3, 22]. Although low complexity can be guaranteed, the expressivity is every restricted.

Moreover, [11, 13, 24, 25, 30] study typed meta-modeling extension or extension based on Henkin semantics [12]. The referred DLs contain SHOIQ, SROIQ, ALCQ and SHIQ. In the typed situation, names are typed with layer or level information (non-negative integer numbers) with the intention to describe levels of classes and roles. However, describing axioms (assertions) referring names with different types is limited or not allowed. Furthermore, the actual OWL 2 Full

ontologies do not contain such way of meta-modeling. On the other hand, Henkin semantics, which deals with higher-order structures via a hierarchy of power sets, is stronger than HiLog-style semantics. In this semantics, the class extensions of names are the sets in the domains they mapped into. Thus the claim $\mathcal{KB} \models a \approx b$ iff $\mathcal{KB} \models a =_c b$ holds. However, for both OWL 2 RDF Based Semantics and HiLog-style semantics, just the "if-then" direction holds. Moreover, under this semantics, some undesired conclusions may be entailed. For example, from the knowledge (RDF triple (a), (b) and (c)) described in Linked Open Data:

`(geosp:Country, owl:equivalentClass, geo:Pays)`	(a)
`(geosp:Country, rdfs:isDefinedBy, geospecies.owl)`	(b)
`(geo:Pays, rdfs:isDefinedBy, geo−fr)`	(c)
`(geo:Pays, rdfs:isDefinedBy, geospecies.owl)`	(d)
`(geosp:Country, rdfs:isDefinedBy, geo−fr)`	(e)

under this semantics, we can obtain the conclusion (d) and (e). However, neither of them are desired.

Finally, we mention [7] and [21]. [7] discusses ontology-inherent meta-modeling for classes in OWL 2 DL by adding extra names, axioms and assertions. The limitation of this approach is that just two layers (classes and meta-classes) of meta-modeling can be supported. By combing and unifying some of the above mentioned approaches, [21] investigates a variant of higher-order DLs, containing (a) a fixly interpreted instanceOf role connecting individuals with classes; (b) promiscuous classes may have roles and other classes as individuals; (c) strictly typed classes allowing only a certain type of individuals.

6 Discussion and conclusions

In this paper, we discuss the problem of efficiently reasoning and query answering in expressive domain KBs that are in OWL 2 Full and contain large individual datasets. For this, we introduce Hi(Horn-SROIQ) and meta-queries, and provide concrete methods of reasoning with Hi(Horn-SROIQ). For the future research, our work can be extended in the following two aspects. Firstly, theorem 2 indicates that evaluating a meta-query Q over a Hi(Horn-SROIQ) KB $\mathcal{K}$ may need to answer a lot of conjunctive queries over a Horn-SROIQ KB. Although having low data complexity, heuristics, such as query partition, are needed to optimize such meta-query answering procedure. And secondly, the main motivation of this paper is to discuss the theoretic features of Hi(Horn-SROIQ), thus, experiments on actual OWL 2 Full ontologies, such as OpenCyc and SUMO, are needed to validate the overall provided method.

Acknowledgments

References

1. Bienvenu, M., Lutz, C., and Wolter, F. First-Order Rewritability of Atomic Queries in Horn Description Logics. *In Proc. IJCAI*, 754-760, AAAI Press, 2013.
2. Chen, W., Kifer, M., and Warren, D S. HILOG: a foundation for higher-order logic programming. *J. of Logic Programming*, 15(3):187-230, 1993.
3. De Giacomo, G., Lenzerini, M., and Rosati, R. Higher-Order Description Logics for Domain Metamodeling. *In Proc. AAAI*, 2011.
4. Eiter, T., Ortiz, M., Šimkus, M., Tran, T K., and Xiao, G. Towards Practical Query Answering for Horn-SHIQ⋆. *In Proc. DL*, 2012.
5. Faddoul, J., Farsinia, N., Haarslev, V., and Möller, R. A Hybrid Tableau Algorithm for ALCQ. *In Proc. DL*, 2008.
6. Glimm, B., and Ogbuji, C. SPARQL 1.1 Entailment Regimes. W3C Recommendation 21 March 2013, http://www.w3.org/TR/sparql11-entailment/, 2013.
7. Glimm, B., Rudolph, S., and Völker J. Integrated Metamodeling and Diagnosis in OWL 2. *In Proc. ISWC*, 257-272, 2010.
8. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., and Wang, Z. HermiT: An OWL 2 Reasoner. *J. of Automated Reasoning*, 53(3): 245-269, 2014.
9. Golbreich, C., and Wallace, E K. OWL 2 Web Ontology Language New Features and Rationale (Second Edition). W3C Recommendation 11 December 2012, http://www.w3.org/TR/owl2-new-features/, 2012.
10. Haarslev, V., Hidde, K., Möller, R., and Wessel, M. The RacerPro knowledge representation and reasoning system. *Semantic Web Journal*, 3(3):267-277, 2012.
11. Homola, M., Kl'uka, J., Svátek, V., and Vacura, M. Typed High-Order variant of SROIQ-Why Not? *In Proc. DL*, 2014.
12. Henkin, L. Completeness in the theory of types. *J. Symbolic Logic*, 15:81-91, 1950.
13. Homola, M., Kl'uka, J., Svátek, V, and Vacura, M. Towards Typed Higher-Order Description Logics. *In Proc. DL*, 2013.
14. Horrocks, I., Kutz, O., and Sattler, U. The Even More Irresistible SROIQ. *In Proc. KR*, 57-67, 2006.
15. Horrocks, I., and Sattler, U. A Tableau Decision Procedure for SHOIQ. *J. of Automated Reasoning*, 39(3):249-276, 2007.
16. Horrocks, I., and Sattler, U. Decidability of SHIQ with complex role inclusion axioms. *Artificial Intelligence*, 160(1):79-104, 2004.
17. Horrocks, I., Sattler, U., Tobies, S. Practical Reasoning for Very Expressive Description Logics. *Logic Journal of the IGPL*, 8(3): 239-263, 2000.
18. Ibañez-García, Y., Lutz, C., and Schneider, T. Finite Model Reasoning in Horn Description Logics. *In Proc. KR*, 2014.
19. Kazakov, Y. Consequence-driven reasoning for Horn-SHIQ ontologies. *In Proc. IJCAI*, 2040-2045, 2009.
20. Krötzsch, M., Rudolph, S., and Hitzler, P. Complexities of Horn Description Logics. *ACM Transactions on Computational Logic*, 14(1):2:1-2:36,, 2013.
21. Kubincová P, Kl'uka J, Homola M. Towards Expressive Metamodelling with Instantiation. *In Proc. DL*, 2015.
22. Lenzerini, M., Lepore, L., and Poggi, A. Practical query answering over Hi(DL-Lite$_R$) knowledge bases. *In Proc. DL*, 2014.
23. Matuszek, C., Cabral, J., Witbrock, M., and DeOliveira, J. An Introduction to the Syntax and Content of Cyc. *In AAAI Spring Symposium*, 2006.
24. Motz, R., Rohrer, E., and Severi, P. Reasoning for ALCQ Extended with a Flexible Meta-Modelling Hierarchy. *In Proc. JIST*, 47-62, 2014.

25. Motz, R., Rohrer, E., and Severi, P. The description logic SHIQ with a flexible meta-modeling hierarchy. *Web Semantics: Science, Services and Agents on the World Wide Web.* 35(4):214-234, 2015.

26. Motik, B. On the Properties of Metamodeling in OWL. *J. of Logic and Computation*, 17(4):617-637, 2007.

27. Noy, N F., Rubin, D L. Translating the foundational model of anatomy into OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(2):133-136, 2008.

28. Ortiz, M., Rudolph, S., and Šimkus, M. Worst-case Optimal Reasoning for the Horn-DL fragements of OWL 1 and OWL 2. *In Proc. KR*, 2010.

29. Ortiz, M., Rudolph, S., and Šimkus, M. Query Answering in the Horn Fragments of the Description Logics SHOIQ and SROIQ. *In Proc. IJCAI*, 2011.

30. Pan, J Z., and Horrocks, I. OWL FA: A Metamodeling Extension of OWL DL. *In Proc. WWW*, 1065-1066, 2006.

31. Pease, A., Niles, I., and Li, J. The suggested upper merged ontology: A large ontology for the semantic web and its applications. *In Proc. AAAI*, 2002.

32. Schneider, M. OWL 2 Web Ontology Language RDF-Based Semantics (Second Edition). W3C Recommendation 11 December 2012, http://www.w3.org/TR/owl2-rdf-based-semantics/, 2012.

33. Sirin, E., Parsia, B., Grau, B C., Kalyanpur, A., and Katz, Y. Pellet: A practical OWL-DL reasoner. *J. of Web Semantics* 5(2): 51-53, 2007.