

Attributed Description Logics: Reasoning on Knowledge Graphs*

Markus Krötzsch¹, Maximilian Marx¹, Ana Ozaki¹, and Veronika Thost²

¹ Center for Advancing Electronics Dresden (cfaed), TU Dresden, Dresden, Germany

² MIT-IBM Watson AI Lab, IBM Research

markus.kroetzsch@tu-dresden.de, maximilian.marx@tu-dresden.de,

ana.ozaki@tu-dresden.de, veronika.thost@ibm.com

Abstract

In modelling real-world knowledge, there often arises a need to represent and reason with meta-knowledge. To equip description logics (DLs) for dealing with such ontologies, we enrich DL concepts and roles with finite sets of attribute–value pairs, called annotations, and allow concept inclusions to express constraints on annotations. We investigate a range of DLs starting from the lightweight description logic \mathcal{EL} , covering the prototypical \mathcal{ALCH} , and extending to the very expressive \mathcal{SROIQ} , the DL underlying OWL 2 DL.

1 Introduction

Description logics (DLs) can express complex schema information on graph-like models, while supporting incomplete information via the open world assumption. While OWL is often used in RDF-based knowledge graphs developed in academia, such as DBpedia [Bizer *et al.*, 2009] and Bio2RDF [Belleau *et al.*, 2008], it has almost no impact on other applications of graph-structured data. This might in part be due to a format mismatch. Like DLs, many knowledge graphs use directed, labelled graph models, but unlike DLs they often add (*sets of*) annotations to vertices and edges. For example, the fact that Lise Meitner received the Max Planck Medal can be described by an assertion `awarded(meitner, planck_medal)`, but in practice we may also wish to record the year and the location. We may write this as follows:

`awarded(meitner, planck_medal)@[year:1949, loc:berlin]`.

Such annotated graph edges today are widespread in practice. Prominent representatives include *Property Graph*, the data model used in many graph databases [Rodriguez and Neubauer, 2010], and *Wikidata*, the knowledge graph used by Wikipedia [Vrandečić and Krötzsch, 2014]. One can encode annotated (multi-)graphs as directed graphs, e.g., using reification [Erxleben *et al.*, 2014], but DLs cannot express much over such models. For example, one cannot say that the spouse relation is symmetric, where annotations are the same

in both directions [Marx *et al.*, 2017]. Other traditional KR formalisms are similarly challenged in this situation.

In a recent work, we have therefore proposed to develop logics that support sets of attribute–value annotations natively [Marx *et al.*, 2017]. The according generalisation of first-order logic, called *multi-attribute predicate logic* (MAPL), is expressive enough to capture weak second-order logic, making reasoning non-semi-decidable. For that reason, we have developed the Datalog-like *MAPL rule language* (MARPL) as a decidable fragment.

In this paper, we explore the use of DLs as a basis for decidable, and even tractable, fragments of MAPL. The resulting family of *attributed DLs* allows statements such as `spouse@X ⊆ spouse^-@X` to say that spouse is symmetric; `spouse^-` denotes the inverse spouse relation. We introduce set variables (X in the example) to refer to annotations. References to variables are used to express constraints over annotations and to compare attribute values between them. A challenge is to add functionality of this type without giving up the nature of a DL, i.e., without sacrificing the variable-free notation of DLs in favour of rule languages. Another challenge is that these extensions may greatly increase the complexity of DLs.

We show that reasoning for attributed \mathcal{ELH} , a DL close to OWL 2 EL, is ExpTime -complete, and identify syntactic constraints under which we recover the PTime -completeness of \mathcal{ELH} . We also establish complexity results for the attributed variant of the prototypical DL \mathcal{ALCH} : reasoning is 2ExpTime -complete in general, and ExpTime -complete under our syntactic constraints. Finally, we show that reasoning is N2ExpTime -complete for attributed \mathcal{SROIQ} , the DL underlying OWL 2 DL. Additional proofs can be found in [Krötzsch *et al.*, 2017].

2 Attributed Description Logics

We introduce attributed description logics by defining the syntax and semantics of attributed \mathcal{ALCH} , denoted $\mathcal{ALCH}_@$. This allows us to illustrate the central ideas without having to deal with the full generality of \mathcal{SROIQ} , which we introduce in Section 4. We note that fact entailment can be polynomially reduced to satisfiability checking in the DLs we study.

2.1 Syntax and Intuition

We first give the syntax and intuitive semantics of $\mathcal{ALCH}_@$; the semantics will be formalised thereafter.

*This invited contribution to IJCAI 2018 summarises and discusses results of our previous ISWC 2017 paper “Attributed description logics: Ontologies for knowledge graphs” [Krötzsch *et al.*, 2017].

Example 1 We start with a guiding example, which will be formally explained when we define $\mathcal{ALCH}_@$. Wikidata contains assertions of the form:

awarded(a_person, an_award)@[year : a_year, url : a_url].

This motivates the following $\mathcal{ALCH}_@$ axiom.

$$\exists \text{awarded}@X.\text{Award} \sqsubseteq \exists \text{known_for}@[\text{year}:X.\text{year}].\top \quad (1)$$

The underlying DL axiom $\exists \text{awarded}.\text{Award} \sqsubseteq \exists \text{known_for}.\top$ states that anybody who received an award is known for some achievement.

Attributed DLs are defined over the usual DL signature with sets of *concept names* N_C , *role names* N_R , and *individual names* N_I . In OWL terminology, concepts correspond to classes, roles correspond to properties, and individual names correspond to individuals. We consider an additional set N_V of (*set*) *variables*. Following the definition of multi-attributed predicate logic (MAPL, [Marx et al., 2017]), we define annotation sets as finite binary relations, understood as sets of attribute–value pairs. In particular, *attributes* refer to domain elements and are syntactically denoted by individual names. To describe annotation sets, we introduce *specifiers*. The set \mathbf{S} of specifiers contains the following expressions:

- set variables $X \in N_V$;
- *closed specifiers* $[a_1 : v_1, \dots, a_n : v_n]$; and
- *open specifiers* $[a_1 : v_1, \dots, a_n : v_n]$,

where $a_i \in N_I$ and v_i is either an individual name in N_I or an expression of the form $X.c$, with X a set variable in N_V and c an individual name in N_I . Intuitively, closed specifiers define specific annotation sets whereas open specifiers merely provide lower bounds. A *ground specifier* is a specifier that does not contain expressions of the form $X.c$.

Example 2 The open specifier $[\text{year}:X.\text{year}]$ in Example 1 describes all those annotation sets that contain at least every attribute–value pair with attribute *year* that occurs in X .

The set \mathbf{R} of $\mathcal{ALCH}_@$ *role expressions* contains all expressions $r@S$ with $r \in N_R$ and $S \in \mathbf{S}$. The set \mathbf{C} of $\mathcal{ALCH}_@$ *concept expressions* is defined as follows.

$$\mathbf{C} ::= \top \mid \perp \mid N_C @ \mathbf{S} \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \exists R.C \mid \forall R.C$$

An $\mathcal{ALCH}_@$ *concept* (or *role*) *assertion* is an expression $A(a)@S$ (or $r(a, b)@S$), with $A \in N_C$ (or $r \in N_R$), $a, b \in N_I$, and $S \in \mathbf{S}$ a specifier that is not a set variable. An $\mathcal{ALCH}_@$ *concept inclusion* is an expression of the form

$$X_1 : S_1, \dots, X_n : S_n \quad (C \sqsubseteq D), \quad (2)$$

where $S_1, \dots, S_n \in \mathbf{S}$ are specifiers, $C, D \in \mathbf{C}$ are $\mathcal{ALCH}_@$ concept expressions, and $X_1, \dots, X_n \in N_V$ are set variables occurring in C, D or in S_1, \dots, S_n . $\mathcal{ALCH}_@$ *role inclusions* are defined analogously, but with role expressions instead of the concept expressions. An $\mathcal{ALCH}_@$ *ontology* is a set of $\mathcal{ALCH}_@$ assertions, and role and concept inclusions.

Example 3 Let α be the concept inclusion (1) from Example 1. Using a specifier for X , we can restrict the applicability of this axiom, e.g., to prizes awarded by the Nobel Foundation:

$$X : [\text{conferred_by}: \text{nobel_foundation}] (\alpha) \quad (3)$$

To simplify notation, we omit the specifier $[\]$ (meaning “any annotation set”) in role or concept expressions, as done for Award in Example 1. In this sense, any $\mathcal{ALCH}_@$ axiom is also an $\mathcal{ALCH}_@$ axiom. We also omit prefixes of the form $X : [\]$, allowing X to be any annotation set, as done in (1).

We follow the usual DL notation for referring to other attributed DLs. In particular, $\mathcal{ELH}_@$ is the fragment of $\mathcal{ALCH}_@$ which disallows negation.

2.2 Formal Semantics

An *interpretation* $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ of attributed logic consists of a non-empty domain $\Delta^{\mathcal{I}}$ and a function $\cdot^{\mathcal{I}}$. Individual names $a \in N_I$ are interpreted as elements $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$. To interpret annotation sets, we use the set $\Phi^{\mathcal{I}} := \{\Sigma \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid \Sigma \text{ is finite}\}$ of all finite binary relations over $\Delta^{\mathcal{I}}$. Now each concept name $C \in N_C$ is interpreted as a set $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Phi^{\mathcal{I}}$ of elements with annotations, and each role name $r \in N_R$ is interpreted as a set $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \times \Phi^{\mathcal{I}}$ of pairs of elements with annotations. Note that each element (pair of elements) may appear with multiple different annotations.

Expressions with free set variables are interpreted using variable assignments $\mathcal{Z} : N_V \rightarrow \Phi^{\mathcal{I}}$. A specifier $S \in \mathbf{S}$ is interpreted as a set $S^{\mathcal{I}, \mathcal{Z}} \subseteq \Phi^{\mathcal{I}}$ of matching annotation sets. We set $X^{\mathcal{I}, \mathcal{Z}} := \{\mathcal{Z}(X)\}$ for variables $X \in N_V$. The semantics of closed specifiers is defined as follows:

- (i) $[a : b]^{\mathcal{I}, \mathcal{Z}} := \{\{\langle a^{\mathcal{I}}, b^{\mathcal{I}} \rangle\}\}$
- (ii) $[a : X.b]^{\mathcal{I}, \mathcal{Z}} := \{\{\langle a^{\mathcal{I}}, \delta \rangle \mid \langle b^{\mathcal{I}}, \delta \rangle \in \mathcal{Z}(X)\}\}$
- (iii) $[a_1 : v_1, \dots, a_n : v_n]^{\mathcal{I}, \mathcal{Z}} := \{\cup_{i=1}^n F_i\}$ where $\{F_i\} = [a_i : v_i]^{\mathcal{I}, \mathcal{Z}}$ for all $i \in \{1, \dots, n\}$.

$S^{\mathcal{I}, \mathcal{Z}}$ therefore is a singleton set for variables and closed specifiers. For open specifiers, however, we define $[a_1 : v_1, \dots, a_n : v_n]^{\mathcal{I}, \mathcal{Z}}$ to be the set

$$\{F \in \Phi^{\mathcal{I}} \mid F \supseteq G \text{ for } \{G\} = [a_1 : v_1, \dots, a_n : v_n]^{\mathcal{I}, \mathcal{Z}}\}.$$

Now, given $A \in N_C$, $r \in N_R$, and $S \in \mathbf{S}$, we define:

$$(A@S)^{\mathcal{I}, \mathcal{Z}} := \{\delta \mid \langle \delta, F \rangle \in A^{\mathcal{I}} \text{ for some } F \in S^{\mathcal{I}, \mathcal{Z}}\},$$

$$(r@S)^{\mathcal{I}, \mathcal{Z}} := \{\langle \delta, \epsilon \rangle \mid \langle \delta, \epsilon, F \rangle \in r^{\mathcal{I}} \text{ for some } F \in S^{\mathcal{I}, \mathcal{Z}}\}.$$

Further DL expressions are defined as usual: $\top^{\mathcal{I}, \mathcal{Z}} = \Delta^{\mathcal{I}}$, $\neg C^{\mathcal{I}, \mathcal{Z}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}, \mathcal{Z}}$, $(C \sqcap D)^{\mathcal{I}, \mathcal{Z}} = C^{\mathcal{I}, \mathcal{Z}} \cap D^{\mathcal{I}, \mathcal{Z}}$, and $(\exists R.C)^{\mathcal{I}, \mathcal{Z}} = \{\delta \mid \text{there is } \langle \delta, \epsilon \rangle \in R^{\mathcal{I}, \mathcal{Z}} \text{ with } \epsilon \in C^{\mathcal{I}, \mathcal{Z}}\}$.

Now \mathcal{I} satisfies an $\mathcal{ALCH}_@$ concept inclusion α of the form (2), written $\mathcal{I} \models \alpha$, if, for all variable assignments \mathcal{Z} such that $\mathcal{Z}(X_i) \in S_i^{\mathcal{I}, \mathcal{Z}}$ for all $i \in \{1, \dots, n\}$, we have $C^{\mathcal{I}, \mathcal{Z}} \subseteq D^{\mathcal{I}, \mathcal{Z}}$. Satisfaction of role inclusions is defined analogously. Moreover, \mathcal{I} satisfies an $\mathcal{ALCH}_@$ concept assertion $A(a)@S$ if $\langle a^{\mathcal{I}}, \Psi \rangle \in A^{\mathcal{I}}$ for some $\Psi \in S^{\mathcal{I}}$ (the latter is well-defined since S contains no variables), and likewise for role assertions. \mathcal{I} satisfies an ontology if it satisfies all of its axioms. Based on this model theory, logical entailment is defined as usual.

Example 4 Consider the concept inclusion α of Example 1 and the following interpretation \mathcal{I} over domain $\Delta^{\mathcal{I}} = \{\text{meitner}, \text{planck_medal}, 1949, \text{year}, \text{nuclear_fission}\}$ with

$$\text{Award}^{\mathcal{I}} = \{\text{planck_medal}\},$$

$$\text{awarded}^{\mathcal{I}} = \{\langle \text{meitner}, \text{planck_medal}, \{\langle \text{year}, 1949 \rangle\} \rangle\}, \text{ and}$$

$$\text{known_for}^{\mathcal{I}} = \{\langle \text{meitner}, \text{nuclear_fission}, \{\langle \text{year}, 1949 \rangle\} \rangle\}.$$

Then $\mathcal{I} \models \alpha$, i.e., \mathcal{I} satisfies α .

3 Tractable Reasoning in Attributed \mathcal{ELH}

We now introduce the two central reasoning techniques for attributed DLs: *grounding* eliminates set variables from an attributed ontology, and *renaming* constructs an equisatisfiable ontology in the underlying non-attributed DL.

Let KB be a ground $\mathcal{ELH}_@$ ontology, i.e., an $\mathcal{ELH}_@$ ontology that does not contain set variables. We translate KB into an \mathcal{ELH} ontology KB^\dagger by replacing every concept expression $A@S$ with $A \in N_C$ and $S \in \mathbf{S}$ occurring in KB by a fresh concept name A_S , and similarly for role expressions $r@S$. To capture the interactions between specifiers such as $[a : b]$ and $[\]$, we extend KB^\dagger by concept and role inclusion axioms $A_S \sqsubseteq A_T$ and $r_S \sqsubseteq r_T$, respectively, where A_S and A_T (or r_S and r_T , respectively) occur in KB^\dagger and T is an open specifier *subsuming* the (open or closed) specifier S , i.e., S contains all attribute–value pairs $a : b$ in T .

Then KB^\dagger is polynomial in the size of KB , and together with the PTIME-completeness of \mathcal{ELH} [Baader *et al.*, 2005], we obtain the following theorem.

Theorem 1 *Let KB be a ground $\mathcal{ELH}_@$ ontology. Checking satisfiability of KB is PTIME-complete.*

For a general $\mathcal{ELH}_@$ ontology KB , the *grounding* $\text{ground}(KB)$ is the ground $\mathcal{ELH}_@$ ontology consisting of all the assertions of KB , and, for each inclusion axiom $X_1 : S_1, \dots, X_n : S_n$ ($C \sqsubseteq D$) by all possible ground instances $C' \sqsubseteq D'$: concept expressions $A@X_i$ are replaced by $A@T_i$, where T_i is some specifier subsumed by S_i , and assignments $a : X_i.b$ are replaced by all assignments $a : c$ such that $b : c$ occurs in T_i . Role inclusion axioms are treated analogously.

In general, there may be exponentially many ground instances of a single inclusion axiom. Yet, tractable reasoning in $\mathcal{ELH}_@$ is possible under further restrictions, as the next theorem shows.

Theorem 2 *Let $\ell, k \in \mathbb{N}$, and let KB be an $\mathcal{ELH}_@$ ontology satisfying all of the following conditions:*

- (A) *each axiom of KB contains at most ℓ set variables,*
- (B) *each specifier occurring in KB contains at most k expressions of the form $X.a$, and,*
- (C) *if a specifier S contains an assignment $a : X.b$, no further assignment for the same attribute a occurs in S .*

Then checking satisfiability of KB is PTIME-complete.

Together, these conditions guarantee that the grounding $\text{ground}(KB)$ remains polynomial in the size of KB . Indeed, relaxing any of the conditions immediately yields intractability.

Theorem 3 *Let KB be an $\mathcal{ELH}_@$ ontology and consider conditions (A)–(C) of Theorem 2 with $\ell = 1$ and $k = 2$. Then deciding satisfiability of KB is*

- (1) *EXPTIME-complete if KB violates only condition (A),*
- (2) *EXPTIME-complete if KB violates only condition (B), and*
- (3) *PSPACE-hard if KB violates only condition (C).*

It remains open whether the PSPACE-bound in the third case is tight. Nevertheless, reasoning in this case is intractable.

The same pattern repeats for more expressive DLs: satisfiability of ground ontologies retains the complexity of the underlying non-attributed DL, and the grounding technique results in an exponentially larger ontology.

Theorem 4 *Consider conditions (A)–(C) of Theorem 2. Checking satisfiability of an $\mathcal{ALCH}_@$ ontology KB is*

- (1) *EXPTIME-complete if KB is ground,*
- (2) *EXPTIME-complete if KB satisfies conditions (A)–(C), and*
- (3) *2EXPTIME-complete in general.*

4 Attributed $SRIOQ$

$SRIOQ$ extends \mathcal{ALCH} with nominals $\{c\}$ (concepts with just one individual), number restrictions $\leq n R.C$ and $\geq n R.C$ (bounding the number of R -successors in C), inverse roles r^- , and generalised role inclusions $R \circ S \sqsubseteq T$ (generalising transitivity). Since axioms such as $\{c\} \equiv A_c @ [\]$ can be used to introduce a concept name A_c for any nominal c , we do not add annotations on nominals. This allows us to reuse the notion of interpretation from $\mathcal{ALCH}_@$. We lift the definitions of assertions, and of concept and role inclusions by extending them to include the new constructors. Furthermore, $SRIOQ_@$ allows complex role inclusion axioms of the following form:

$$X_1 : S_1, \dots, X_n : S_n \quad (R_1 \circ \dots \circ R_\ell \sqsubseteq T), \quad (4)$$

where $R_i, T \in \mathbf{R}$ are $SRIOQ_@$ role expressions, $S_1, \dots, S_n \in \mathbf{S}$ are specifiers, and $X_1, \dots, X_n \in N_V$ are set variables occurring in R_i, T, S_1, \dots, S_n . As before, a $SRIOQ_@$ ontology is a set consisting of $SRIOQ_@$ assertions, and of role and concept inclusion axioms. We extend the usual semantics [Horrocks *et al.*, 2006] to $SRIOQ_@$ as for $\mathcal{ALCH}_@$ by accounting for annotations in the interpretation of concepts and roles.

Example 5 *$SRIOQ$ can express multiplicities of relations, e.g., to state that an award with more than one recipient in one particular year is a shared award for this year:*

$$\geq 2 \text{ awarded}^- @ X.T \sqsubseteq \text{Shared_Award} @ [\text{year} : X.\text{year}].$$

The following example uses a nominal (without annotations, as explained above):

$$\exists \text{ awarded}^- @ [\text{conferred_by} : \text{nobel_foundation}, \text{loc} : \text{oslo}].T \sqsubseteq \{\text{nobel_peace_prize}\}.$$

This can be viewed as an annotation-dependent range restriction axiom, in this case defining the range of awarded to be a singleton class.

$SRIOQ$ ensures decidability of reasoning by imposing two additional restrictions on ontologies – *simplicity* and *regularity* [Horrocks *et al.*, 2006] – which we adapt to $SRIOQ_@$. Simple roles are defined as in $SRIOQ$, ignoring all annotations. They are the only roles that may appear in the role expressions of number restrictions. A $SRIOQ_@$ ontology is *regular* if the set $N_R^\pm = N_R \cup \{r^- \mid r \in N_R\}$ can be ordered by a strict partial order $<$ such that

- (1) $r < t$ iff $r^- < t$ for all $t \in N_R^\pm$ and all $r \in N_R$, and

(2) the inclusion $R_1 \circ \dots \circ R_\ell \sqsubseteq T$ in each complex role inclusion axiom (4) is of one of the following forms:

$$\begin{aligned} t@S \circ t@S \sqsubseteq t@S, & \quad R_1 \circ \dots \circ R_{\ell-1} \circ t@S \sqsubseteq t@S, \\ R_1 \circ \dots \circ R_\ell \sqsubseteq t@S, & \quad t@S \circ R_2 \circ \dots \circ R_\ell \sqsubseteq t@S, \\ r^-@S \sqsubseteq r@S, & \end{aligned}$$

where $S \in \mathbf{S}$, $t \in \mathbf{N}_R^\pm$, $r \in \mathbf{N}_R$, and $R_1, \dots, R_\ell \in \mathbf{R}$ are of the form $r_1@S_1, \dots, r_\ell@S_\ell$ with $r_i < t$ for all $1 \leq i \leq \ell$.

In addition to adopting the usual conditions from $SRQIQ$ for (inverted) role names, we also require that, whenever the same role name occurs on both sides of the inclusion, also the specifiers are the same in both cases. Regularity can be verified in polynomial time by computing a minimal relation $<$ satisfying the conditions and then checking whether $<$ is a strict partial order.

In regards to reasoning, the new features of $SRQIQ$ over $ALCH$ lead to several difficulties. In the presence of nominals and number restrictions, equalities $a \approx b$ of individuals may be entailed. In turn, this leads to equalities of annotation sets, e.g., $A@[c : a] \equiv A@[c : b]$. Furthermore, nominals can be used to bound the overall size of the domain, e.g., by stating $\top \sqsubseteq \{a\}$. This gives rise to further equalities, but, more importantly, it also affects the semantics of open specifiers (e.g., we then have $A@[a : a] \sqsubseteq A@[a : a]$).

Our translation of $SRQIQ$ ontologies thus requires further axioms to handle these two effects. A naive translation, however, requires exponentially many auxiliary axioms to handle either of them, which would lead to an $N3\text{ExpTime}$ upper bound even for ground $SRQIQ$. Instead, we use an auxiliary DL $SRQIQ_\approx$, which allows for a more concise translation. $SRQIQ_\approx$ can in turn be reduced to \mathcal{C}^2 , the two-variable fragment with counting [Pratt-Hartmann, 2005], in exponential time, which yields an $N2\text{ExpTime}$ upper bound for ground $SRQIQ$. Even though the grounding of a $SRQIQ$ ontology still results in an exponentially larger ontology, this does not increase the complexity any further.

Besides standard $SRQIQ$ axioms, $SRQIQ_\approx$ also supports concept inclusion of the form $a \approx b \Rightarrow C \sqsubseteq D$ and role inclusions of the form $a \approx b \Rightarrow R_1 \circ \dots \circ R_\ell \sqsubseteq T$. An interpretation \mathcal{I} satisfies such an axiom $a \approx b \Rightarrow \alpha$ when either $a^{\mathcal{I}} \neq b^{\mathcal{I}}$ or $\mathcal{I} \models \alpha$. In translating a ground $SRQIQ$ ontology KB into a $SRQIQ_\approx$ ontology KB^\ddagger , we proceed as before, but introduce names $A_S \in \mathbf{N}_C$ and $r_S \in \mathbf{N}_R$ for all possible open and closed specifiers over the individual names occurring in KB . Then, to handle equality, for every $A \in \mathbf{N}_C$, we add axioms $a \approx b \Rightarrow A_S \sqsubseteq A_T$ for all pairs S, T of distinct ground specifiers that are either both open or both closed, and where S and T are the same when replacing every occurrence of a by b . We add similar axioms for each role name $r \in \mathbf{N}_R$.

To deal with the bounded domain size, we use an individual name z not occurring in KB as a canary: equalities of the form $z \approx a$ allow us to detect the bounded domain case. We let \mathbf{N}_I^{KB} be the set of all individual names in KB , and add, for each $a, b, c \in \mathbf{N}_I^{KB}$ and all $A \in \mathbf{N}_C$ and all $r \in \mathbf{N}_R$ in KB , axioms

$$z \approx a \Rightarrow \top \sqsubseteq \bigsqcup_{c \in \mathbf{N}_I^{KB}} \{c\}, \quad (5)$$

$$z \approx a \Rightarrow A_S \sqsubseteq \bigsqcup_{T \supseteq_c S} A_T, \quad \text{and} \quad (6)$$

$$z \approx a \Rightarrow \{b\} \sqcap \exists r_S.\{c\} \sqsubseteq \bigsqcup_{T \supseteq_c S} \exists r_T.\{c\}, \quad (7)$$

where S is a ground open specifier and $T \supseteq_c S$ means that T is a ground closed specifier subsuming S . Intuitively, (7) states that any fact $r_S(b, c)$ entails some fact of the form $r_T(b, c)$.

Together with the auxiliary axioms from the $ALCH$ translation, we obtain KB^\ddagger , an equisatisfiable $SRQIQ_\approx$ ontology. The construction can be performed in exponential time.

The grounding technique does not increase the $<$ -depth of the ontology KB , and the above translation of ground $SRQIQ$ ontologies into $SRQIQ_\approx$ is polynomial in the number of possible ground annotation sets. Thus, we find that $\text{ground}(KB)^\ddagger$ is still only exponentially larger than KB , even for non-ground $SRQIQ$ ontologies. Together with the bounded $<$ -depth, we thus obtain an $N2\text{ExpTime}$ upper bound.

Theorem 5 *Checking satisfiability of $SRQIQ$ ontologies is $N2\text{ExpTime}$ -complete.*

5 Discussion and Outlook

Current graph-based knowledge representation formalisms suffer from an inability to handle meta-data in the form of sets of attribute–value pairs. These limitations show up even when dealing with purely abstract data and are orthogonal to datatype support in the formalisms. Our family of attributed description logics approaches the problem in the context of DLs. We have presented a grounding-based decision procedure and identified the special cases of ground ontologies and structural restrictions on axioms, for which this overhead can be avoided. Those restrictions ensure tractability of attributed \mathcal{EL} .

Related approaches. Our goal was to bring features of the recent attributed rule language $MARPL$ to DLs [Marx *et al.*, 2017]. In practice, further rule-like expressivity is probably needed, and DL-rule extensions should be considered. Our grounding approach naturally fits to DL-safe rules [Motik *et al.*, 2005] and nominal schemas [Krötzsch *et al.*, 2011], but many other DL-rule mergers exist. We are not aware of any other extension of DL to property graphs. However, \mathcal{DLR}^+ , a recent extension of \mathcal{DLR} [Calvanese *et al.*, 2008], extends DL to relational data models [Artale *et al.*, 2017]. Our annotation sets can be encoded in \mathcal{DLR}^+ using auxiliary relations (just like in RDF), but it remains open how the expressive powers compare under such an encoding. Semiring semantics can also express certain “annotations” [Green *et al.*, 2007], e.g., in annotated RDF [Straccia *et al.*, 2010]. Many approaches could be re-phrased using semirings, but this yields neither a concrete syntax nor specific insights, e.g., on decidability and complexity.

Outlook. We are currently working on developing reasoning support for attributed logics. A variant of $MARPL$ rules is used in ontological modelling for Wikidata; inferred facts are shown in the $SQID$ data browser (<https://tools.wmflabs.org/sqid/>, [Marx and Krötzsch, 2017]). This rule language was extended

with expressions of the form $X.a$, introduced herein. Future research will be necessary to design attributed logics with the right level of expressivity for such applications, and to develop general-purpose reasoning tools to fully support them.

Acknowledgements

This work is partly supported by the German Research Foundation (DFG) within the Cluster of Excellence “Center for Advancing Electronics Dresden” (cfaed), the Collaborative Research Center SFB 912 (HAEC), and in Emmy Noether grant KR 4381/1-1 (DIAMOND).

References

- [Artale *et al.*, 2017] Alessandro Artale, Enrico Franconi, Rafael Peñaloza, and Francesco Sportelli. A decidable very expressive description logic for databases. In *The Semantic Web - ISWC 2017 - 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings, Part I*, pages 37–52, 2017.
- [Baader *et al.*, 2005] Franz Baader, Sebastian Brandt, and Carsten Lutz. Pushing the \mathcal{EL} envelope. In Leslie Pack Kaelbling and Alessandro Saffiotti, editors, *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI'05)*, pages 364–369. Professional Book Center, 2005.
- [Belleau *et al.*, 2008] François Belleau, Marc-Alexandre Nolin, Nicole Tourigny, Philippe Rigault, and Jean Morissette. Bio2RDF: Towards a mashup to build bioinformatics knowledge systems. *J. of Biomedical Inf.*, 41(5):706–716, 2008.
- [Bizer *et al.*, 2009] Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. DBpedia – A crystallization point for the Web of Data. *J. of Web Semantics*, 7(3):154–165, 2009.
- [Calvanese *et al.*, 2008] Diego Calvanese, Giuseppe De Giacomo, and Maurizio Lenzerini. Conjunctive query containment and answering under description logic constraints. *ACM Trans. Comput. Log.*, 9(3):22:1–22:31, 2008.
- [Erxleben *et al.*, 2014] Fredo Erxleben, Michael Günther, Markus Krötzsch, Julian Mendez, and Denny Vrandečić. Introducing Wikidata to the linked data web. In *Proc. 13th Int. Semantic Web Conf. (ISWC'14)*, volume 8796 of LNCS, pages 50–65. Springer, 2014.
- [Green *et al.*, 2007] Todd J. Green, Gregory Karvounarakis, and Val Tannen. Provenance semirings. In *Proceedings of the Twenty-Sixth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 11-13, 2007, Beijing, China*, pages 31–40, 2007.
- [Horrocks *et al.*, 2006] Ian Horrocks, Oliver Kutz, and Ulrike Sattler. The even more irresistible *SR_{OLQ}*. In Patrick Doherty, John Mylopoulos, and Christopher A. Welty, editors, *Proc. 10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'06)*, pages 57–67. AAAI Press, 2006.
- [Krötzsch *et al.*, 2011] Markus Krötzsch, Frederick Maier, Adila A. Krisnadhi, and Pascal Hitzler. A better uncle for OWL: Nominal schemas for integrating rules and ontologies. In *Proc. 20th Int. Conf. on World Wide Web (WWW'11)*, pages 645–654. ACM, 2011.
- [Krötzsch *et al.*, 2017] Markus Krötzsch, Maximilian Marx, Ana Ozaki, and Veronika Thost. Attributed description logics: Ontologies for knowledge graphs. In *The Semantic Web - ISWC 2017 - 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings, Part I*, pages 418–435, 2017.
- [Marx and Krötzsch, 2017] Maximilian Marx and Markus Krötzsch. SQID: Towards ontological reasoning for Wikidata. In Nadeschda Nikitina and Dezhao Song, editors, *Proceedings of the ISWC 2017 Posters & Demonstrations Track*, CEUR Workshop Proceedings. CEUR-WS.org, October 2017.
- [Marx *et al.*, 2017] Maximilian Marx, Markus Krötzsch, and Veronika Thost. Logic on MARS: Ontologies for generalised property graphs. In Carles Sierra, editor, *Proc. 26th Int. Joint Conf. on Artificial Intelligence (IJCAI'17)*, pages 1188–1194. International Joint Conferences on Artificial Intelligence, 2017.
- [Motik *et al.*, 2005] Boris Motik, Ulrike Sattler, and Rudi Studer. Query answering for OWL DL with rules. *J. of Web Semantics*, 3(1):41–60, 2005.
- [Pratt-Hartmann, 2005] Ian Pratt-Hartmann. Complexity of the two-variable fragment with counting quantifiers. *J. of Logic, Language and Information*, 14:369–395, 2005.
- [Rodriguez and Neubauer, 2010] Marko A. Rodriguez and Peter Neubauer. Constructions from dots and lines. *Bulletin of the American Society for Information Science and Technology*, 36(6):35–41, 2010.
- [Straccia *et al.*, 2010] Umberto Straccia, Nuno Lopes, Gergely Lukacsy, and Axel Polleres. A general framework for representing and reasoning with annotated Semantic Web data. In Maria Fox and David Poole, editors, *Proc. 24th AAAI Conf. on Artificial Intelligence (AAAI'10)*. AAAI Press, 2010.
- [Vrandečić and Krötzsch, 2014] Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. *Commun. ACM*, 57(10), 2014.