



An Optimized KE-Tableau-Based System for Reasoning in the Description Logic $\mathcal{DL}_D^{4,\times}$

Domenico Cantone, Marianna Nicolosi-Asmundo,
and Daniele Francesco Santamaria^(✉)

Department of Mathematics and Computer Science,
University of Catania, Catania, Italy
{cantone,nicolosi,santamaria}@dmi.unict.it

Abstract. We present a KE-tableau-based procedure for the main TBox and ABox reasoning tasks for the description logic $\mathcal{DL}\langle 4\text{LQS}^{\text{R},\times}\rangle(\mathbf{D})$, in short $\mathcal{DL}_D^{4,\times}$. The logic $\mathcal{DL}_D^{4,\times}$, representable in the decidable multi-sorted quantified set-theoretic fragment 4LQS^{R} , combines the high scalability and efficiency of rule languages such as the Semantic Web Rule Language (SWRL) with the expressivity of description logics.

Our algorithm is based on a variant of the KE-tableau system for sets of universally quantified clauses, where the KE-elimination rule is generalized in such a way as to incorporate the γ -rule. The novel system, called KE $^\gamma$ -tableau, turns out to improve both the system introduced in [3] and the standard first-order KE-tableaux [10]. Suitable benchmark test sets executed on C++ implementations of the three mentioned systems show that the performances of the KE $^\gamma$ -tableau-based reasoner are often up to about 400% better than the ones of the other two systems. This a first step towards the construction of efficient reasoners for expressive OWL ontologies based on fragments of computable set theory.

1 Introduction

Recently, decidability results in Computable Set Theory have been used for knowledge representation and reasoning, in particular, in the context of description logics (DLs) and rule languages for the Semantic Web. These efforts are motivated by the fact that there exists a natural translation function between set-theoretic fragments and languages for the Semantic Web.

In particular, the decidable four-level stratified set-theoretic fragment 4LQS^{R} , involving variables of four sorts, pair terms, and a restricted form of quantification over variables of the first three sorts (cf. [1]), has been used in [2] to represent the description logic $\mathcal{DL}\langle 4\text{LQS}^{\text{R},\times}\rangle(\mathbf{D})$, in short $\mathcal{DL}_D^{4,\times}$.

The description logic $\mathcal{DL}_D^{4,\times}$ supports various constructs on concepts and roles, and it also admits data types, relevant in real word applications. In addition, it permits to express the Semantic Web Rule Language (SWRL), an extension of the Ontology Web Language (OWL). Decidability of the *Conjunctive Query Answering* (CQA) problem for $\mathcal{DL}_D^{4,\times}$ has been proved in [2] via a

reduction to the CQA problem for $4LQS^R$, whose decidability easily follows from that of $4LQS^R$ (see [1]). In [2], the authors provided a terminating KE-tableau based procedure that, given a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -query Q and a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -knowledge base \mathcal{KB} represented in set-theoretic terms, determines the answer set of Q with respect to \mathcal{KB} . Such an algorithm serves also as a decision procedure for the consistency problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -knowledge bases (KBs). We recall that KE-tableaux systems [10] construct tableaux whose distinct branches define mutually exclusive situations, thus preventing the proliferation of redundant branches, typical of semantic tableaux.

The results presented in [2] have been extended in [3] to the main ABox reasoning tasks for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$, such as instance checking and concept retrieval, by defining the Higher-Order Conjunctive Query Answering (HOCQA) problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$. Such problem, instantiable to the principal reasoning tasks for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -ABoxes, has been defined by introducing Higher Order (HO) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -conjunctive queries, admitting variables of three sorts: individual and data type variables, concept variables, and role variables. Decidability of the HOCQA problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ has been proved via a reduction to the HOCQA problem for the set-theoretic fragment $4LQS^R$.

In [4], an implementation of the KE-tableau procedure defined in [3] has been presented. Such prototype, written in C++, supports $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -KBs serialized in OWL/XML. It has been implemented only for TBox-reasoning services, namely, for verifying the consistency of given ontologies. Purely universal quantifiers are eliminated by the reasoner during a preprocessing phase, in which each quantified formula is instantiated in a systematic way with the individuals of the KB. The resulting instances are then suitably handled by applying to them the KE-elimination and bivalence rules. In the light of the benchmarking of the prototype, it turned out that the preprocessing phase of the universally quantified formulae is more and more expensive as the size of the KB grows.

In this paper, the KE-tableau-based procedure defined in [3] is modified, by eliminating the preprocessing phase for universally quantified formulae and replacing the standard KE-elimination rule with a novel elimination rule, called E^γ -rule, that incorporates the standard rule for treating universally quantified formulae (γ -rule). The resulting system¹ turns out to be more efficient than the KE-system in [4] and the First-Order (FO) KE-system in [10], as shown by suitable benchmarking tests executed on C++ implementations of the three systems. The main reason for such a speed-up relies on the fact that the E^γ -rule does not need to store the instances of the quantified formulae on the KE-tableau.

2 Preliminaries

2.1 The Set-Theoretic Fragment

It is convenient to recall the main set-theoretic notions behind the description logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ and its reasoning problems. For space reasons, we do not report the

¹ The source code is available at <https://github.com/dfsantamaria/DL4xD-Reasoner>.

syntax and semantics of the whole 4LQS^R : the interested reader can find it in [1], together with the decision procedure for its satisfiability problem. Thus, we just focus on the class of 4LQS^R -formulae actually involved in the set-theoretic representation of $\mathcal{DL}_D^{4,\times}$, namely, the propositional combinations of 4LQS^R -quantifier-free literals (atomic formulae and their negations) and of 4LQS^R purely universal formulae of the types displayed in Table 1. For the sake of conciseness, we refer to such class of 4LQS^R -formulae as $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$.

We recall that the fragment 4LQS^R admits four collections, Var_i , of variables of sort $i = 0, 1, 2, 3$, which are denoted by X^i, Y^i, Z^i, \dots (in particular, variables of sort 0 are also denoted by x, y, z, \dots). In addition to variables, also *pair terms* of the form $\langle x, y \rangle$, with $x, y \in \text{Var}_0$, are allowed. Since the types of formulae illustrated in Table 1 do not involve variables of sort 2, notions and definitions concerning $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formulae will refer to variables of sorts 0, 1, and 3 only.

Table 1. Types of literals and quantified formulae admitted in $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$.

Quantifier-free literals of level 0	Purely universal quantified formulae of level 1
$x = y, x \in X^1, \langle x, y \rangle \in X^3$ $\neg(x = y), \neg(x \in X^1), \neg(\langle x, y \rangle \in X^3)$	$(\forall z_1) \dots (\forall z_n)\varphi_0$, where $z_1, \dots, z_n \in \text{Var}_0$ and φ_0 is any propositional combination of quantifier-free atomic formulae of level 0

The variables z_1, \dots, z_n are said to occur *quantified* in $(\forall z_1) \dots (\forall z_n)\varphi_0$. A variable occurs *free* in a $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula φ if it does not occur quantified in any subformula of φ . For $i = 0, 1, 3$, we denote with $\text{Var}_i(\varphi)$ the collections of variables of sort i occurring free in φ .

For space reasons, the notions of $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -*substitution* and of $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -*interpretation* are not included here, but they can be found in [5].

2.2 The Logic $\mathcal{DL}_D^{4,\times}$

In what follows, we present the syntax and semantics of the description logic $\mathcal{DL}_D^{4,\times}$.

Let $\mathbf{R}_A, \mathbf{R}_D, \mathbf{C}, \mathbf{I}$ be denumerable pairwise disjoint sets of abstract role names, concrete role names, concept names, and individual names, respectively.

The definition of data types relies on the notion of data type map, given according to [11] as follows. A *data type map* is a quadruple $\mathbf{D} = (N_D, N_C, N_F, \cdot^{\mathbf{D}})$, where N_D is a finite set of data types, N_C is a function assigning a set of constants $N_C(d)$ to each data type $d \in N_D$, N_F is a function assigning a set of facets $N_F(d)$ to each $d \in N_D$, and $\cdot^{\mathbf{D}}$ is a function assigning a data type interpretation $d^{\mathbf{D}}$ to each $d \in N_D$, a facet interpretation $f^{\mathbf{D}} \subseteq d^{\mathbf{D}}$ to each facet $f \in N_F(d)$, and a data value $e_d^{\mathbf{D}} \in d^{\mathbf{D}}$ to every constant $e_d \in N_C(d)$. We shall assume that the interpretations of the data types in N_D are nonempty pairwise disjoint sets.

(a) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*data type*, (b) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*concept*, (c) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*abstract role*, and (d) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*concrete role terms* are constructed according to the following syntax rules:

- (a) $t_1, t_2 \longrightarrow dr \mid \neg t_1 \mid t_1 \sqcap t_2 \mid t_1 \sqcup t_2 \mid \{e_d\}$,
- (b) $C_1, C_2 \longrightarrow A \mid \top \mid \perp \mid \neg C_1 \mid C_1 \sqcup C_2 \mid C_1 \sqcap C_2 \mid \{a\} \mid \exists R. Self \mid \exists R. \{a\} \mid \exists P. \{e_d\}$,
- (c) $R_1, R_2 \longrightarrow S \mid U \mid R_1^{-1} \mid \neg R_1 \mid R_1 \sqcup R_2 \mid R_1 \sqcap R_2 \mid R_{C_1} \mid R_{|C_1} \mid R_{C_1} \mid C_2 \mid id(C) \mid C_1 \times C_2$,
- (d) $P_1, P_2 \longrightarrow T \mid \neg P_1 \mid P_1 \sqcup P_2 \mid P_1 \sqcap P_2 \mid P_{C_1} \mid P_{|t_1} \mid P_{C_1|t_1}$,

where dr is a data range for \mathbf{D} , t_1, t_2 are data type terms, e_d is a constant in $N_C(d)$, a is an individual name, A is a concept name, C_1, C_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -concept terms, S is an abstract role name, U is an abstract role name denoting the universal role, R, R_1, R_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -abstract role terms, T is a concrete role name, and P, P_1, P_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -concrete role terms. We remark that data type terms are introduced in order to represent derived data types.

A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -KB is a triple $\mathcal{K} = (\mathcal{R}, \mathcal{T}, \mathcal{A})$ such that \mathcal{R} is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*RBox*, \mathcal{T} is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*TBox*, and \mathcal{A} is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*ABox*. For space constraints, the definitions of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*RBox*, $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*TBox*, $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -*ABox* and the semantics of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ are omitted here, but the interested reader can find them in [5].

Expressiveness of the Description Logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$. Despite the fact that the description logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ is limited as far as the introduction of new individuals is concerned, it is more liberal than $\mathcal{SROIQ}(\mathbf{D})$ [7] in the construction of role inclusion axioms, since the roles involved are not restricted by any ordering relationship, the notion of simple role is not needed, and Boolean operations on roles and role constructs such as the product of concepts are admitted. Moreover, $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ supports more OWL constructs than the DLs underpinning the profiles OWL QL, OWL RL, and OWL EL [8], such as disjoint union of concepts and union of data ranges. Furthermore, basic and derived data types can be used inside inclusion axioms involving concrete roles. In addition, concerning the expressiveness of rules, the set-theoretic fragment $4\text{LQS}_{\mathcal{DL}_{\mathbf{D}}^{4,\times}}^R$ underpinning $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ allows one to express the disjunctive Datalog fragment admitting negation, equality and constraints, subject to no safety condition, and supporting for data types.

Reasoning with the Description Logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$. Next, we introduce the reasoning services available for the description logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$, i.e., the type of inferences that can be drawn from what is explicitly asserted in a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -KB. Specifically, we focus on two families of reasoning tasks, concerning respectively TBoxes and ABoxes. Among the main TBox reasoning problems, such as *satisfiability of a concept*, *subsumption of concepts*, *equivalence of concepts*, and *disjunction of concepts*, the problem of deciding the consistency of a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -KB is the most representative one, as it includes the majority of them.² In [2],

² A separate analysis, to be addressed in a future work, is required by the classification problem of a TBox, consisting in the computation of ancestor and descendant concepts of a given concept in a TBox.

we proved the decidability of the consistency problem of a $\mathcal{DL}_D^{4,\times}$ -KB and of a relevant ABox reasoning task, namely the *Conjunctive Query Answering* (CQA) problem for $\mathcal{DL}_D^{4,\times}$ consisting in computing the answer set of a $\mathcal{DL}_D^{4,\times}$ -conjunctive query with respect to a $\mathcal{DL}_D^{4,\times}$ -KB. In [3], we generalized the problem introducing the *Higher Order Conjunctive Query Answering* (HOCQA) problem for $\mathcal{DL}_D^{4,\times}$. Such problem is characterized by *Higher Order* (HO) $\mathcal{DL}_D^{4,\times}$ -conjunctive queries admitting variables of three sorts: individual and data type variables, concept variables, and role variables (the interested reader can find in [3] the definitions of HO-conjunctive queries, HO-substitutions, and HOCQA problem).

As illustrated in [3], the HOCQA problem can be instantiated to significant ABox reasoning problems such as (A) *role filler retrieval*, the problem of retrieving all the fillers x such that the pair (a, x) is an instance of a role R ; (B) *concept retrieval*, the problem of retrieving all concepts which an individual is an instance of; (C) *role instance retrieval*, the problem of retrieving all roles which a pair of individuals (a, b) is an instance of; and (D) *conjunctive query answering*, the problem of finding the answer set of a conjunctive query.

In [3], we solved the HOCQA problem just stated by reducing it to the analogous problem formulated in the context of the fragment $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ (and in turn of the decision procedure for 4LQS^R presented in [1]).

The HOCQA problem in the $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ context can be stated as follows. Let ϕ be a $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula and ψ a conjunction of $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -literals. The *HOCQA problem for ψ w.r.t. ϕ* consists in computing the HO *answer set of ψ w.r.t. ϕ* , namely, the collection Σ' of all the substitutions σ' such that $\mathcal{M} \models \phi \wedge \psi\sigma'$, for some $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -interpretation \mathcal{M} .

In view of the decidability of the satisfiability problem for 4LQS^R -formulae, the HOCQA problem for $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formulae is decidable as well. The reduction is carried out by means of a function θ that maps the $\mathcal{DL}_D^{4,\times}$ -KB \mathcal{KB} in a $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula $\phi_{\mathcal{KB}}$ in Conjunctive Normal Form (CNF) and the HO $\mathcal{DL}_D^{4,\times}$ -conjunctive query Q in the $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula ψ_Q (see [2] for details). Specifically,

$$\phi_{\mathcal{KB}} := \bigwedge_{H \in \mathcal{KB}} \theta(H) \wedge \bigwedge_{i=1}^{12} \xi_i, \quad \text{and} \quad \psi_Q := \theta(Q).$$

Let Σ be the HO-answer set of Q w.r.t. \mathcal{KB} (see [3] for the definition of HO-answer set) and Σ' the HO-answer set of ψ_Q w.r.t. $\phi_{\mathcal{KB}}$. Then Σ consists of all substitutions σ (involving exactly the variables occurring in Q) such that $\theta(\sigma) \in \Sigma'$. By Lemma 1 in [3], Σ' can be effectively computed and thus Σ can be effectively computed as well.

3 A KE-Tableau Based Algorithm for Reasoning in $\mathcal{DL}_D^{4,\times}$

In what follows, we briefly discuss the procedures *Consistency- $\mathcal{DL}_D^{4,\times}$* and *HOCQA $^\gamma$ - $\mathcal{DL}_D^{4,\times}$* . The procedure *Consistency- $\mathcal{DL}_D^{4,\times}$* checks the consistency

of its input $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula $\phi_{\mathcal{KB}}$, representing a $\mathcal{DL}_D^{4,\times}$ -KB. When $\phi_{\mathcal{KB}}$ is consistent, it builds a KE-tableau $\mathcal{T}_{\mathcal{KB}}$ whose distinct open and complete branches induce the models of $\phi_{\mathcal{KB}}$. Then the procedure $\text{HOCQA}^\gamma\text{-}\mathcal{DL}_D^{4,\times}$ computes the HO-answer set of a given $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -formula ψ_Q (representing a $\mathcal{DL}_D^{4,\times}$ -HO conjunctive query Q) w.r.t. $\phi_{\mathcal{KB}}$, by means of a forest of decision trees based on the branches of the KE-tableau $\mathcal{T}_{\mathcal{KB}}$ computed by the procedure $\text{Consistency-}\mathcal{DL}_D^{4,\times}$ with input $\phi_{\mathcal{KB}}$.

We now shortly introduce a variant of KE-tableau called KE^γ -tableau.³ Let $\Phi := \{C_1, \dots, C_p\}$, where each C_i is either a $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -literal of the types illustrated in Table 1 or a $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -purely universal quantified formula of the form $(\forall x_1) \dots (\forall x_m)(\beta_1 \vee \dots \vee \beta_n)$, with β_1, \dots, β_n $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$ -literals. \mathcal{T} is a KE^γ -tableau for Φ if there exists a finite sequence $\mathcal{T}_1, \dots, \mathcal{T}_t$ such that (i) \mathcal{T}_1 is the one-branch tree consisting of the sequence C_1, \dots, C_p , (ii) \mathcal{T}_{i+1} is obtained from \mathcal{T}_i , for each $i < t$, either by an application of one of the rules (E $^\gamma$ -rule or PB-rule) in Fig. 1 or by applying a substitution σ to a branch ϑ of \mathcal{T}_i (in particular, the substitution σ is applied to each formula X of ϑ ; the resulting branch will be denoted with $\vartheta\sigma$), and (iii) $\mathcal{T}_t = \mathcal{T}$. In the definition of the E $^\gamma$ -rule reported in Fig. 1 we have: (a) $\tau := \{x_1/x_{o_1} \dots x_m/x_{o_m}\}$ is a substitution such that x_1, \dots, x_m are the quantified variables in ψ and $x_{o_1}, \dots, x_{o_m} \in \text{Var}_0(\Phi)$; (b) $\mathcal{S}^{\beta_i\tau} := \{\bar{\beta}_1\tau, \dots, \bar{\beta}_n\tau\} \setminus \{\bar{\beta}_i\tau\}$ is a set containing the complements of all the disjuncts β_1, \dots, β_n to which the substitution τ is applied, with the exception of the disjunct β_i .

$$\frac{\psi \quad \mathcal{S}^{\bar{\beta}_i\tau}}{\beta_i\tau} \quad \text{E}^\gamma\text{-rule} \qquad \frac{}{A \mid A} \quad \text{PB-rule}$$

where A is a literal

where
 $\psi = (\forall x_1) \dots (\forall x_m)(\beta_1 \vee \dots \vee \beta_n)$,
 $\tau := \{x_1/x_{o_1} \dots x_m/x_{o_m}\}$,
and $\mathcal{S}^{\bar{\beta}_i\tau} := \{\bar{\beta}_1\tau, \dots, \bar{\beta}_n\tau\} \setminus \{\bar{\beta}_i\tau\}$,
for $i = 1, \dots, n$

Fig. 1. Expansion rules for the KE^γ -tableau.

Let \mathcal{T} be a KE^γ -tableau. A formula $\psi = (\forall x_1) \dots (\forall x_m)(\beta_1 \vee \dots \vee \beta_n)$ is *fulfilled* in a branch ϑ , if ϑ contains $\beta_i\tau$ for some $i = 1, \dots, n$ and for all τ having as domain the set $Q\text{Var}_0(\psi) = \{x_1, \dots, x_m\}$ of the quantified variables occurring in ψ and as range the set $\text{Var}_0(\vartheta)$ of the variables of sort 0 occurring free in ϑ . Notice that since the procedure $\text{Consistency-}\mathcal{DL}_D^{4,\times}$ does not introduce any new variable, $\text{Var}_0(\vartheta)$ coincides with $\text{Var}_0(\phi_{\mathcal{KB}})$, for every branch ϑ . A branch ϑ is *fulfilled* if every formula $\psi = (\forall x_1) \dots (\forall x_m)(\beta_1 \vee \dots \vee \beta_n)$ occurring in ϑ is fulfilled. A KE^γ -tableau is *fulfilled* when all its branches are fulfilled. A branch

³ KE-tableaux are a refutation system inspired to Smullyan’s semantic tableaux [13] (see [10] for details).

ϑ of \mathcal{T} is *closed* if either it contains both A and $\neg A$, for some formula A , or a literal of type $\neg(x = x)$; otherwise, it is *open*. A KE^γ -tableau is *closed* when all its branches are closed. A branch ϑ is *complete* if either it is closed or it is open, fulfilled, and it does not contain any literal of type $x = y$, with x and y distinct variables. A KE^γ -tableau is *complete* (resp., *fulfilled*) if all its branches are complete (resp., fulfilled or closed). The notions of branch and KE^γ -tableau *satisfiability* are the standard ones.

The procedure *Consistency- $\mathcal{DL}_{\mathbf{D}}^{4,\times}$* takes care of the literals of type $x = y$ occurring in the branches of $\mathcal{T}_{\mathcal{KB}}$ by constructing, for each open and fulfilled branch ϑ of $\mathcal{T}_{\mathcal{KB}}$, a substitution σ_ϑ such that $\vartheta\sigma_\vartheta$ does not contain literals of type $x = y$ with distinct x, y . Then, for every open and complete branch $\vartheta' := \vartheta\sigma_\vartheta$ of $\mathcal{T}_{\mathcal{KB}}$, the procedure *HOCQA $^\gamma$ - $\mathcal{DL}_{\mathbf{D}}^{4,\times}$* constructs a decision tree $\mathcal{D}_{\vartheta'}$ whose maximal branches induce substitutions σ' such that $\sigma_\vartheta\sigma'$ belongs to the HO-answer set of ψ_Q with respect to $\phi_{\mathcal{KB}}$ (the definition of $\mathcal{D}_{\vartheta'}$ can be found in [5]).⁴

We recall that the HOCQA problem can be solved in exponential-time when the $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -KBs do not contain qualified cardinality restrictions and role chain axioms, otherwise it can be solved in double-exponential time (see [5] for details).

Remarks on Different Versions of the Algorithm. The C++ implementation of the algorithm presented in this paper, called KE^γ -system, is more efficient than the prototype KE-system proposed in [4]. The main motivation behind this performance improvement relies on the introduction of the E^γ -rule (see Fig. 1) that acts on the $4\text{LQS}_{\mathcal{DL}_{\mathbf{D}}^{4,\times}}^R$ -purely universal quantified formulae in the KB by systematically instantiating them and applying the standard E-rule (elimination rule) on-the-fly. The E^γ -rule replaces the preliminary phase of systematic expansion of the $4\text{LQS}_{\mathcal{DL}_{\mathbf{D}}^{4,\times}}^R$ -purely universal quantified formulae in the KB and the subsequent application of the E-rule implemented by the KE-system presented in [3]. In addition, the KE^γ -system turns out to be also more efficient than the implementation FO KE-system of the FO KE-tableau in [10], that applies the standard γ - and E-rules. Incidentally, it turns out that the KE-system and the FO KE-system have similar performances. As shown in [5, Fig. 3], the KE^γ -system has a better performance, up to about 400%, than the other two, even if in some cases (lowest part of the plot) the performances of the three systems are comparable. Thus the KE^γ -system is always convenient, also because the collection of the expansions of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ -purely universal quantified formulae of level 1 (exponential in the size of the KB) is not stored in memory.

The benchmarking process is based on a huge amount of KBs of different sizes and kinds, constructed *ad hoc* just for the purpose of comparing the three mentioned systems, and on some real-world ontologies developed by the authors.

⁴ The procedures *Consistency- $\mathcal{DL}_{\mathbf{D}}^{4,\times}$* and *HOCQA $^\gamma$ - $\mathcal{DL}_{\mathbf{D}}^{4,\times}$* can be found in [5] together with their correctness, completeness, and termination proofs.

4 Conclusions and Future Work

We presented an improvement, called KE^γ -tableau, of the KE-tableau in [3] for the most widespread reasoning tasks for $\mathcal{DL}_D^{4,\times}$ -TBoxes and $\mathcal{DL}_D^{4,\times}$ -ABoxes. These reasoning problems are addressed by translating $\mathcal{DL}_D^{4,\times}$ -KBs and queries in terms of formulae of the set-theoretic language $4\text{LQS}_{\mathcal{DL}_D^{4,\times}}^R$. The procedure introduced in this paper generalizes the KE-elimination rule in such way as to incorporate the γ -rule, namely, the expansion rule for handling universally quantified formulae. The KE^γ -tableau procedure has remarkable aftermath, since its implementation is markedly more efficient in terms of space and execution time than the KE-system [4] and the implementation FO KE-system of the FO KE-tableau [10], as observed in our experimental tests.

In order to be able to reason with description logics admitting full existential and universal quantification, we plan to extend the set-theoretic fragment underpinning our reasoner with a restricted version of the operator of relational composition. Results and notions presented in [9] will be of inspiration for such a task. We also intend to improve our reasoner so as to deal with the reasoning problem of ontology classification. We shall compare the resulting reasoner with existing well-known reasoners such as Hermit [6] and Pellet [12], providing also some benchmarking. In addition, we plan to allow data type reasoning by either integrating existing solvers for the Satisfiability Modulo Theories (SMT) problem or by designing *ad hoc* new solvers. Finally, as each branch of a KE^γ -tableau can be independently computed by a single processing unit, we intend to implement a parallel version of the software by using the Nvidia CUDA framework.

References

1. Cantone, D., Nicolosi-Asmundo, M.: On the satisfiability problem for a 4-level quantified syllogistic and some applications to modal logic. *Fundamenta Informaticae* **124**(4), 427–448 (2013)
2. Cantone, D., Nicolosi-Asmundo, M., Santamaria, D.F.: Conjunctive query answering via a fragment of set theory. In: *Proceedings of ICTCS 2016, Lecce, 7–9 September*, vol. 1720, pp. 23–35. CEUR-WS (2016)
3. Cantone, D., Nicolosi-Asmundo, M., Santamaria, D.F.: A set-theoretic approach to abox reasoning services. In: Costantini, S., Franconi, E., Van Woensel, W., Kontchakov, R., Sadri, F., Roman, D. (eds.) *RuleML+RR 2017*. LNCS, vol. 10364, pp. 87–102. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61252-2_7
4. Cantone, D., Nicolosi-Asmundo, M., Santamaria, D.F.: A C++ reasoner for the description logic $\mathcal{DL}_D^{4,\times}$. In: *Proceedings of CILC 2017, Naples, Italy, 26–29 September 2017*, vol. 1949, pp. 276–280. CEUR WS (2017). ISSN 1613-0073
5. Cantone, D., Nicolosi-Asmundo, M., Santamaria, D.F.: An optimized KE-tableau-based system for reasoning in the description logic $\mathcal{DL}_D^{4,\times}$. *CoRR*, 1804.11222 (2018). Extended version
6. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z.: Hermit: an OWL 2 reasoner. *J. Autom. Reason.* **53**(3), 245–269 (2014)

7. Horrocks, I., Kutz, O., Sattler, U.: The even more irresistible SROIQ. In: Doherty, P., Mylopoulos, J., Welty, C.A. (eds.) Proceedings of 10th International Conference on Principle of Knowledge Representation and Reasoning, pp. 57–67. AAAI Press (2006)
8. Krötzsch, M.: OWL 2 profiles: an introduction to lightweight ontology languages. In: Eiter, T., Krennwallner, T. (eds.) Reasoning Web 2012. LNCS, vol. 7487, pp. 112–183. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33158-9_4
9. Cristiá, M., Rossi, G.: A decision procedure for restricted intensional sets. In: de Moura, L. (ed.) CADE 2017. LNCS, vol. 10395, pp. 185–201. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63046-5_12
10. Mondadori, M., D’Agostino, M.: The taming of the cut. Classical refutations with analytic cut. *J. Logic Comput.* **4**, 285–319 (1994)
11. Motik, B., Horrocks, I.: OWL datatypes: design and implementation. In: Sheth, A., et al. (eds.) ISWC 2008. LNCS, vol. 5318, pp. 307–322. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88564-1_20
12. Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y.: Pellet: a practical OWL-DL reasoner. *J. Web Semant.* **5**(2), 51–53 (2007)
13. Smullyan, R.M.: First-order Logic. Dover Books on Advanced Mathematics. Dover, New York (1995)