Scalable Geo-thematic Query Answering

Özgür Lütfü Özçep and Ralf Möller

Institute for Software Systems (STS)
Hamburg University of Technology
Hamburg, Germany
{oezguer.oezcep,moeller}@tu-harburg.de

Abstract. First order logic (FOL) rewritability is a desirable feature for query answering over geo-thematic ontologies because in most geo-processing scenarios one has to cope with large data volumes. Hence, there is a need for combined geo-thematic logics that provide a sufficiently expressive query language allowing for FOL rewritability. The DL-Lite family of description logics is tailored towards FOL rewritability of query answering for unions of conjunctive queries, hence it is a suitable candidate for the thematic component of a combined geo-thematic logic. We show that a weak coupling of DL-Lite with the expressive region connection calculus RCC8 allows for FOL rewritability under a spatial completeness condition for the ABox. Stronger couplings allowing for FOL rewritability are possible only for spatial calculi as weak as the low-resolution calculus RCC2. Already a strong combination of DL-Lite with the low-resolution calculus RCC3 does not allow for FOL rewritability.

Keywords: FOL rewritability, description logics, region connection calculus, qualitative spatial reasoning, GIS, combined logic

1 Introduction

Query answering over a database becomes far more difficult if the extensional knowledge in the database is extended by constraints in an ontology. The reason is that a database plus an ontology may have many different models, hence ontology based query answering has to compute the answers w.r.t. to all models and build their intersection (*certain answer* semantics). But in some cases—when using a lightweight logic like DL-Lite for the representation of the ontology and a restricted query language like unions of conjunctive queries—query answering w.r.t. an ontology can be reduced to model checking. This is formalized by the notion of *FOL (first order logic) rewritability*: a given query can be rewritten into a FOL query in which the intensional knowledge of the ontology is captured. Though the rewritten queries may become exponentially bigger than the original ones, there exist optimizations based on semantic indexes which encode entailed knowledge of the terminological part of the ontology [15]. So, FOL rewritability means a benefit.

DL-Lite per se [3] is not sufficient for use in scenarios of geographic information processing, as these demand, among others, the representation and deduction over spatial concepts. Though constraint-based spatial reasoning [14] offers a well developed and well proven theory for spatial domains, it does not fill in the need for a system that combines reasoning over a spatial and a non-spatial (thematic) domain. Though constraint databases [7] are good candidate frameworks for reasoning over a mixed domain of geo-thematic objects, the investigations on constraint databases so far did not incorporate terminological reasoning in the OBDA (ontology based data access) paradigm. But even in case of related work which equally considers spatial and thematic reasoning [5], [17], [8], it is not aimed at FOL rewritability. Hence, there is still a need for investigating combinations of logics that, on the one hand, are sufficiently expressive to match the representation requirements in geographical information processing and that, on the other hand, allow for computationally feasible (in particular FOL rewritable) satisfiability checking and query answering.

We would like to illustrate the use of the logics of this paper by a simple scenario in which an engineering bureau plans additional parks in a city [10]. Assume, the bureau has stored geographical data in some database (DB) and declares relevant concepts in the terminological part of his knowledge base, the TBox. The engineer gives necessary conditions for a concept *Park+Lake* which is a park containing a lake that touches it from within, i.e., using the terminology of the region connection calculus (RCC) [11], the lake is a tangential proper part of the park. Similarly, a necessary condition for the concept *Park4Playing* is given which is a park containing a playing ground (for children) that is a tangential proper part.

We assume that the data are mapped to a logical pendant of the DB called the ABox (assertional box). In particular the data should generate the fact that there is an object a which is both a park with a lake and with a playing area, that is *Park+Lake*(a) and *Park4Playing*(a) are contained in the ABox. But the location of a is not known. Think of a as an object whose architectural design is determined but the place where a is going to be localized is not determined yet.

Now, the engineering bureau asks for all parks with lakes and playing areas such that the playing area is not contained as island in the lake. These kinds of parks can be thought of as secure as the playing ground can be directly reached from the park (without a bridge). All objects that fall into the answer set of this query w.r.t. to the TBox and the data can have one of the configurations A to C illustrated in Figure 1 (and many more) but are not allowed to have the configuration D. The object a has to be in the answer set to the original query as the TBox together with the ABox and some deduction on the spatial configuration implies that a is an object which excludes the island configuration D. Remember that a is "abstract" in so far as its geographical location is not known. So in fact deduction is needed to see that a does not have configuration D. Later on we will formalize this example in the logic DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) and the query language GCQ^+ and show that the deduction needed for the inclusion of a into the answer set can be mimicked in a perfect rewriting algorithm.
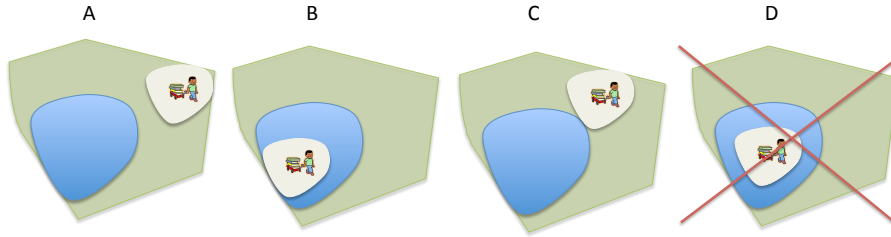
Fig. 1. (Dis-)Allowed spatial configurations for query in engineering bureau scenario

Continuing previous work [10], we investigate combinations of logics in the DL-Lite family with different members of RCC family [11], a well-known family of calculi for qualitative spatial reasoning. In previous work [10], we focussed on the FOL rewritability aspects for weak combinations of DL-Lite with RCC8; these combinations are weak in so far as they do not allow for the construction of arbitrary RCC8 constraint networks in the intensional part (TBox) of the ontology. In this paper we extend these investigations by enlarging the expressivity of DL-Lite to one that allows for concept conjunctions on the left-hand side of general inclusion axioms [1], and we give a proof including a rewriting algorithm for the main result stating that the weak combination of DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap}$ with RCC8 allows for FOL rewriting w.r.t. to the query language GCQ^+.

Moreover, in this paper, we consider strong combinations of DL-Lite with the weaker RCC fragments RCC3 and RCC2, and prove that DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap}$(RCC3) does not allow for FOL rewritability of satisfiability checking while the weaker DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap}$(RCC2) does [9].

The paper is structured as follows. Section 2 collects technical details on the region connection calculus and the DL-Lite family of description logics. Weak combinations of DL-Lite with the region connection calculus are described in Sect. 3. In Sect. 4, the last section before the conclusion, we consider strong combinations of DL-Lite with weaker fragments of the region connection calculus.

Full proofs can be found in an extended version of this paper available under the URL http://dl.dropbox.com/u/65078815/oezcepMoellerISWC2012Ext.pdf.

2 Logical Preliminaries

We recapitulate the main logical notation and concepts used in this paper; the region connection calculus and DL-Lite.

2.1 The Region Connection Calculus

We will consider different fragments of the region connection calculus [11] as potential candidates for the spatial logic to be combined with DL-Lite. Randell and colleagues' axiom system [11] is based on a primitive binary relation C intended to denote a connectedness relation which is specified to be reflexive and

symmetric. On the basis of C other binary relations between regions which are called *base relations* are explained. One set of base relations is the set $\mathcal{B}_{RCC8}$, which is the main component of the most expressive region connection calculus RCC8. The base relations of $\mathcal{B}_{RCC8}$ and their intended meanings are given as follows: $\mathcal{B}_{RCC8} = \{\mathsf{DC}$ (disconnected), EC (externally connected), EQ (equal), PO (partially overlapping), NTPP (non-tangential proper part), TPP (tangential proper part), NTPPi (inverse of NTPP), TPPi (inverse of TPP)$\}$. We skip the concrete definitions of the base relations by the connectedness relation C (see, e.g., [13, p. 45]), as we—in contrast to the axiom system of Randell and colleagues—consider the following axiom system schema Ax_{RCCi}, which directly specifies the properties of the base relations in $\mathcal{B}_{RCCi}$.

Definition 1 (Axiom system schema Ax_{RCCi}). *For all $i \in \{2,3,5,8\}$ the axiom set Ax_{RCCi} contains the following axioms:*

$$\{\forall x,y. \bigvee_{r \in \mathcal{B}_{RCCi}} r(x,y)\} \cup \qquad\qquad\qquad\qquad\text{(joint exhaustivity)}$$
$$\{\forall x,y. \bigwedge_{r_1,r_2 \in \mathcal{B}_{RCCi}, r_1 \neq r_2} r_1(x,y) \to \neg r_2(x,y)\} \cup \qquad\text{(pairwise disjointness)}$$
$$\{\forall x,y,z. r_1(x,y) \wedge r_2(y,z) \to r_3^1(x,z) \vee \cdots \vee r_3^k(x,z) \mid r_1; r_2 = \{r_3^1, \ldots, r_3^k\}\}$$
$$\text{(weak composition axioms)}$$

For $i \in \{3,5,8\}$ additionally the axiom $\forall x \mathsf{EQ}(x,x)$ (reflexivity of EQ) is contained. For $i = 2$ the axiom $\forall x \mathsf{O}(x,x)$ (reflexivity of O) is contained.

In particular, the axioms state the JEPD-property of the base relations (each pair of regions x, y is related over exactly one base relation) and describe the (weak) composition of two base relations (denoted by ;) according to the composition table for RCCi. With the composition of two base relations, in most cases, only indefinite knowledge of spatial configurations follows. The spatial configuration $r_1(x,z) \vee \cdots \vee r_n(x,z)$ for base relations r_j in $\mathcal{B}_{RCCi}$ is also written as $\{r_1, \ldots, r_n\}(x,z)$, and the set $\{r_1, \ldots, r_n\}$ is called a general RCCi relation. Let Rel_{RCCi} be the set of all 2^i general RCCi relations. An RCCi *(constraint) network* consists of assertions of the form $\{r_1, \ldots, r_n\}(x,y)$.

We mention here the composition table for the low resolution logics RCC2 and RCC3. Their base relations are given by the sets $\mathcal{B}_{RCC3} = \{\mathsf{DR}, \mathsf{EQ}, \mathsf{ONE}\}$ and $\mathcal{B}_{RCC2} = \{\mathsf{DR}, \mathsf{O}\}$, and their weak compositions are defined as shown in Fig. 2. The discreteness relation DR is the same as $\{\mathsf{DC}, \mathsf{EC}\}$, the overlapping-but-not-equal relation ONE is equal to $\{\mathsf{PO}, \mathsf{NTPP}, \mathsf{TPP}, \mathsf{NTPPi}, \mathsf{TPPi}\}$ and the overlapping relation O is given by $\{\mathsf{ONE}, \mathsf{EQ}\}$.

;	DR	O
DR	$\mathcal{B}_{RCC2}$	$\mathcal{B}_{RCC2}$
O	$\mathcal{B}_{RCC2}$	$\mathcal{B}_{RCC2}$

;	DR	ONE	EQ
DR	$\mathcal{B}_{RCC3}$	$\{\mathsf{DR}, \mathsf{ONE}\}$	DR
ONE	$\{\mathsf{DR}, \mathsf{ONE}\}$	$\mathcal{B}_{RCC3}$	ONE
EQ	DR	ONE	EQ

Fig. 2. Composition tables for RCC2 and RCC3

Note that in the definitions of the base relations (of RCC3 and RCC2) we followed the author of [16] and not [4]. But the composition tables for both definitions are identical. For the composition tables of RCC5 and RCC8 confer [14, p. 45]. As an example entry for RCC8, which is relevant for the engineering bureau scenario, we mention the table entry for the pair $(\mathsf{tpp}, \mathsf{tppi})$: $\mathsf{tpp}; \mathsf{tppi} = \{\mathsf{dc}, \mathsf{ec}, \mathsf{po}, \mathsf{tpp}, \mathsf{tppi}, \mathsf{eq}\}$ which is described in Ax_{RCC8} by $\forall x, y, z.\mathsf{tpp}(x, y) \wedge \mathsf{tppi}(y, z) \rightarrow \{\mathsf{dc}, \mathsf{ec}, \mathsf{po}, \mathsf{tpp}, \mathsf{tppi}, \mathsf{eq}\}(x, z)$. In case of the engineering bureau scenario from the introduction the constraint of this composition entry is demonstrated for x being the lake, y being the park and z being the playing ground.
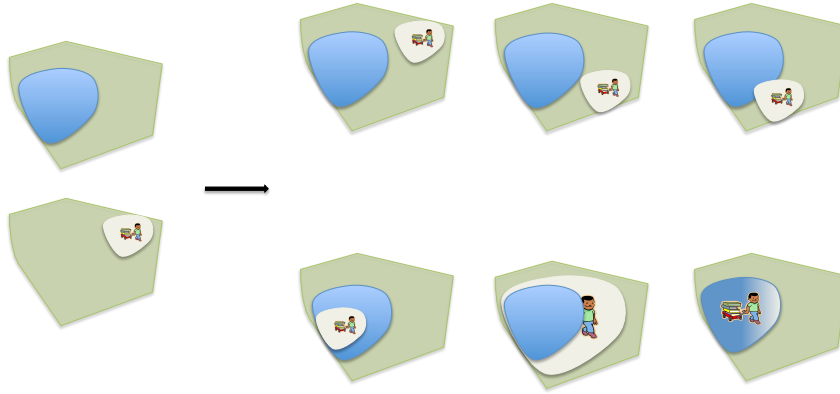

Fig. 3. Illustration for composition entry $\mathsf{tpp}; \mathsf{tppi}$.

2.2 DL-Lite

The family of DL-Lite description logics [3] is an appropriate candidate for the thematic component of the envisioned geo-thematic logic as it offers computationally feasible satisfiability checking and query answering over ontologies and data stored in a relational database. More concretely, satisfiability checking and query answering (for unions of conjunctive queries) are FOL rewritable. In this paper, we mainly deal with a member of the extended DL-Lite family DL-Lite$_{\mathcal{F}, \mathcal{R}}^{\sqcap}$; it allows for functional roles, role hierarchies, role inverses and conjunction of basic concepts on the left-hand side of GCIs (general concept inclusions). The syntax is given in Def. 2. The semantics of this logic is defined in the usual way—but imposing the unique name assumption (UNA).

Definition 2 (DL-Lite$_{\mathcal{F}, \mathcal{R}}^{\sqcap}$). *Let RN be the set of role symbols and $P \in RN$, CN be a set of concept symbols and $A \in CN$, Const be a set of individual constants and $a, b \in Const$.*

$$R \longrightarrow P \mid P^- \quad B \longrightarrow A \mid \exists R \quad C_l \longrightarrow B \mid C_l \sqcap B \quad C_r \longrightarrow B \mid \neg B$$
$$TBox^{*)}: \qquad C_l \sqsubseteq C_r, (\text{funct } R), R_1 \sqsubseteq R_2$$
$$ABox: \qquad A(a), R(a, b)$$

*) *Restriction: If R occurs in a functionality axiom, then R and its inverse do not occur on the right-hand side of a role inclusion axiom $R_1 \sqsubseteq R_2$.*

FOL rewritability also holds for the logic DL-Lite$^{\square}_{\mathcal{F},\mathcal{R}}$ which follows from the corresponding FOL rewritability results for the Datalog extension Datalog$^{\pm}$ [2]. We recapitulate the technical notions needed for defining FOL rewritability. An ontology $\mathcal{O}$ is a tuple $(Sig, \mathcal{T}, \mathcal{A})$, with a signature Sig (i.e., a set of concept symbols, role symbols and constants also denoted by $Sig(\mathcal{O})$), with a TBox $\mathcal{T}$, and with an ABox $\mathcal{A}$. An *FOL query* $Q = \psi(\boldsymbol{x})$ is an FOL formula $\psi(\boldsymbol{x})$ with free variables $\boldsymbol{x}$ called *distinguished variables*. If $\boldsymbol{x}$ is empty, the query is called *boolean*. Let $\boldsymbol{a}$ be a vector of constants from $Sig(\mathcal{O})$. The set of answers w.r.t. $\mathcal{I}$ is defined by $Q^{\mathcal{I}} = \{\boldsymbol{d} \in (\Delta^{\mathcal{I}})^n \mid \mathcal{I}_{[\boldsymbol{x} \mapsto \boldsymbol{d}]} \models \psi(\boldsymbol{x})\}$. (We use $Q^{\mathcal{I}}$ later on for a specific model $\mathcal{I}$, namely a Herbrand model.) The set of certain answers w.r.t. to $\mathcal{O}$ is defined by $cert(Q, \mathcal{T} \cup \mathcal{A}) = \{\boldsymbol{a} \mid \mathcal{T} \cup \mathcal{A} \models \psi[\boldsymbol{x}/\boldsymbol{a}]\}$. A *conjunctive query (CQ)* is a FOL query in which $\psi(\boldsymbol{x})$ is an existentially quantified conjunction of atomic formulas $at(\cdot)$, $\psi(\boldsymbol{x}) = \exists \boldsymbol{y} \bigwedge_i at_i(\boldsymbol{x}, \boldsymbol{y})$. A *union of conjunctive queries (UCQ)* is a disjunction of CQs, i.e., a formula of the form $\exists \boldsymbol{y_1} \bigwedge_{i_1} at_{i_1}(\boldsymbol{x}, \boldsymbol{y_1}) \vee \cdots \vee \exists \boldsymbol{y_n} \bigwedge_{i_n} at_{i_n}(\boldsymbol{x}, \boldsymbol{y_n})$. We conceive a UCQ as a set of CQs. The existential quantifiers in UCQs are interpreted in the same way as for FOL formulas (natural domain semantics) and not w.r.t. a given set of constants mentioned in the signature (active domain semantics).

Let $DB(\mathcal{A})$ be the minimal Herbrand model of $\mathcal{A}$. *Checking the satisfiability of ontologies is FOL rewritable* iff for all TBoxes $\mathcal{T}$ there is a boolean FOL query $Q_{\mathcal{T}}$ s.t. for all ABoxes $\mathcal{A}$: the ontology $\mathcal{T} \cup \mathcal{A}$ is satisfiable iff $DB(\mathcal{A}) \not\models Q_{\mathcal{T}}$. *Answering queries from a subclass $\mathcal{C}$ of FOL queries w.r.t. to ontologies is FOL rewritable* iff for all TBoxes $\mathcal{T}$ and queries $Q = \psi(\boldsymbol{x})$ in $\mathcal{C}$ there is a FOL query $Q_{\mathcal{T}}$ such that for all ABoxes $\mathcal{A}$ it is the case that $cert(Q, \mathcal{T} \cup \mathcal{A}) = Q_{\mathcal{T}}^{DB(\mathcal{A})}$. For DL-Lite, FOL-rewritability can be proved w.r.t. to satisfiability as well as w.r.t. answering UCQs [3, Thm 4.14, Thm 5.15].

The rewritability results are proved with the so called chase construction known from database theory. The idea of the chase construction is to repair the ABox with respect to the constraints formulated in the TBox. If, e.g., the TBox contains the axiom $A_1 \sqsubseteq A_2$ and the ABox contains $A_1(a)$ but not $A_2(a)$, then it is enriched by the atom $A_2(a)$. This procedure is applied stepwise to yield a sequence of ABoxes S_i starting with the original ABox as S_0. The resulting set of ABox axioms $\bigcup S_i$ may be infinite but induces a canonical model $can(\mathcal{O})$ for the ABox and the TBox axioms being used in the chasing process (see below). We will summarize the chase construction for DL-Lite.

Let $\mathcal{T}$ be a DL-Lite TBox, let $\mathcal{T}_p$ be the subset of positive inclusion (PI) axioms in $\mathcal{T}$ (no negation symbol allowed) and let $\mathcal{A}$ be an ABox and $\mathcal{O} = \mathcal{T} \cup \mathcal{A}$. Chasing will be carried out with respect to PIs only. Let $S_0 = \mathcal{A}$. Let S_i be the set of ABox axioms constructed so far and α be a PI axiom in $\mathcal{T}_p$. Let α be of the form $A_1 \sqsubseteq A_2$ and let $\beta \in S_i$ (resp. $\beta \subseteq S_i$) be an ABox axiom (resp. set

of ABox axioms). The PI axiom α is called applicable to β if β is of the form $A_1(a)$ and $A_2(a)$ is not in S_i. The applicability of other PI axioms of the form $B \sqsubseteq C$ is defined similarly [3, Def. 4.1, p. 287]. If the left-hand side of the PI is a conjunction of base concepts, e.g., if the PI is of the form $A_1 \sqcap \cdots \sqcap A_n \sqsubseteq A_0$, and if β is $\{A_1(a), \ldots, A_n(a)\}$ and $A_0(a)$ is not in S_i, then PI is applicable to β.

As there may be many possible applications of PI axioms to atoms and sets of atoms, one has to impose an order on the TBox axioms and the (finite) subsets of the ABox. So we assume that all strings over the signature $Sig(\mathcal{O})$ of the ontology and some countably infinite set of new constants C_{ch} are well ordered. Such a well ordering exists and has the order type of the natural numbers $\mathbb{N}$. This ordering is different from the one of [3]; but it can also be used also for infinite ABoxes and it can handle concept conjunction. If there is a PI axiom α applicable to an atom β in S_i, one takes the minimal pair (α, β) with respect to the ordering and produces the next level $S_{i+1} = S_i \cup \{\beta_{new}\}$; here β_{new} is the atom that results from applying the chase rule for (α, β) as listed in Def. 3. The primed constants are the chasing constants from C_{ch}.

Definition 3 (Chasing rules for DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap}$).

If $\alpha = A_1 \sqsubseteq A_2$ and $\beta = A_1(a)$ then $\beta_{new} = A_2(a)$
If $\alpha = A_1 \sqsubseteq \exists R$ and $\beta = A_1(a)$ then $\beta_{new} = R(a, a')$
If $\alpha = \exists R \sqsubseteq A$ and $\beta = R(a, b)$ then $\beta_{new} = A(a)$
If $\alpha = \exists R_1 \sqsubseteq \exists R_2$ and $\beta = R_1(a, b)$ then $\beta_{new} = R_2(a, a')$
If $\alpha = R_1 \sqsubseteq R_2$ and $\beta = R_1(a, b)$ then $\beta_{new} = R_2(a, b)$
If $\alpha = A_1 \sqcap \cdots \sqcap A_n \sqsubseteq A_0$ and $\beta = \{A_1(a), \ldots, A_n(a)\}$ then $\beta_{new} = A_2(a)$
(and similarly for other PIs of the form $B_1 \sqcap \cdots \sqcap B_n \sqsubseteq C$)

The chase is defined by $chase(\mathcal{O}) = chase(\mathcal{T}_p \cup \mathcal{A}) = \bigcup_{i \in \mathbb{N}} S_i$. The canonical model $can(\mathcal{O})$ is the minimal Herbrand model of $chase(\mathcal{O})$. The canonical model $can(\mathcal{O})$ is a universal model of $\mathcal{T}_p \cup \mathcal{A}$ with respect to homomorphisms. In particular this implies that answering a UCQ Q w.r.t. to $\mathcal{T}_p \cup \mathcal{A}$ can be reduced to answering $Q^{can(\mathcal{O})}$ w.r.t. to $DB(\mathcal{A})$. More concretely, (some finite closure $cln(\mathcal{T})$ of) the negative inclusions axioms and the functionality axioms are only relevant for checking the satisfiability of the ontology which can be tested by a simple FOL query. We leave out the details (see the extended version of this paper).

3   Weak Combinations of DL-Lite with RCC

In this section, we extend the results concerning a weak coupling of DL-Lite with the most expressive region connection calculus fragment RCC8, which we introduced in [10], and explain its use(fulness) with a formalization of the example scenario from the introduction. This will give us the opportunity to introduce further concepts that are necessary to understand the discussions on stronger couplings of DL-Lite with the weaker region connection calculi RCC2 and RCC3.

The combination paradigm follows that of Lutz and Miličič [8] who combine $\mathcal{ALC}$ with the RCC8 and, more generally, with ω-admissible concrete domains

[8, Def. 5, p. 7]. The combined logic $\mathcal{ALC}(RCC8)$ of [8] is well behaved in so far as testing concept subsumption is decidable. As we aim at FOL rewritability we have to be even more careful in choosing the right combination method.

We use an axiom set T_ω with corresponding properties of an ω-admissible domain for coupling with DL-Lite because axioms are more appropriate for rewriting investigations. The axiom sets Ax_{RCCi} will instantiate T_ω.

We recapitulate the syntax and the semantics of the constructors of [8] that are used for the coupling of the thematic and the spatial domain. A path U (of length at most 2) is defined as l for a fixed attribute l ("has location") or as $R \circ l$, the composition of the role symbol R with l. We abbreviate $R \circ l$ with $\tilde{R}$ in this paper. The usual notion of an interpretation $\mathcal{I}$ in our combined logic is slightly modified by using two separate domains $\Delta^\mathcal{I}$ and $(\Delta^*)^\mathcal{I}$. All symbols of the theory T_ω are interpreted relative to $(\Delta^*)^\mathcal{I}$. Let r be an RCC-relation of some RCC-fragment. That is, let be given a set of base relations $\mathcal{B}_{RCCi}$ and $r = \{r_1, \ldots r_n\} \equiv r_1 \vee \cdots \vee r_n$ for $r_i \in \mathcal{B}_{RCCi}$. Then $l^\mathcal{I} \subseteq \Delta^\mathcal{I} \times (\Delta^*)^\mathcal{I}$; $r^\mathcal{I} = r_1^\mathcal{I} \cup \cdots \cup r_n^\mathcal{I}$; $(R \circ l)^\mathcal{I} = \{(d, e^*) \in \Delta^\mathcal{I} \times (\Delta^*)^\mathcal{I} \mid$ there is an e s.t. $(d, e) \in R^\mathcal{I}$ and $(e, e^*) \in l^\mathcal{I}\}$; $(\exists U_1, U_2.r)^\mathcal{I} = \{d \in \Delta^\mathcal{I} \mid$ there exist e_1^*, e_2^* s.t. $(d, e_1^*) \in U_1^\mathcal{I}, (d, e_2^*) \in U_2^\mathcal{I}$ and $(e_1^*, e_2^*) \in r^\mathcal{I}\}$; $(\forall U_1, U_2.r)^\mathcal{I} = \{d \in \Delta^\mathcal{I} \mid$ for all e_1^*, e_2^* s.t. $(d, e_1^*) \in U_1^\mathcal{I}, (d, e_2^*) \in U_2^\mathcal{I}$ it holds that $(e_1^*, e_2^*) \in r^\mathcal{I}\}$.

Now we can define the following combined geo-thematic logic (where a^*, b^* stand for constants intended to be interpreted by regions):

Definition 4 (DL-Lite$_{\mathcal{F}, \mathcal{R}}^\square$(RCC8)). *Let $r \in Rel_{RCC8}$ and $T_\omega = Ax_{RCC8}$.*

$$R \longrightarrow P \mid P^- \qquad U \longrightarrow R \mid \tilde{R} \qquad B \longrightarrow A \mid \exists R \mid \exists l$$
$$C_l \longrightarrow B \mid C_l \sqcap B \qquad C_r \longrightarrow B \mid \neg B \mid \exists U_1, U_2.r$$
$$TBox^{*)}: \qquad C_l \sqsubseteq C_r, (\text{funct } l), (\text{funct } R), R_1 \sqsubseteq R_2$$
$$ABox: \qquad A(a), R(a, b), l(a, a^*), r(a^*, b^*)$$

*) Restriction: If (funct R) $\in \mathcal{T}$, *then R and R^- do not occur on the right-hand side of a role inclusion axiom or in a concept of the form $\exists U_1, U_2.r$.*

As satisfiability checking of RCC8 constraint networks is NP-complete, there is only a chance to reach FOL rewritability if we assume within the ABox a constraint network which is consistent and complete, i.e., it has a exactly one solution and it is a clique with base relations as labels; in this case the ABox is called *spatially complete*. For cadastral maps or maps containing areas of administration one can assume pretty safely (almost) spatial completeness. The coupling with RCC8 is so weak that FOL rewritability of satisfiability follows.

Proposition 1. *Checking the satisfiability of $DL-Lite_{\mathcal{F}, \mathcal{R}}^\square$ (RCC8) ontologies that have a spatially complete ABox is FOL rewritable.*

Testing whether FOL rewritability holds for satisfiability tests is necessary for tests whether FOL rewritability is provable for query answering w.r.t. a sufficiently expressive query language. The query language which we consider is derived from grounded conjunctive queries and is denoted by GCQ^+. This query language is explicitly constructed for use with DL-Lite$_{\mathcal{F}, \mathcal{R}}^\square$(RCC8) and so provides only means for qualitative spatial queries. But it could be extended to allow also for quantitative spatial queries.

Definition 5. *A GCQ^+ atom w.r.t. DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) is a formula of one of the following forms:*

- *$C(x)$, where C is a DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) concept without the negation symbol and x is a variable or a constant.*
- *$(\exists R_1 \ldots R_n.C)(x)$ for role symbols or their inverses R_i, a DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) concept C without the negation symbol, and a variable or a constant x*
- *$R(x,y)$ for a role symbol R or an inverse thereof*
- *$l(x,y^*)$, where x is a variable or constant and y^* is a variable or constant intended to denote elements of Ax_{RCC8}*
- *$r(x^*,y^*)$, where $r \in Rel_{RCC8}$ and x^*, y^* are variables or constants intended to denote elements of Ax_{RCC8}*

A GCQ^+ query w.r.t. DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) is a query $\tilde{\exists}\boldsymbol{yz}^* \bigwedge C_i(\boldsymbol{x},\boldsymbol{w}^*,\boldsymbol{y},\boldsymbol{z}^*)$ where all $C_i(\boldsymbol{x},\boldsymbol{w}^*,\boldsymbol{y},\boldsymbol{z}^*)$ are GCQ^+ atoms and $\tilde{\exists}\boldsymbol{yz}^* = \tilde{\exists}y_1 \ldots \tilde{\exists}y_n\tilde{\exists}z_1^* \ldots \tilde{\exists}z_m^*$ is a sequence of $\exists$-quantifiers interpreted w.r.t. the active domain semantics.

Our perfect rewriting algorithm is an an adaptation of the algorithm PerfectRef [3, Fig. 13] for reformulating UCQs w.r.t. DL-Lite ontologies to our setting in which GCQ^+-queries are asked to DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$(RCC8) ontologies. We give a description of our adapted algorithm in the following.

Given a query GCQ^+ Q, we transform it to a special form; $\tau_1(Q)$ is the result of the transformation to a UCQ and $\tau_2(Q)$ is the result of transforming Q in a hybrid UCQ whose conjuncts are either classical predicate logical atoms or GCQ^+-atoms which are not further transformed. We use the notation "$g = F$" for "g is of the form F".

The original algorithm PerfectRef operates on the PI axioms of a DL-Lite ontology by using them as rewriting aids for the atomic formulas in the UCQ. Lines 5–12 and 28–34 of our adapted algorithm (Algorithm 1) make up the original PerfectRef. Roughly, the PerfectRef algorithm acts in the inverse direction with respect to the chasing process. For example, if the TBox contains the PI axiom $A_1 \sqcap A_2 \sqsubseteq A_3$, and the UCQ contains the atom $A_3(x)$ in a CQ, then the new rewritten UCQ query contains a CQ in which $A_3(x)$ is substituted by $A_1(x) \wedge A_2(x)$. The applicability of a PI axiom to an atom is restricted in those cases where the variables of an atom are either distinguished variables or also appear in another atom of the CQ at hand. To handle these cases, PerfectRef—as well as also our adapted version—uses anonymous variables $\_$ to denote all non-distinguished variables in an atom that do not occur in other atoms of the same CQ. The function anon (line 31 in Algorithm 1) implements the anonymization. The application conditions for PI axioms α and atoms are as follows: α is applicable to $A(x)$ if A occurs on the right-hand side; and α is applicable to $R(x_1, x_2)$, if $x_2 = \_$ and the right-hand side of α is $\exists R$; or $x_1 = \_$ and the right-hand side of α is $\exists R^-$; or α is a role inclusion assertion and its right-hand side is either R or R^-. The outcome $gr(g,\alpha)$ of applying an applicable PI α to an atom g corresponds to the outcome of resolving α with g. For example, if α is $A \sqsubseteq \exists R$ and g is $R(x, \_)$, the result of the application is $gr(g,\alpha) = A(x)$. We leave out the details [3, Fig.12, p. 307]. In PerfectRef, atoms in a CQ are rewritten with

input : a hybrid query $\tau_1(Q) \cup \tau_2(Q)$, DL-Lite(RCC8) TBox $\mathcal{T}$
output: a UCQ pr

1 $pr := \tau_1(Q) \cup \tau_2(Q)$;
2 repeat
3 $pr' := pr$;
4   **foreach** *query* $q' \in pr'$ **do**
5   **foreach** *atom* g *in* q' **do**
6   **if** g *is a FOL-atom* **then**
7   **foreach** *PI* α *in* $\mathcal{T}$ **do**
8   **if** α *is applicable to* g **then**
9 $pr := pr \cup \{q'[g/gr(g,\alpha)]\}$;
10   **end**
11   **end**
12   **else**
13   **if** $g = \exists \tilde{R}_1, \tilde{R}_2.r_3(x)$ *and* $r_1; r_2 \subseteq r_3$ **then**
14 $X := q'[g/(\exists \tilde{R}_1, l.r_1(x) \wedge \exists l, \tilde{R}_2.r_2(x))]$;
15 $pr := pr \cup \{X\} \cup \{\tau_2\big(X, \{\exists \tilde{R}_1, l.r_1(x), \exists l, \tilde{R}_2.r_2(x)\}\big)\}$
16   **end**
17   **if** $g = \exists U_1, U_2.r_1(x)$ *and* $B \sqsubseteq \exists U_1, U_2.r_2(x) \in \mathcal{T}$ *for* $r_2 \subseteq r_1$
                      then
18 $pr := pr \cup \{q'[g/B(x)]\}$;
19   **end**
20   **if** $g = \exists U_1, U_2.r_1(x)$ *and* $B \sqsubseteq \exists U_1, U_2.r_2(x) \in \mathcal{T}$ *for* $r_2^{-1} \subseteq r_1$
                      then
21 $pr := pr \cup \{q'[g/B(x)]\}$;
22   **end**
23   **if** $g = \exists \tilde{R}_1, U_1.r(x)$ *(resp.* $\exists U_1, \tilde{R}_1.r(x)$*) and (*$R_2 \sqsubseteq R_1 \in \mathcal{T}$ *or*
                      $R_2^{-1} \sqsubseteq R_1^{-1} \in \mathcal{T}$*)* **then**
24 $X := q'[g/(g[R_1/R_2])]$;
25 $pr := pr \cup \{X\} \cup \{\tau_2\big(X, \{g[R_1/R_2]\}\big)\}$;
26   **end**
27   **end**
28   **end**
29   **foreach** *pair of FOL-atoms* g_1, g_2 *in* q' **do**
30   **if** g_1 *and* g_2 *unify* **then**
31 $pr := pr \cup \{anon(reduce(q', g_1, g_2))\}$;
32   **end**
33   **end**
34   **end**
35 until $pr' = pr$;
36 return $drop(pr)$

Algorithm 1: Adapted PerfectRef

the PI axioms (lines 6–11) and if possible merged by the function reduce (line 31) which unifies the atoms with the most general unifier (lines 28–34).

The modification of PerfectRef concerns the handling of GCQ^+-atoms of the form $\exists U_1, U_2.r(x)$. These atoms may have additional implications that are accounted for with four cases (lines 12–26 of the algorithm). At the end of the adapted algorithm PerfectRef (Algorithm 1, line 35) these atoms are deleted by calling the function *drop*. The algorithm returns a classical UCQ, which can be evaluated as a SQL query on the database $DB(\mathcal{A})$.

Let us demonstrate the rewriting algorithm with a formalization of the simple scenario from the beginning. The TBox of the engineering bureau contains the following axioms of DL-Lite$_{\mathcal{F},\mathcal{R}}^{\Box}$(RCC8): $Park+Lake \sqsubseteq Park$; $Park4Playing \sqsubseteq Park$; $Park+Lake \sqsubseteq \exists hasLake \circ l, l.\mathsf{tpp}$; $Park4Playing \sqsubseteq \exists hasPlAr \circ l, l.\mathsf{tpp}$. The ABox $\mathcal{A}$ contains $\{Park+Lake(a), Park4Playing(a)\} \subseteq \mathcal{A}$. The query of the engineer, which asks for all parks with lakes and playing areas such that the playing area is not a tangential proper part of the lake, can be formalized by the following GCQ^+: $Q = Park(x) \wedge \exists hasLake \circ l, hasPlAr \circ l.(\mathcal{B}_{RCC8} \setminus \{\mathsf{ntpp}\})(x)$. Using the composition $\mathsf{tpp};\mathsf{tppi} = \{\mathsf{dc}, \mathsf{ec}, \mathsf{po}, \mathsf{tpp}, \mathsf{tppi}, \mathsf{eq}\} \subseteq \mathcal{B}_{RCC8} \setminus \{\mathsf{ntpp}\}$, the reformulation algorithm introduced above (lines 13–15) produces a UCQ that contains the following CQ: $Q' = (\exists hasLake \circ l, l.\mathsf{tpp})(x) \wedge (\exists l, hasPlAr \circ l.\mathsf{tppi})(x)$. Rewriting $\exists l, hasPlAr \circ l.\mathsf{tppi}$ to $\exists hasPlAr \circ l, l.\mathsf{tpp}$ (lines 20–21) in combination with the rewriting rule for $A_1 \sqsubseteq A_2$ (Def. 3) we get another CQ $Q'' = Park+Lake(x) \wedge Park4Playing(x)$. Now, Q'' captures (as desired) the object a.

That the rewriting given in Algorithm 1 is indeed correct and complete follows from Theorem 1.

Theorem 1. *Answering GCQ^+-queries w.r.t. $DL-Lite_{\mathcal{F},\mathcal{R}}^{\Box}$(RCC8) ontologies that have a spatially complete ABox is FOL-rewritable.*

We give a proof sketch. The proof follows the proof of Theorem 5.15 for pure DL-Lite ontologies [3]. We adapt the chase construction to account for the RCC8 relations $r \in Rel_{RCC8}$. The main observation is that the disjunctions in r can be nearly handled as if they were predicate symbols.

Because of Prop. 1 we may assume that $\mathcal{O}$ is satisfiable. Let pr be the UCQ resulting from applying the algorithm to Q and $\mathcal{O}$. We have to show that $cert(Q, \mathcal{O}) = (pr)^{DB(\mathcal{A})}$. These can be done in two main steps of which the first will be sketched here as it contains the main chase-like construction $chase^*(\mathcal{O})$. After the construction, one has to describe what it means to answer Q with respect to $chase^*(\mathcal{O})$ is, resulting in the set $ans(Q, chase^*(\mathcal{O}))$, and then show that $ans(Q, chase^*(\mathcal{O})) = cert(Q, \mathcal{O})$. In the second step, which we leave out here (see the extended version of this paper) one has to show that $ans(Q, chase^*(\mathcal{O})) = (pr)^{DB(\mathcal{A})}$.

For the construction of $chase^*(\mathcal{O})$ one uses the chase rules of Def. 3 and the special rule (R).

Chasing Rule (R)
If $B(x) \in S_i$ and there are no y, y^*, x^* s.t. $\{R_1(x, y), l(y, y^*), l(x, x^*), r_1(y^*, x^*)\}$

is contained in S_i, then let $S_{i+1} = S_i \cup \{R_1(x,y), l(y,y^*), l(x,x^*), r_1(y^*,x^*)\}$. The constants y, y^* are completely new constants not appearing in S_i. The constant x^* is the old x^* if already in S_i, otherwise it is also a completely new constant symbol.

Every time (R) is applied to yield a new ABox S_i, the resulting constraint network in S_i is saturated by calculating the minimal labels between the new added region constants and the other region constants. The application of (R) does not constrain the RCC8-relations between the old regions and even stronger: Let (R) be applied to a TBox axiom of the form $A \sqsubseteq \exists \tilde{R}, l.r$ and $A(a) \in S_i$ resulting in the addition of $R(a,b)$, $l(b,b^*)$ and $r(b^*,a^*)$. Then it is enough to consider all $c^* \in S_i$ and all relations r_{c^*,a^*} such that $r_{c^*,a^*}(c^*,a^*) \in S_i$. The composition table gives the outcome $r_{c^*,a^*}; r = r'_{c^*,b^*}$ and one adds $r'_{c^*,b^*}(c^*,b^*)$ to S_i. After this step, which we call triangulation step, one closes the assertions up with respect to the subset relation between RCC8-relations and with respect to symmetry. I.e., if $r_1(x^*,y^*)$ is added to S_i, then one also adds $r_2(x^*,y^*)$ for all r_2 such that $r_1 \subseteq r_2$ and $r_2^{-1}(y^*,x^*)$. For different c_1^*, c_2^*, assertions of the form $r_{c_1^*,b^*}(c_1^*,b^*)$ and $r_{c_2^*,b^*}(c_2^*,b^*)$ do not constrain each other (because of the patch work property). The saturation leads to a finite set S_{i+k} (for some $k \in \mathbb{N}$) that is a superset of S_i. Let $chase^*(\mathcal{O}) = \bigcup S_i$. The set $chase^*(\mathcal{O})$ does not induce a single canonical model. But it is universal in the following sense: For every model $\mathcal{I}$ of $\mathcal{O}$ define a model $\mathcal{I}_c$ out of $chase^*(\mathcal{O})$ by taking a (consistent) configuration of the contained RCC8-network and taking the minimal model of this configuration and the thematic part of $chase^*(\mathcal{O})$. Then $\mathcal{I}_c$ maps homomorphically to $\mathcal{I}$. Now one can define that answers of GCQ^+-queries with respect to $chase^*(\mathcal{O})$ are given by homomorphic embeddings and show that these answers are exactly the certain answers w.r.t. the ontology $\mathcal{O}$.

4 Strong Combinations of DL-Lite with RCC

Another way of reaching FOL rewritability for combinations of DL-Lite with RCC is weakening the expressivity of the spatial component. Hence, one may ask whether a combination with the calculus RCC3 or RCC2 [17], both fragments with weak expressibility, allows for weak FOL rewritability w.r.t. satisfiability checks (and query answering). Their potential use as logics for approximating [6] ontologies in more expressible combined logics like $\mathcal{ALC}(\text{RCC8})$ makes the investigation valuable. The logics DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}(\text{RCC2})$ and DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}(\text{RCC3})$ are defined as follows ('+' indicates the strong combination):

Definition 6 (DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) and DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC3)). *Let* $T_\omega = Ax_{RCC2}$ *resp.* $T_\omega = Ax_{RCC3}$ *and* $r \in \mathcal{B}_{RCC2}$ *resp.* $r \in \mathcal{B}_{RCC3}$

$$R \longrightarrow P \mid P^- \qquad U \longrightarrow l \mid \tilde{R} \qquad B \longrightarrow A \mid \exists R$$
$$C_l \longrightarrow B \mid C_l \sqcap B \qquad C_r \longrightarrow B \mid \neg B \mid \exists U_1, U_2.r$$
$$TBox^{*)}: \qquad C_l \sqsubseteq C_r, (\text{funct } l, R), R_1 \sqsubseteq R_2$$
$$ABox: \qquad A(a), R(a,b), l(a,a^*), r(a^*,b^*)$$

*) Restriction: If $(\text{funct } R) \in \mathcal{T}$, then R and R^- do not occur on the right-hand side of a role inclusion axiom.

For RCC3 the strong combination with DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square}$ leads to non-FOL rewritability. The reason lies in the fact that testing the satisfiability of RCC3 is not in the complexity class AC^0 as shown by the following lemma.

Lemma 1. *Checking satisfiability of RCC3 networks is* LOGSPACE *hard.*

Proof. As is known, the reachability problem in symmetric (undirected) graphs is logspace complete [12]—where graph reachability asks whether for nodes s, t in G there is a path between s and t. By reducing this problem to the satisfiability test for RCC3 networks we will have shown that the latter problem is LOGSPACE hard itself. So let be given a (symmetric) graph $G = (V, E)$ and nodes $s, t \in V$. We define the network N in the following way (see Figure 4): Let $V = \{v_1, \ldots, v_n\}$ be an enumeration of the nodes of G; w.l.o.g. let $s = v_1$ and $t = v_n$ and let $\mathcal{B} = \mathcal{B}_{RCC3}$. Nodes of N are given by $V \cup \{a\}$ where $a \notin V$. Labelled edges of N are given by: $s\{\mathsf{DR}\}a$; $t\{\mathsf{ONE}\}a$; $v_i\{\mathcal{B}\}a$ for all $i \neq 1, n$; $v_i\{\mathsf{EQ}\}v_j$ if $E(v_i, v_j)$; $v_i\{\mathcal{B}\}v_j$ if $\neg E(v_i, v_j)$. Now we show that the network N is satisfiable iff s and t are
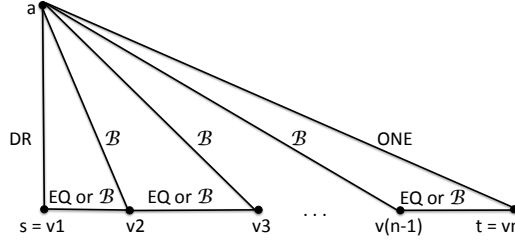
Fig. 4. Network N used in proof of Lemma 1

connected in G. Assume that s and t are connected; then there is an EQ-path in N between them, hence $s\{\mathsf{EQ}\}t$ follows. But this contradicts $s\{\mathsf{DR}\}a$ and $t\{\mathsf{ONE}\}a$. Now assume that s and t are not connected; then there is no path consisting only of EQ-labels between s and t. The graph G consists of at least 2 components, and s, t are in different components. We define a consistent configuration as follows: For all nodes v, v' in the component in which s is contained, let $v\{\mathsf{DR}\}a$ and $v\{\mathsf{EQ}\}v'$. For all nodes v, v' in the component of t let $v\{\mathsf{ONE}\}a$ and $v\{\mathsf{EQ}\}v'$. For all nodes v, v' in the other components let $v\{\mathsf{DR}\}a$ and $v\{\mathsf{EQ}\}v'$. For all nodes v, v' which have not a label yet, let $v\{\mathsf{DR}\}v'$. (Two remarks : 1) EQ-edges for edges $E(v_i, v_j)$ in G with $j > i + 1$ are not shown in Fig. 4. 2) We inserted edges labelled $\mathcal{B}$ for better illustrations. But these are not needed.)

This lemma immediately entails the fact that satisfiability checking for ontologies over the logic DL-Lite$_{\mathcal{F},\mathcal{R}}^{\square,+}$(RCC3) is not FOL rewritable. This problem does not vanish if we presuppose that the ABox $\mathcal{A}$ is spatially complete—as shown by the following proposition.

Proposition 2. *Satisfiability checking of ontologies in DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$ (RCC3) with spatially complete ABoxes is not FOL rewritable.*

Proof. We construct a generic TBox $\mathcal{T}_g$ that allows one to encode any RCC3 constraint network so that checking the consistency of RCC3 constraint networks is reducible to a satisfiability check of this TBox and a spatially complete ABox. Let for every $r \in Rel_{RCC3}$ be given role symbols R_r^1, R_r^2. The generic TBox $\mathcal{T}_g$ has for every $r \in Rel_{RCC3}$ a concept symbol A_r and a corresponding axiom with the content that all instances of A_r have paths over the abstract features R_1 resp. R_2 to regions that are r-related.

$$\mathcal{T}_g = \{A_r \sqsubseteq \exists \tilde{R}_r^1, \tilde{R}_r^2.r, (\text{funct } l, R_r^1, R_r^2) \mid r \in Rel_{RCC3}\} \tag{1}$$

Now, let N be an arbitrary RCC3 constraint network which has to be tested for relational consistency. Let $\mathcal{A}_N$ be an ABox such that for every $r(a,b)$ in N three new constants are introduced: x_{ab}, x_a, x_b.

$$\mathcal{A}_N = \{A_r(x_{ab}), R_r^1(x_{ab}, x_a), R_r^2(x_{ab}, x_b) \mid r(a,b) \in N\} \tag{2}$$

The construction entails: $\mathcal{T}_g \cup \mathcal{A}_N \cup Ax_{RCC3}$ is satisfiable iff $N \cup Ax_{RCC3}$ is satisfiable. If the data complexity of the satisfiability check for DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC3)-ontologies were in AC^0, then the consistency of constraint networks could be tested in AC^0, too. (Note that $\mathcal{T}_g$ is a fixed TBox.) But checking the consistency of RCC3 constraint networks is LOGSPACE-hard and $AC^0 \subsetneq$ LOGSPACE.

As a corollary to this proposition we get the assertion that strong combinations of RCC5 and RCC8 into DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC5) and DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC8), respectively—which are defined in the same manner as in Definition 6—do not allow for FOL rewritability of satisfiability checking.

The low resolution calculus RCC2 is quite more inexpressive than RCC3 due to the fact that the composition table does not allow for the propagation of information: All compositions of DR, O result in the maximally unspecified relation $\{\mathsf{DR}, \mathsf{O}\}$. Hence, FOL rewritability of satisfiability testing follows easily considering the query $Q = \exists x, y[\mathsf{O}(x,y) \wedge \mathsf{DR}(x,y)] \vee \exists x[\mathsf{DR}(x,x)]$.

Proposition 3. *Testing the satisfiability of RCC2 networks is FOL rewritable.*

But in combination with functionality axioms of the TBox one could have the problem that the ABox may lead to identifications of regions. The identified regions are not allowed to have edges labelled O, DR resp. to the same region. Though this can be tested, the problem arises when a chain of regions is identified by the TBox and the ABox, because we do not know the length of the chain in advance. More formally: In addition to RCC2 constraint-network assertions we allow identity assertions $v = w$ for regions v, w. As we can assume that all nodes in a RCC2 network are connected by an edge labelled O, DR or $\mathcal{B}_{RCC2}$ we use a more intuitive formalism where, for every assertion $v = w$, the label of the edge between v and w is marked with an $=$; e.g., an edge between v, w with label $\mathsf{DR}^=$ stands for $\mathsf{DR}(v,w) \wedge v = w$. We call such a network an $=$-marked RCC2 network (a RCC$^=$2 network for short). Let $\mathcal{B} = \mathcal{B}_{RCC2}$ in the following.

Proposition 4. *An RCC$^=$2 constraint network N is unsatisfiable iff*

1. *N contains* $\mathsf{DR}(v,v)$ *or* $\mathsf{DR}^=(v,v)$ *for some node v; or*
2. *N contains* $\mathsf{DR}^=(v,w)$*; or*
3. *N contains a cycle in which there is* $\mathsf{DR}(v,w)$ *and in which there is a path from v to w such that every label on the path is* $\mathcal{B}^=$ *or* $\mathsf{O}^=$*; or*
4. *N contains a cycle in which there is* $\mathsf{DR}(v,w)$ *and in which there is a path from v to w s.t. every label on the path is* $\mathcal{B}^=$ *or* $\mathsf{O}^=$ *except one which is* O*.*

Proposition 4 shows that adding identity assertions to an RCC2 network may require checking the existence of identity chains of arbitrary length. Hence, in principle it is possible that the functional roles used in DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) may lead to identity chains. But as the following proposition show, this cannot be the case: The identity paths induced by functionalities in DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) can have only a maximal length of one.

Proposition 5. *Satisfiability checking of ontologies in DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) is FOL rewritable.*

5 Conclusion

As proved in this paper, combining DL-Lite with expressive fragments of the region calculus like RCC8 into logics that preserve the property of FOL rewritability is possible if the coupling is weak: Constraints of the RCC8 network contained in the ABox are not transported over to the implicitly constructed constraint network resulting from the constructors of the form $\exists U_1, U_2.r$. In this paper we further dealt with strong combinations for weaker calculi like RCC2 or RCC3. As we have shown by a reduction proof, a strong combination with RCC3 destroys the FOL rewritability of satisfiability checking. The reason is that checking the satisfiability of RCC3 networks needs to test for reachability along EQ paths, which can be reproduced by the TBox. For the low resolution calculus RCC2, FOL rewritability of satisfiability checking is provable—though checking the satisfiability of RCC2 networks with additional identity assertions is at least as hard as checking RCC3 networks. We plan to investigate whether DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) and DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap}$(RCC8) can be used for approximation—following the complete but not necessarily correct approximation method of [6]. Moreover we want to check whether DL-Lite$_{\mathcal{F},\mathcal{R}}^{\sqcap,+}$(RCC2) allows for FOL rewritability of query answering w.r.t. unions of conjunctive queries.

References

1. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The dl-lite family and relations. J. Artif. Intell. Res. (JAIR) 36, 1–69 (2009)
2. Calì, A., Gottlob, G., Lukasiewicz, T.: A general datalog-based framework for tractable query answering over ontologies. Technical Report CL-RR-10-21, Oxford University Computing Laboratory (November 2010)

3. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Poggi, A., Rodríguez-Muro, M., Rosati, R.: Ontologies and databases: The DL-Lite approach. In: Tessaris, S., Franconi, E. (eds.) Semantic Technologies for Informations Systems – 5th Int. Reasoning Web Summer School (RW-2009), LNCS, vol. 5689, pp. 255–356. Springer (2009)

4. Grigni, M., Papadias, D., Papadimitriou, C.H.: Topological inference. In: IJCAI (1). pp. 901–907 (1995)

5. Haarslev, V., Lutz, C., Möller, R.: A description logic with concrete domains and a role-forming predicate operator. J. Log. Comput. 9(3), 351–384 (1999)

6. Kaplunova, A., Moeller, R., Wandelt, S., Wessel, M.: Towards scalable instance retrieval over ontologies. In: Yaxin, B., Mary-Anne, W. (eds.) Knowledge Science, Engineering and Management, Fourth International Conference, KSEM 2010, Proceedings. Lecture Notes in Computer Science, vol. 6291. Springer (2010)

7. Kuper, G.M., Libkin, L., Paredaens, J. (eds.): Constraint Databases. Springer (2000)

8. Lutz, C., Miličić, M.: A tableau algorithm for description logics with concrete domains and general TBoxes. J. Autom. Reasoning 38(1-3), 227–259 (2007)

9. Özçep, Ö.L., Möller, R.: Combining DL-Lite with spatial calculi for feasible geothematic query answering. In: Kazakov, Y., Lembo, D., Wolter, F. (eds.) Proceedings of the 25th Iternational Workshop on Description Logics (DL 2012). vol. 846 (2012), `http://ceur-ws.org/Vol-846/`

10. Özçep, O.L., Möller, R.: Computationally feasible query answering over spatio-thematic ontologies. In: Proceedings of GEOProcessing 2012, The Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services (2012), available online at `http://www.thinkmind.org/index.php?view=article&articleid=geoprocessing_2012_7_10_30059`

11. Randell, D.A., Cui, Z., Cohn, A.G.: A spatial logic based on regions and connection. In: Proceedings of the 3rd International Conferecence on Knowledge Representation and Reasoning. pp. 165–176 (1992)

12. Reingold, O.: Undirected connectivity in log-space. J. ACM 55(4), 17:1–17:24 (Sep 2008), `http://doi.acm.org/10.1145/1391289.1391291`

13. Renz, J.: Qualitative Spatial Reasoning with Topological Information, Lecture Notes in Computer Science, vol. 2293. Springer (2002)

14. Renz, J., Nebel, B.: Qualitative spatial reasoning using constraint calculi. In: Aiello, M., Pratt-Hartmann, I., Benthem, J. (eds.) Handbook of Spatial Logics, pp. 161–215. Springer Netherlands (2007)

15. Rodriguez-Muro, M., Calvanese, D.: Semantic index: Scalable query answering without forward chaining or exponential rewritings. In: Posters of the 10th Int. Semantic Web Conf. (ISWC 2011) (2011)

16. Wessel, M.: On spatial reasoning with description logics - position paper. In: Ian Horrocks, S.T. (ed.) Proceedings of the International Workshop on Description Logics (DL-2002). CEUR Workshop Proceedings, vol. 53 (2002), http://CEUR-WS.org/Vol-53/

17. Wessel, M.: Qualitative spatial reasoning with the $\mathcal{ALCI}_{RCC}$- family — first results and unanswered questions. Technical Report FBI–HH–M–324/03, University of Hamburg, Department for Informatics (2003)