

Deciding FO-rewritability of Ontology-Mediated Queries in Linear Temporal Logic

Vladislav Ryzhikov¹, Yury Savateev^{1,2}, and Michael Zakharyashev¹

¹ Department of Computer Science, Birkbeck, University of London, U.K.

² National Research University Higher School of Economics, Moscow, Russia

Abstract. Aiming at ontology-based data access to temporal data, we investigate the problems of determining the data complexity of answering an ontology-mediated query (OMQ) given in linear temporal logic *LTL* and deciding whether it is rewritable to an FO($<$)-formula, possibly with extra built-in predicates. Using known facts about the complexity of regular languages, we show that OMQ answering in AC^0 coincides with FO($<$, $\equiv_{\mathbb{N}}$)-rewritability, which admits unary predicates $x \equiv 0 \pmod{n}$, and that deciding FO($<$)- and FO($<$, $\equiv_{\mathbb{N}}$)-rewritability of *LTL* OMQs is EXPSpace-complete. We further observe that answering any OMQ is either in ACC^0 , in which case it is FO($<$, MOD)-rewritable, or NC^1 -complete, and prove that distinguishing between these two cases can be done in EXPSpace. Finally, we identify fragments of *LTL* for which some of these decision problems become PSPACE-, Π_2^p - and coNP-complete.

1 Introduction

Classical ontology-based data access (OBDA) [8,20] was launched by identifying ontology and query languages that *uniformly* guarantee FO-rewritability of all ontology-mediated queries (OMQs) given in those languages. Thus, by design, OBDA ontologies are rather inexpressive. An alternative, *non-uniform* approach to OBDA would be—at least in theory—to develop algorithms that, given any OMQ in some expressive languages, could recognise the data complexity of answering that OMQ and construct its rewriting of optimal type. The datalog community has been investigating FO- and linear-datalog-rewritability (aka boundedness and linearisability) of datalog programs since the 1980s [26,25,11,19]. The data complexity and rewritability of individual OMQs in various description logics have become an active research area in the past decade [17,7,16,18,15].

Here we take first steps towards extending the non-uniform analysis to OBDA over temporal data (see [3] for a survey of results in uniform temporal OBDA). We consider OMQs given in linear temporal logic *LTL*, which were uniformly classified in [2,4] according to their data complexity and rewritability type.

Example 1. Let \mathcal{O} be an *LTL* ontology with the following axioms containing the temporal operators \diamond_F (eventually) and \circ_F/\circ_P (next/previous minute):

$$\text{Malfunction} \rightarrow \diamond_F \text{Fixed}, \quad (1)$$

$$\text{Fixed} \rightarrow \circ_F \text{InOperation}, \quad (2)$$

$$Malfunction \wedge \circ_P Malfunction \wedge \circ_F^2 Malfunction \rightarrow \neg \circ_F InOperation. \quad (3)$$

We query temporal data, say the ABox

$$\mathcal{A} = \{Malfunction(2), Malfunction(5), Malfunction(6), Fixed(6), Malfunction(7)\}$$

using *LTL*-formulas such as

$$\varkappa = Malfunction \wedge \bigvee_{1 \leq i \leq 5} \circ_F^i (Fixed \wedge \bigvee_{1 \leq j \leq 5} \neg \circ_F^j InOperation),$$

which can be understood as a Boolean query asking whether there was a malfunction that was fixed in $\leq 5m$ but within the next 5m the equipment went out of operation again. The certain answer to the OMQ (\mathcal{O}, \varkappa) over \mathcal{A} is yes.

Our aim in this paper is to understand the complexity of deciding whether a given *LTL* OMQ is rewritable to an $FO(<)$ -formula with the order relation $<$ over timestamps and possibly other built-in predicates. As shown in [4], every *LTL* OMQ is rewritable to an $FO(<)$ -formula extended with relational primitive recursion (or to a monadic second-order formula), whose evaluation over data instances is in the complexity class NC^1 . Here, we first establish a connection between *LTL* OMQs and regular languages, and then use it to prove that deciding $FO(<)$ -rewritability of such OMQs is $EXSPACE$ -complete, with the lower bound shown for Horn *LTL* ontologies with the next-time operator \circ_F and atomic queries, and also for Krom ontologies with positive *LTL* queries. For atomic OMQs (OMAQs, for short) with linear Horn *LTL* ontologies that contain \circ_P only, deciding $FO(<)$ -rewritability turns out to be $PSPACE$ -complete; the complexity goes down to $CONP$ for OMAQs with Krom ontologies and the next- and previous-time operators. On the other hand, deciding $FO(<)$ -rewritability becomes Π_2^P -complete for core (that is, both Horn and Krom) ontologies and *positive existential* temporal queries, which do not contain negation and the operators always in the future/past, since and until. OMQs with such queries are referred to as OMPEQs.

Using the connection with regular languages and the seminal results of [5], we show that OMQ answering in AC^0 coincides with rewritability to $FO(<, \equiv_{\mathbb{N}})$ -formulas, which admit unary predicates $x \equiv 0 \pmod{n}$, and that deciding $FO(<, \equiv_{\mathbb{N}})$ -rewritability of *LTL* OMQs is $EXSPACE$ -complete. We further observe that answering any OMQ is either in ACC^0 , in which case it is $FO(<, MOD)$ -rewritable, or NC^1 -complete, and prove that distinguishing between these two cases can be done in $EXSPACE$. For OMAQs with linear Horn *LTL* ontologies with \circ_P only, these problems become decidable in $PSPACE$. All our complexity results for circuit complexity and rewritability of OMQs are summarised below:

class of OMQs	in AC^0		in ACC^0 / NC^1 -comp.
	$FO(<)$	$FO(<, \equiv_{\mathbb{N}})$	$FO(<, MOD)/FO(RPR)$
<i>LTL</i> OMQs			
$LTL_{horn}^{\circ_P}$ OMAQs	$EXSPACE$ [Th.1,5,6]	$EXSPACE$ [Th.4,5,6]	$\leq EXSPACE$ [Th.4]
LTL_{krom}° OMPEQs			
lin. $LTL_{horn}^{\circ_P}$ OMAQs	$PSPACE$ [Th. 7]	$PSPACE$ [Th. 7]	$\leq PSPACE$ [Th. 7]
LTL_{krom}° OMAQs	$CONP$ [Th. 8]	all in AC^0 [4]	–
LTL_{core}° OMPEQs	Π_2^P [Th. 9]	all in AC^0 [4]	–

2 Preliminaries

In our setting, the alphabet of linear temporal logic *LTL* comprises a set of *atomic concepts* A_i , $i < \omega$. *Basic temporal concepts*, C , are defined by the grammar

$$C ::= A_i \mid \Box_F C \mid \Box_P C \mid \bigcirc_F C \mid \bigcirc_P C$$

with the *operators* \Box_F/\Box_P (always in the future/past) and \bigcirc_F/\bigcirc_P (at the next/previous moment). A *temporal ontology*, \mathcal{O} , is a finite set of *axioms* of the form

$$C_1 \wedge \cdots \wedge C_k \rightarrow C_{k+1} \vee \cdots \vee C_{k+m}, \quad (4)$$

where $k, m \geq 0$, the C_i are basic temporal concepts, the empty \wedge is \top , and the empty \vee is \perp . Following the *DL-Lite* convention [1,2], we classify ontologies by the shape of their axioms and the temporal operators that can occur in them. Suppose $\mathbf{c} \in \{\text{horn}, \text{krom}, \text{core}, \text{bool}\}$ and $\mathbf{o} \in \{\Box, \bigcirc, \Box\bigcirc\}$. The axioms of an *LTL $_{\mathbf{c}}^{\mathbf{o}}$ -ontology* may only contain occurrences of the (future and past) temporal operators in \mathbf{o} and satisfy the following restrictions on k and m in (4) indicated by \mathbf{c} : *horn* requires $m \leq 1$, *krom* requires $k + m \leq 2$, *core* both $k + m \leq 2$ and $m \leq 1$, while *bool* imposes no restrictions. For example, axiom (2) from Example 1 is allowed in all of these fragments, (3) is equivalent to a *horn* axiom (with \perp on the right-hand side), and (1) can be expressed in *krom* as explained in Remark 1. A basic concept is called an *IDB* (intensional database) *concept* in an ontology \mathcal{O} if its atomic concept occurs on the right-hand side of an axiom in \mathcal{O} . The set of IDB atomic concepts in \mathcal{O} is denoted by $\text{idb}(\mathcal{O})$. An *LTL $_{\text{horn}}^{\bigcirc}$ -ontology* is called *linear* if each of its axioms $C_1 \wedge \cdots \wedge C_k \rightarrow B$ is such that B is either \perp or *atomic* and *at most one* C_i , $1 \leq i \leq k$, is an IDB concept.

An *ABox* is a finite set \mathcal{A} of atoms $A_i(\ell)$, for $\ell \in \mathbb{Z}$, together with a finite interval $\text{tem}(\mathcal{A}) = [\min \mathcal{A}, \max \mathcal{A}]$ of integers such that $\min \mathcal{A} < \max \mathcal{A}$ and whenever $A_i(\ell) \in \mathcal{A}$ then $\min \mathcal{A} \leq \ell \leq \max \mathcal{A}$. Without loss of generality, we always assume that $\min \mathcal{A} = 0$. The interval $\text{tem}(\mathcal{A})$ is called the *active domain* of \mathcal{A} . If $\text{tem}(\mathcal{A})$ is not specified explicitly, it is assumed to be $[0, m]$, where m is the maximal timestamp in \mathcal{A} . By a *signature*, Ξ , we mean any finite set of atomic concepts. An ABox \mathcal{A} is said to be a Ξ -ABox if $A_i(\ell) \in \mathcal{A}$ implies $A_i \in \Xi$.

We query ABoxes by means of *temporal concepts*, \varkappa , which are *LTL*-formulas built from the atoms A_i , Booleans \wedge, \vee, \neg , temporal operators $\bigcirc_F, \Box_F, \diamond_F$ (eventually), \mathcal{U} (until), and their past-time counterparts $\bigcirc_P, \Box_P, \diamond_P$ (some time in the past) and \mathcal{S} (since). If \varkappa does not contain $\neg, \Box_F, \Box_P, \mathcal{U}$ and \mathcal{S} , we call it a *positive existential temporal concept*.

A (temporal) *interpretation* is a structure $\mathcal{I} = (\mathbb{Z}, A_0^{\mathcal{I}}, A_1^{\mathcal{I}}, \dots)$ with $A_i^{\mathcal{I}} \subseteq \mathbb{Z}$, for every $i < \omega$. The *extension* $\varkappa^{\mathcal{I}}$ of a temporal concept \varkappa in \mathcal{I} is defined inductively as usual in *LTL* under the ‘strict semantics’ [14,12]:

$$\begin{aligned} (\bigcirc_F \varkappa)^{\mathcal{I}} &= \{n \in \mathbb{Z} \mid n+1 \in \varkappa^{\mathcal{I}}\}, & (\Box_F \varkappa)^{\mathcal{I}} &= \{n \in \mathbb{Z} \mid k \in \varkappa^{\mathcal{I}}, \text{ for all } k > n\}, \\ (\diamond_F \varkappa)^{\mathcal{I}} &= \{n \in \mathbb{Z} \mid \text{there is } k > n \text{ with } k \in \varkappa^{\mathcal{I}}\}, \\ (\varkappa_1 \mathcal{U} \varkappa_2)^{\mathcal{I}} &= \{n \in \mathbb{Z} \mid \text{there is } k > n \text{ with } k \in \varkappa_2^{\mathcal{I}} \text{ and } m \in \varkappa_1^{\mathcal{I}} \text{ for } n < m < k\}, \end{aligned}$$

and symmetrically for the past operators. An axiom (4) is *true* in \mathcal{I} if we have $C_1^{\mathcal{I}} \cap \dots \cap C_k^{\mathcal{I}} \subseteq C_{k+1}^{\mathcal{I}} \cup \dots \cup C_{k+m}^{\mathcal{I}}$. An interpretation \mathcal{I} is a *model* of \mathcal{O} if all axioms of \mathcal{O} are true in \mathcal{I} ; it is a *model* of \mathcal{A} if $A_i(\ell) \in \mathcal{A}$ implies $\ell \in A_i^{\mathcal{I}}$.

An *LTL_c^o ontology-mediated query* (OMQ) is a pair of the form $\mathbf{q} = (\mathcal{O}, \varkappa)$, where \mathcal{O} is an *LTL_c^o* ontology and \varkappa a temporal concept. If \varkappa is a positive existential temporal concept, we call \mathbf{q} a *positive existential OMQ* (OMPEQ), and if \varkappa is an atomic concept, we call \mathbf{q} *atomic* (OMAQ). The set of atomic concepts occurring in an OMQ \mathbf{q} is denoted by $\text{sig}(\mathbf{q})$. We can treat \mathbf{q} as a *Boolean* OMQ, which returns yes/no as an answer, or as a *specific* OMQ, which returns timestamps from the ABox in question assigned to the free variable, say x , in the standard FO-translation of \varkappa . In the latter case, we write $\mathbf{q}(x) = (\mathcal{O}, \varkappa(x))$.

More precisely, a *certain answer* to a Boolean OMQ $\mathbf{q} = (\mathcal{O}, \varkappa)$ over an ABox \mathcal{A} is *yes* if, for every model \mathcal{I} of \mathcal{O} and \mathcal{A} , there is $k \in \mathbb{Z}$ such that $k \in \varkappa^{\mathcal{I}}$, in which case we write $(\mathcal{O}, \mathcal{A}) \models \exists x \varkappa(x)$. If $(\mathcal{O}, \mathcal{A}) \not\models \exists x \varkappa(x)$, the certain answer to \mathbf{q} over \mathcal{A} is *no*. We write $(\mathcal{O}, \mathcal{A}) \models \varkappa(k)$, for $k \in \mathbb{Z}$, if $k \in \varkappa^{\mathcal{I}}$ in all models \mathcal{I} of \mathcal{O} and \mathcal{A} . A *certain answer* to a specific OMQ $\mathbf{q}(x) = (\mathcal{O}, \varkappa(x))$ over \mathcal{A} is any $k \in \text{tem}(\mathcal{A})$ with $(\mathcal{O}, \mathcal{A}) \models \varkappa(k)$. By the *evaluation* (or *answering*) *problems* for \mathbf{q} and $\mathbf{q}(x)$ we understand the decision problems ‘ $(\mathcal{O}, \mathcal{A}) \models^? \exists x \varkappa(x)$ ’ and ‘ $(\mathcal{O}, \mathcal{A}) \models^? \varkappa(k)$ ’ with input \mathcal{A} and, resp., \mathcal{A} and $k \in \text{tem}(\mathcal{A})$. We say that \mathbf{q} or $\mathbf{q}(x)$ is in a complexity class \mathcal{C} if the corresponding evaluation problem is in \mathcal{C} .

Example 2. (i) Let $\mathbf{q}_1 = (\mathcal{O}_1, C \wedge D)$ with $\mathcal{O}_1 = \{\diamond_P A \rightarrow B, \square_F B \rightarrow C\}$. The certain answer to \mathbf{q}_1 over $\mathcal{A}_1 = \{D(0), B(1), A(1)\}$ is *yes*, but over $\mathcal{A}_2 = \{D(0), A(1)\}$ it is *no*. The only answer to $\mathbf{q}_1(x) = (\mathcal{O}_1, (C \wedge D)(x))$ over \mathcal{A}_1 is 0. (ii) Let $\mathcal{O}_2 = \{\circ_P A \rightarrow B, \circ_P B \rightarrow A, A \wedge B \rightarrow \perp\}$. The answer to $\mathbf{q}_2 = (\mathcal{O}_2, C)$ over $\mathcal{A}_1 = \{A(0)\}$ is *no*, but over $\mathcal{A}_2 = \{A(0), A(1)\}$ it is *yes*. There are no answers to $\mathbf{q}_2(x) = (\mathcal{O}_2, C(x))$ over \mathcal{A}_1 , while over \mathcal{A}_2 there are two of them: 0 and 1. (iii) Let $\mathcal{O}_3 = \{\circ_P B_k \wedge A_0 \rightarrow B_k, \circ_P B_{1-k} \wedge A_1 \rightarrow B_k \mid k = 0, 1\}$. For any word $e = e_1 \dots e_n \in \{0, 1\}^n$, let $\mathcal{A}_e = \{B_0(0)\} \cup \{A_{e_i}(i) \mid 0 < i \leq n\} \cup \{E(n)\}$. The answer to $\mathbf{q}_3 = (\mathcal{O}_3, B_0 \wedge E)$ over \mathcal{A}_e is *yes* iff the number of 1s in e is even. (iv) Let $\mathcal{O}_4 = \{A \rightarrow \circ_F B\}$ and $\mathbf{q}_4 = (\mathcal{O}_4, B)$. Then, the answer to \mathbf{q}_4 over $\mathcal{A} = \{A(0)\}$ is *yes*; however, there are no certain answers to $\mathbf{q}_4(x) = (\mathcal{O}_4, B(x))$ over \mathcal{A} . (v) Let $\mathcal{O}_5 = \{A \rightarrow B \vee \circ_F B\}$. The certain answer to $\mathbf{q}_5 = (\mathcal{O}_5, B)$ over $\mathcal{A} = \{A(0), C(1)\}$ is *yes*; however, there are no certain answers to $\mathbf{q}_5(x)$ over \mathcal{A} .

Remark 1. Let \mathcal{O} be as in Example 1 and let \mathcal{O}' be the result of replacing (1) in \mathcal{O} by *Malfunction* $\wedge \square_F X \rightarrow \perp$ and $\top \rightarrow X \vee \text{Fixed}$, for a fresh concept name X . Then the OMQ $\mathbf{q} = (\mathcal{O}, \varkappa)$ in Example 1 is ‘equivalent’ to $\mathbf{q}' = (\mathcal{O}', \varkappa)$ in the sense that \mathbf{q} and \mathbf{q}' have the same certain answers over any $\text{sig}(\mathbf{q})$ -ABox \mathcal{A} .

Let \mathcal{L} be a class of FO-formulas that can be interpreted over finite linear orders. A Boolean OMQ \mathbf{q} is *\mathcal{L} -rewritable over Ξ -ABoxes* if there is an \mathcal{L} -sentence \mathbf{Q} such that, for any Ξ -ABox \mathcal{A} , the certain answer to \mathbf{q} over \mathcal{A} is *yes* iff $\mathfrak{S}_{\mathcal{A}} \models \mathbf{Q}$. Here, $\mathfrak{S}_{\mathcal{A}}$ is a structure with domain $\text{tem}(\mathcal{A})$ ordered by $<$, in which $\mathfrak{S}_{\mathcal{A}} \models A_i(\ell)$ iff $A_i(\ell) \in \mathcal{A}$. A specific OMQ $\mathbf{q}(x)$ is *\mathcal{L} -rewritable over Ξ -ABoxes* if there is an \mathcal{L} -formula $\mathbf{Q}(x)$ with one free variable x such that, for any Ξ -ABox \mathcal{A} , k is a certain answer to \mathbf{q} over \mathcal{A} iff $\mathfrak{S}_{\mathcal{A}} \models \mathbf{Q}(k)$. The sentence \mathbf{Q} and the formula $\mathbf{Q}(x)$ are called *\mathcal{L} -rewritings* of the Boolean and specific OMQ \mathbf{q} , respectively.

We require four languages \mathcal{L} for rewriting *LTL* OMQs:

- FO(<)** (monadic) first-order formulas with the built-in predicate $<$ for order;
- FO(<, $\equiv_{\mathbb{N}}$)** FO(<)-formulas with predicates $x \equiv 0 \pmod{N}$, for any $N > 1$;
- FO(<, MOD)** FO(<)-formulas with quantifiers \exists^N , for $N > 1$, defined by taking $\mathfrak{S}_{\mathcal{A}} \models \exists^N x \psi(x)$ iff the cardinality of the set $\{n \in \text{tem}(\mathcal{A}) \mid \mathfrak{S}_{\mathcal{A}} \models \psi(n)\}$ is divisible by N ($x \equiv 0 \pmod{N}$ can obviously be defined as $\exists^N y (y < x)$);
- FO(RPR)** FO(<) with relational primitive recursion [10].

Example 3. (i) An FO(<)-rewriting of $\mathbf{q}'_1(x)$ is

$$\varphi_1(x) = D(x) \wedge [C(x) \vee \exists y (A(y) \wedge \forall z ((x < z \leq y) \rightarrow B(z)))],$$

$\exists x \varphi_1(x)$ is an FO(<)-rewriting of \mathbf{q}_1 . (ii) An FO(<, $\equiv_{\mathbb{N}}$)-rewriting of $\mathbf{q}_2(x)$ is

$$\varphi_2(x) = C(x) \vee \exists x, y [(A(x) \wedge A(y) \wedge \text{odd}(x, y)) \vee (B(x) \wedge B(y) \wedge \text{odd}(x, y)) \vee (A(x) \wedge B(y) \wedge \neg \text{odd}(x, y))],$$

where $\text{odd}(x, y) = (x \equiv 0 \pmod{2}) \leftrightarrow y \not\equiv 0 \pmod{2}$ implies that the distance between x and y is odd, and an FO(<, $\equiv_{\mathbb{N}}$)-rewriting of \mathbf{q}_2 is $\exists x \varphi_2(x)$. (iii) The OMQ \mathbf{q}_3 is not rewritable to an FO-formula with any numeric predicates as PARITY is not in AC^0 [13]; the following is an FO(<, MOD)-rewriting of \mathbf{q}_3 :

$$\varphi_3 = \exists x, y (E(x) \wedge (x \leq y) \wedge \forall z ((y < z \leq x) \rightarrow (A_0(z) \vee A_1(z))) \wedge ((B_0(y) \wedge \exists^2 z ((y < z \leq x) \wedge A_1(z))) \vee (B_1(y) \wedge \neg \exists^2 z ((y < z \leq x) \wedge A_1(z)))).$$

(iv) An FO(<)-rewriting of $\mathbf{q}_4(x)$ is $\varphi_4(x) = B(x) \vee A(x-1)$; an FO(<)-rewriting of \mathbf{q}_4 is $\varphi_4 = \exists x (A(x) \vee B(x))$. (v) The same φ_4 is an FO(<)-rewriting of \mathbf{q}_5 , and $B(x)$ is a rewriting of $\mathbf{q}_5(x)$.

A *uniform* classification of specific *LTL* OMQs by their rewritability type has been obtained in [4]. Here, we only mention in passing that all (Boolean and specific) *LTL* OMQs are FO(RPR)-rewritable and can be answered in NC^1 . In this paper, we take a *non-uniform* approach to rewritability, aiming to understand how (complex it is) to decide the optimal type of FO-rewritability for a given *LTL* OMQ \mathbf{q} over Ξ -ABoxes. Clearly, we can always assume that $\Xi \subseteq \text{sig}(\mathbf{q})$.

For any \mathbf{q} and $\Xi \subseteq \text{sig}(\mathbf{q})$, we regard the set $\Sigma_{\Xi} = 2^{\Xi}$ as an *alphabet*. A Ξ -ABox \mathcal{A} can be given as a Σ_{Ξ} -word $w_{\mathcal{A}} = a_0 \dots a_n$ with $a_i = \{A \mid A(i) \in \mathcal{A}\}$. Conversely, any Σ_{Ξ} -word $w = a_0 \dots a_n$ can be understood as an ABox \mathcal{A}_w with $\text{tem}(\mathcal{A}_w) = [0, n]$ and $A(i) \in \mathcal{A}_w$ iff $A \in a_i$. The *language* $\mathbf{L}_{\Xi}(\mathbf{q})$ of the Boolean OMQ \mathbf{q} is the set of Σ_{Ξ} -words $w_{\mathcal{A}}$ such that the certain answer to \mathbf{q} over \mathcal{A} is **yes**. For a specific $\mathbf{q}(x)$, we take $\Gamma_{\Xi} = \Sigma_{\Xi} \cup \Sigma'_{\Xi}$, for a disjoint copy Σ'_{Ξ} of Σ_{Ξ} , and represent a pair (\mathcal{A}, i) with a Ξ -ABox \mathcal{A} and $i \in \text{tem}(\mathcal{A})$ as a Γ_{Ξ} -word $w_{\mathcal{A}, i} = a_0 \dots a'_i \dots a_n$, where $a_j = \{A \mid A(j) \in \mathcal{A}\} \in \Sigma_{\Xi}$ for $j \neq i$, and $a'_i = \{A \mid A(i) \in \mathcal{A}\} \in \Sigma'_{\Xi}$. The *language* $\mathbf{L}_{\Xi}(\mathbf{q}(x))$ is the set of Γ_{Ξ} -words $w_{\mathcal{A}, i}$ such that i is a certain answer to $\mathbf{q}(x)$ over \mathcal{A} .

Proposition 1. *Let \mathcal{L} be one of the classes of FO-formulas introduced above. (i) A Boolean OMQ $\mathbf{q} = (\mathcal{O}, \varkappa)$ is \mathcal{L} -rewritable over Ξ -ABoxes iff $\mathbf{L}_\Xi(\mathbf{q})$ is \mathcal{L} -definable. (ii) A specific OMQ $\mathbf{q}(x) = (\mathcal{O}, \varkappa(x))$ is \mathcal{L} -rewritable over Ξ -ABoxes iff $\mathbf{L}_\Xi(\mathbf{q}(x))$ is \mathcal{L} -definable. Both $\mathbf{L}_\Xi(\mathbf{q})$ and $\mathbf{L}_\Xi(\mathbf{q}(x))$ are regular for any Ξ .*

Proof. We only show that $\mathbf{L}_\Xi(\mathbf{q})$ is regular. Let $\text{sub}_{\mathbf{q}}$ be the set of temporal concepts in \mathbf{q} and their negations. A *type* is any maximal subset τ of $\text{sub}_{\mathbf{q}}$ that is consistent with \mathcal{O} . The set of all types is denoted by \mathbf{T} . We define an NFA \mathfrak{A} over Σ_Ξ^* whose language is $\Sigma_\Xi^* \setminus \mathbf{L}_\Xi(\mathbf{q})$. The states in \mathfrak{A} comprise the set $Q_{\neg\varkappa} = \{\tau \in \mathbf{T} \mid \neg\varkappa \in \tau\}$. The transition relation \rightarrow_a , for $a \in \Sigma_\Xi$, is defined by setting $\tau_1 \rightarrow_a \tau_2$ if the following conditions hold (assuming that the temporal operators are expressed via \mathcal{U} and \mathcal{S}): $a \subseteq \tau_2$, $C_1 \mathcal{U} C_2 \in \tau_1$ iff $C_2 \in \tau_2$ or $C_1 \mathcal{U} C_2 \in \tau_2$ and $C_1 \in \tau_2$, and symmetrically for \mathcal{S} . The set of initial states comprises $\tau \in Q_{\neg\varkappa}$ with $\tau \cup \{\square_F \neg\varkappa\}$ is consistent with \mathcal{O} ; the set of accepting states comprises those $\tau \in Q_{\neg\varkappa}$ for which $\tau \cup \{\square_F \neg\varkappa\}$ is consistent with \mathcal{O} . It is readily seen that, for every $\mathbf{a} \in \Sigma_\Xi^*$ we have $\mathbf{a} \in \mathbf{L}(\mathfrak{A})$ iff $(\mathcal{O}, \mathcal{A}_{\mathbf{a}}) \not\models \exists x \varkappa(x)$. The number of states in \mathfrak{A} does not exceed $O(2^{|\mathbf{q}|})$. Since *LTL*-satisfiability is in PSPACE, the NFA \mathfrak{A} can be constructed in exponential time in $|\mathbf{q}|$.

The following table summarises known results connecting definability of regular languages \mathbf{L} with properties of their syntactic monoids $M(\mathbf{L})$ and syntactic morphisms $\eta_{\mathbf{L}}$ (see Section 3 and [5] for details) and with their circuit complexity (under a reasonable binary encoding of \mathbf{L} 's alphabet):

definability of \mathbf{L}	algebraic characterisation of \mathbf{L}	circuit complexity
FO($<$)	$M(\mathbf{L})$ is aperiodic	in AC^0
FO($<, \equiv_{\mathbb{N}}$)	$\eta_{\mathbf{L}}$ is quasi-aperiodic	
FO($<, \text{MOD}$)	all groups in $M(\mathbf{L})$ are solvable	in ACC^0
FO(RPR)	arbitrary $M(\mathbf{L})$	in NC^1
–	$M(\mathbf{L})$ contains unsolvable group	NC^1 -hard

The statement in the table that all groups in $M(\mathbf{L})$ are solvable iff \mathbf{L} is in ACC^0 holds unless $\text{ACC}^0 = \text{NC}^1$. Using Proposition 1, these results can be extended to rewritability and data complexity of Boolean and specific *LTL* OMQs: (a) an OMQ is FO($<, \equiv_{\mathbb{N}}$)-rewritable iff it can be answered in AC^0 , (b) an OMQ is FO($<, \text{MOD}$)-rewritable iff it can be answered in ACC^0 (unless $\text{ACC}^0 = \text{NC}^1$), (c) an OMQ is FO($<, \text{RPR}$)-rewritable iff it can be answered in NC^1 .

3 Deciding FO-Rewritability: Upper Bounds

Since deciding FO($<$)-definability of regular languages given by NFAs is PSPACE-complete [9,6], we obtain by Proposition 1:

Theorem 1. *Deciding FO($<$)-rewritability of $\text{LTL}_{\text{bool}}^{\square\circ}$ OMQs over Ξ -ABoxes is in EXPSpace.*

The exact complexity of deciding FO($<, \equiv_{\mathbb{N}}$)-definability and NC^1 -hardness of regular languages seems to be open (their decidability was shown in [5].) So

our first aim is to settle these issues. Given an NFA $\mathfrak{A} = (Q, \Sigma, \delta, Q_0, F)$, states $q, q' \in Q$, and $w = a_0 \dots a_n \in \Sigma^*$, we write $q \rightarrow_w q'$ if there is a run of \mathfrak{A} on w that starts with $(q_0, 0)$ and ends with $(q', n + 1)$. We say that a state $q \in Q$ is *accessible* if $q' \rightarrow_w q$, for some $q' \in Q_0$ and $w \in \Sigma^*$. Two states $q_1, q_2 \in Q$ are *equivalent* if, for each $w \in \Sigma^*$, we have $q_1 \rightarrow_w q'$ for some $q' \in F$ iff $q_2 \rightarrow_w q''$ for some $q'' \in F$. A DFA is *minimal* if each of its states is accessible and it has no distinct equivalent states. Every DFA $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$ can be converted to a minimal DFA $\mathfrak{A}' = (Q', \Sigma, \delta', q'_0, F')$ with $\mathbf{L}(\mathfrak{A}) = \mathbf{L}(\mathfrak{A}')$ in the following way [23]. Let $R = \{q \in Q \mid q \text{ is accessible in } \mathfrak{A}\}$ and let \sim be a relation on R defined by taking $q \sim q'$ iff q and q' are equivalent. Clearly, \sim is an equivalence relation; we denote by q/\sim the equivalence class of $q \in R$. Now, we set $Q' = \{q/\sim \mid q \in R\}$ and define δ' by taking $q/\sim \rightarrow_a q'/\sim$, where $\{q'\} = \delta(a, q)$, for all $q \in R$ and $a \in \Sigma$ (which is obviously well-defined). Finally, we set $q'_0 = q_0/\sim$ and $F' = \{q/\sim \mid q \in R \cap F\}$. It is known that, for any regular language \mathbf{L} , all minimal DFAs \mathfrak{A}' with $\mathbf{L}(\mathfrak{A}') = \mathbf{L}$ are *isomorphic*; we call each such \mathfrak{A}' a *minimal DFA of \mathbf{L}* .

A monoid $M = (B, \cdot, e)$ has an associative binary operation \cdot and an identity e with $a \cdot e = e \cdot a = a$, for all $a \in B$. We shorten $a \cdot b$ to ab . Given a DFA $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$ and $w \in \Sigma^*$, define a map $f_w^\mathfrak{A}: Q \rightarrow Q$ by setting $f_w^\mathfrak{A}(q) = q'$ iff $q \rightarrow_w q'$. The *transition monoid* of \mathfrak{A} takes the form $M = (\{f_w^\mathfrak{A} \mid w \in \Sigma^*\}, \cdot, f_\varepsilon^\mathfrak{A})$, where ε is the empty word and $f_w^\mathfrak{A} f_v^\mathfrak{A} = f_{wv}^\mathfrak{A}$, for any $f_w^\mathfrak{A}, f_v^\mathfrak{A}$. The *syntactic monoid* $M(\mathbf{L})$ of a regular language \mathbf{L} is isomorphic to the transition monoid of a minimal DFA accepting \mathbf{L} [23, Chapter V.1]. A monoid is *aperiodic* if it does not contain non-trivial groups (with the monoid operation). Let \mathfrak{A} be a minimal automaton of \mathbf{L} and B the domain of $M(\mathbf{L})$. The map $\eta_\mathbf{L}: \Sigma^* \rightarrow B$ defined by $\eta_\mathbf{L}(w) = f_w^\mathfrak{A}$ is called a *syntactic morphism* of \mathbf{L} . Given a set $W \subseteq \Sigma^*$, we set $\eta_\mathbf{L}(W) = \{\eta_\mathbf{L}(w) \mid w \in W\}$. The syntactic morphism $\eta_\mathbf{L}$ is *quasi-aperiodic* if, for any $t > 0$, the set $\eta_\mathbf{L}(\Sigma^t)$ does not contain non-trivial groups.

Theorem 2. *Deciding $\text{FO}(<, \equiv_{\mathbb{N}})$ -definability of $\mathbf{L}(\mathfrak{A})$, \mathfrak{A} an NFA, is in PSPACE.*

Proof. First, we show the theorem for a minimal DFA \mathfrak{A} , then extend it to an arbitrary DFA and, finally, to an NFA. Let $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$ be minimal. We use the following criterion: $\mathbf{L}(\mathfrak{A})$ is not $\text{FO}(<, \equiv_{\mathbb{N}})$ -definable iff there are w and $n \in \mathbb{N}$ such that $f_w^\mathfrak{A} \neq f_{w^2}^\mathfrak{A}$, $f_w^\mathfrak{A} = f_{w^n}^\mathfrak{A}$, and $f_w^\mathfrak{A} = f_v^\mathfrak{A}$, $f_{w^2}^\mathfrak{A} = f_u^\mathfrak{A}$, for some u and v with $|v| = |u|$. Indeed, let $\mathbf{L} = \mathbf{L}(\mathfrak{A})$. (\Rightarrow) In this case, $\eta_\mathbf{L}$ is quasi-aperiodic, and so there is t such that $\eta_\mathbf{L}(\Sigma^t)$ contains a non-trivial group G . Let e be the identity element of G and let $a \neq e$, $a \in G$. We have $a^{|G|} = e$, $a^{|G|+1} = ae = a$ and, since $a, e \in \eta_\mathbf{L}(\Sigma^t)$, there are $w, u \in \Sigma^t$ such that $a = f_w^\mathfrak{A}$ and $e = f_{w^{|G|}}^\mathfrak{A} = f_u^\mathfrak{A}$. (\Leftarrow) Observe that $f_{w^i}^\mathfrak{A} = f_{w^i w^{n(n-i)}}^\mathfrak{A} = f_{u^i v^{n-i}}^\mathfrak{A}$, for $1 \leq i \leq n$. Therefore, $f_w^\mathfrak{A}, \dots, f_{w^n}^\mathfrak{A}$ form a group in $\eta_\mathbf{L}(\Sigma^{|u| \cdot n})$, and so \mathbf{L} is not $\text{FO}(<, \equiv_{\mathbb{N}})$ -definable.

To check this criterion, we can use the known PSPACE algorithms [9,6] for checking $\text{FO}(<)$ -definability of $\mathbf{L}(\mathfrak{A})$. We now show how to extend this result to any DFA \mathfrak{A} . Let Q^r be the set of *accessible* states in \mathfrak{A} . We call words $w, v \in \Sigma^*$ *equivalent* in \mathfrak{A} and write $w \equiv_{\mathfrak{A}} v$ if whenever $q \rightarrow_w q'$ then there is $q'' \in Q^r$ such that $q'' \sim q'$ and $q \rightarrow_v q''$, and the other way round. Let \mathfrak{A}^* be a minimal

DFA of $\mathbf{L}(\mathfrak{A})$. One can show that $w \equiv_{\mathfrak{A}} v$ iff $f_w^{\mathfrak{A}^*} = f_v^{\mathfrak{A}^*}$. This implies that, in the criterion above, we can replace every $f_x^{\mathfrak{A}} = f_y^{\mathfrak{A}}$ by $x \equiv_{\mathfrak{A}} y$ and obtain the same criterion for an arbitrary DFA, which is checkable in PSPACE. We can finally obtain a criterion of $\text{FO}(<, \equiv_{\mathbb{N}})$ -definability of a language given by an NFA by replacing every $f_x^{\mathfrak{A}} = f_y^{\mathfrak{A}}$ by $x \equiv_{\mathfrak{A}'} y$, where \mathfrak{A}' is the powerset automaton of \mathfrak{A} . To show that the latter criterion can be checked in PSPACE, we observe that each state of \mathfrak{A}' can be stored using polynomial space and then adjust the algorithm for DFAs without increasing its complexity.

Theorem 3. NC^1 -hardness of $\mathbf{L}(\mathfrak{A})$, for an NFA \mathfrak{A} , can be decided in PSPACE.

Proof. We follows that steps of the proof above, using the following criterion. The language $\mathbf{L}(\mathfrak{A})$, for a minimal DFA \mathfrak{A} , is NC^1 -hard iff there are $u, v, w \in \Sigma^*$ such that, for $x \in \{u, v, w\}$, $f_x^{\mathfrak{A}} = f_x^{\mathfrak{A}} f_{uvw}^{\mathfrak{A}}$, $f_x^{\mathfrak{A}} \neq f_{x^2}^{\mathfrak{A}}$ and $f_{uvw}^{\mathfrak{A}} = f_{x^{i_x}}^{\mathfrak{A}}$, for some $i_u, i_v, i_w \in \mathbb{N}$ that are pairwise coprime. Indeed, $\mathbf{L}(\mathfrak{A})$ is NC^1 -hard iff there is a non-solvable group in $M(\mathbf{L}(\mathfrak{A}))$. By [24, Corollary 3], a group is non-solvable iff there are 3 elements with pairwise coprime orders whose product is the identity.

In the PSPACE algorithm checking this criterion, we need to compute i_u, i_v, i_w and check that they are pairwise coprime. It is readily seen that those numbers (if exist) are $\leq |Q|^{|\mathfrak{A}|}$ and can be dealt with in PSPACE.

Using Theorems 2, 3 and Proposition 1, we obtain:

Theorem 4. Both $\text{FO}(<, \equiv_{\mathbb{N}})$ -rewritability and NC^1 -completeness (in data complexity) of $\text{LTL}_{\text{bool}}^{\square\circ}$ OMQs over Ξ -ABoxes are decidable in EXPSPACE.

4 Deciding FO-Rewritability: Lower Bounds

Theorem 5. Deciding $\text{FO}(<)$ - and $\text{FO}(<, \equiv_{\mathbb{N}})$ -rewritability of $\text{LTL}_{\text{horn}}^{\circ}$ OMAQs over Ξ -ABoxes is EXPSPACE-hard.

Proof. The idea of the proof is as follows. Given a Turing machine \mathbf{M} with exponential tape and an input word x , we construct—in a way similar to [9]—two DFAs \mathfrak{A} and \mathfrak{A}' of exponential size whose language is $\text{FO}(<)$ -definable (star-free) and, respectively, $\text{FO}(<, \equiv_{\mathbb{N}})$ -definable (in AC^0) iff \mathbf{M} rejects x . Then we simulate those DFAs by $\text{LTL}_{\text{horn}}^{\circ}$ ontologies of polynomial size.

Let x be a word and \mathbf{M} a Turing machine requiring $N = 2^{n^c}$ tape cells on an input of size n . Let N' be the first prime after $N + 1$. We construct a family $\{\mathfrak{A}_i\}_{0 \leq i < N'}$ of simple star-free minimal DFAs whose intersection represents accepting computations of \mathbf{M} on x . We encode computations as words over $\Sigma \cup \{\sharp, \flat\}$ of the form $\sharp \mathbf{c}_1 \sharp \mathbf{c}_2 \sharp \dots \sharp \mathbf{c}_{k-1} \sharp \mathbf{c}_k \flat$, where the \mathbf{c}_i are configurations.

The DFA \mathfrak{A}_0 checks that an input starts with an initial and ends with an accepting configuration of \mathbf{M} on x . The \mathfrak{A}_i , for $0 < i \leq N$, check that the i th symbol in a configuration ‘follows’ from the $(i-1)$ th, i th, and $(i+1)$ th symbols in the previous configuration (if \mathfrak{A}_1 is constructed, \mathfrak{A}_i can skip the first $i-1$ symbols and run \mathfrak{A}_1). The rest of the family just accept all the words with the

only b at the very end. We then have $\bigcap_{i=0}^{N'} \mathbf{L}(\mathfrak{A}_i) = \emptyset$ iff \mathbf{M} rejects x . We next construct minimal DFAs \mathfrak{A} and \mathfrak{A}' with the languages $(\mathbf{L}(\mathfrak{A}_0) \dots \mathbf{L}(\mathfrak{A}_{N'-1}))^*$ and $((\mathbf{L}(\mathfrak{A}_0) \cup \{b\}) \dots (\mathbf{L}(\mathfrak{A}_{N'-1}) \cup \{b\}))^*$. Thus, we obtain: (i) $\mathbf{L}(\mathfrak{A})$ is star-free iff $\bigcap_{i=0}^{N'} \mathbf{L}(\mathfrak{A}_i) = \emptyset$ (iff \mathbf{M} rejects x); (ii) $\mathbf{L}(\mathfrak{A}')$ is in AC^0 iff $\bigcap_{i=0}^{N'} \mathbf{L}(\mathfrak{A}_i) = \emptyset$.

Now we define LTL_{horn}° ontologies \mathcal{O} and \mathcal{O}' simulating \mathfrak{A} and \mathfrak{A}' . We name the states in \mathfrak{A} by triples (i, t, j) , where i indicates \mathfrak{A}_i the state ‘came from’, t is a ‘type’ of the state (say, where the DFA skips the first $i - 1$ symbols), and j is a counter in t (e.g., saying how many symbols still are to be skipped). The number of types is constant, while $i, j \leq 2^k$, for $k = \lceil \log_2 N' \rceil$.

The ontology \mathcal{O} uses the concepts A_i^j and L_j^i , where $i = 0, 1$ and $j = 1, \dots, k$, the symbols in $\Sigma \cup \{\#, b\}$, Q_t , for a type t , X, Y and F . Let $\Sigma' = \Sigma \cup \{\#, b, X, Y\}$. For any $w = w_1 \dots w_m \in (\Sigma \cup \{\#, b\})^*$, let $\mathcal{A}_w = \{X(0), w_1(1), \dots, w_m(m), Y(m+1)\}$. For a binary word $c = b_k \dots b_1$, set

$$\mathbb{A}_c = A_1^{b_1} \wedge \dots \wedge A_k^{b_k}, \quad \mathbb{A}_{<c} = \bigvee_{b_i=1} (A_i^0 \wedge \bigwedge_{j>i} A_j^{b_j}), \quad \mathbb{A}_{>c} = \bigvee_{b_i=0} (A_i^1 \wedge \bigwedge_{j>i} A_j^{b_j})$$

and let $\mathbb{L}_c, \mathbb{L}_{<c},$ and $\mathbb{L}_{>c}$ be similar concepts for L_j^i . We represent each triple (i, t, j) as the conjunction $\mathbb{A}_i \wedge Q_t \wedge \mathbb{L}_j$. We define \mathcal{O} so that, having read a prefix $w_1 \dots w_l$ of w , the DFA \mathfrak{A} is in state (i, t, j) iff $\mathcal{O}, \mathcal{A}_w \models (\mathbb{A}_i \wedge Q_t \wedge \mathbb{L}_j)(l+1)$. To achieve this, for every transition $(i_1, t_1, j_1) \rightarrow_a (i_2, t_2, j_2)$ of \mathfrak{A} , we need

$$\mathcal{O} \models \mathbb{A}_{i_1} \wedge Q_{t_1} \wedge \mathbb{L}_{j_1} \wedge a \rightarrow \circ_F \mathbb{A}_{i_2} \wedge \circ_F Q_{t_2} \wedge \circ_F \mathbb{L}_{j_2}.$$

As the structure of \mathfrak{A} is repetitive, we can ensure this without writing axioms for all transitions. For example, consider the fragment of \mathfrak{A} corresponding to the part of \mathfrak{A}_0 that, after reading x , checks that the rest of the tape is blank b . All the states in this part have the same type t with a counter j . So, for $n+1 < j < N+1$, there is a transition $(0, t, j) \rightarrow_b (0, t, j+1)$. We capture all these transitions by one formula

$$\begin{aligned} & \mathbb{A}_0 \wedge Q_0 \wedge \mathbb{L}_{>n} \wedge \mathbb{L}_{<N+1} \wedge \mathbf{b} \rightarrow \circ_F \mathbb{A}_0 \wedge \circ_F Q_0 \wedge i_{\mathbb{L}}, \quad \text{where} \\ & i_{\mathbb{L}} = \bigwedge_{l=1}^k (L_l^0 \wedge L_{l-1}^1 \wedge \dots \wedge L_1^l \rightarrow \circ_F L_l^1 \wedge \circ_F L_{l-1}^0 \wedge \dots \wedge \circ_F L_1^0) \wedge \\ & \quad \bigwedge_{l_1 < l_2} ((L_{l_1}^0 \wedge L_{l_2}^0 \rightarrow \circ_F L_{l_2}^0) \wedge (L_{l_1}^0 \wedge L_{l_2}^1 \rightarrow \circ_F L_{l_2}^1)). \end{aligned}$$

As a result, $\mathcal{O} \models (i_{\mathbb{L}} \wedge \mathbb{L}_i) \rightarrow \circ_F \mathbb{L}_{i+1}$. Similarly, we define $i_{\mathbb{A}}, d_{\mathbb{L}}, h_{\mathbb{A}},$ and $\mathbb{L}_{\mathbb{A}}$ so that $\mathcal{O} \models (d_{\mathbb{L}} \wedge \mathbb{L}_i) \rightarrow \circ_F \mathbb{L}_{i-1}$, $\mathcal{O} \models (h_{\mathbb{A}} \wedge \mathbb{A}_i) \rightarrow \circ_F \mathbb{A}_i$, $\mathcal{O} \models (\mathbb{L}_{\mathbb{A}} \wedge \mathbb{A}_i) \rightarrow \circ_F \mathbb{L}_i$. This gives us polynomially many *horn* axioms in \mathcal{O} , to which we add

$$a \wedge b \rightarrow \perp, \text{ for } a, b \in \Sigma', \quad X \rightarrow \circ_F \mathbb{A}_0 \wedge \circ_F Q_0 \wedge \circ_F \mathbb{L}_0, \quad \mathbb{A}_0 \wedge Q_0 \wedge \mathbb{L}_0 \wedge Y \rightarrow F.$$

The ontology \mathcal{O}' is defined in the same way for the DFA \mathfrak{A}' . It follows that the certain answer to (\mathcal{O}, F) over \mathcal{A}_w is yes iff $w \in \mathbf{L}(\mathfrak{A})$, and similarly for (\mathcal{O}', F) .

Lemma 1. *The LTL_{horn}° OMAQs (\mathcal{O}, F) and $(\mathcal{O}, F(x))$ are $FO(<)$ -rewritable over Σ' -ABoxes iff $\mathbf{L}(\mathfrak{A})$ is star-free; (\mathcal{O}', F) and $(\mathcal{O}', F(x))$ are $FO(<, \equiv_{\mathbb{N}})$ -rewritable over Σ' -ABoxes iff $\mathbf{L}(\mathfrak{A}')$ is in AC^0 .*

Proof. We only sketch the proof of the former claim, where (\Rightarrow) is clear. (\Leftarrow) Suppose $\mathbf{L}(\mathfrak{A})$ is star-free and \mathcal{A} is a Σ' -ABox. Then $\mathcal{O}, \mathcal{A} \models F(k)$ only if $(\mathcal{O}, \mathcal{A})$ is inconsistent, and so there are $a(i), b(i) \in \mathcal{A}$ for some $a \neq b$, or \mathcal{A} contains a subset of the form $\{X(i-1), a_1(i), a_2(i+1), a_3(i+2), \dots, a_{k-i}(k-1), Y(k)\}$ such that $a_1 a_2 \dots a_{k-i} \in \mathbf{L}(\mathfrak{A})$. As $\mathbf{L}(\mathfrak{A})$ is star-free, it is definable by an $FO(<)$ -sentence [23, Ch. VI], and so (\mathcal{O}, F) and $(\mathcal{O}, F(x))$ are $FO(<)$ -rewritable.

Theorem 6. *Deciding $FO(<)$ - or $FO(<, \equiv_{\mathbb{N}})$ -rewritability of LTL_{krom}° OMPEQs over Ξ -ABoxes is EXPSPACE-hard.*

Proof. Take an LTL_{horn}° OMAQ $\mathbf{q} = (\mathcal{O}, A)$ and $\Xi \subseteq \text{sig}(\mathbf{q})$, assuming that the axioms in \mathcal{O} are of the form $\mathbf{C} \rightarrow \perp$ or $\mathbf{C} \rightarrow B$, for some $\mathbf{C} = C_1 \wedge \dots \wedge C_n$ and atomic B . We construct an LTL_{krom}° OMPEQ $\mathbf{q}' = (\mathcal{O}', \varkappa)$ with atomic concepts $\{B, \bar{B} \mid B \in \text{sig}(\mathbf{q})\}$. Let $\mathcal{O}' = \{B \wedge \bar{B} \rightarrow \perp, \top \rightarrow B \vee \bar{B} \mid B \in \text{sig}(\mathbf{q})\}$ and

$$\varkappa = A \vee \bigvee_{\mathbf{C} \rightarrow \perp \text{ in } \mathcal{O}} \diamond_F \diamond_P \mathbf{C} \vee \bigvee_{\mathbf{C} \rightarrow B \text{ in } \mathcal{O}} \diamond_F \diamond_P (\mathbf{C} \wedge \bar{B}).$$

For any Ξ -ABox \mathcal{A} , the certain answers to \mathbf{q} and \mathbf{q}' (and to $\mathbf{q}(x)$ and $\mathbf{q}'(x)$) over \mathcal{A} coincide. It follows that \mathbf{q}' and $\mathbf{q}'(x)$ are $FO(<)$ - or $FO(<, \equiv_{\mathbb{N}})$ -rewritable over Ξ -ABoxes iff \mathbf{q} and $\mathbf{q}(x)$ are $FO(<)$ - or $FO(<, \equiv_{\mathbb{N}})$ -rewritable, respectively.

5 Linear, Krom and core OMAQs and OMPEQs

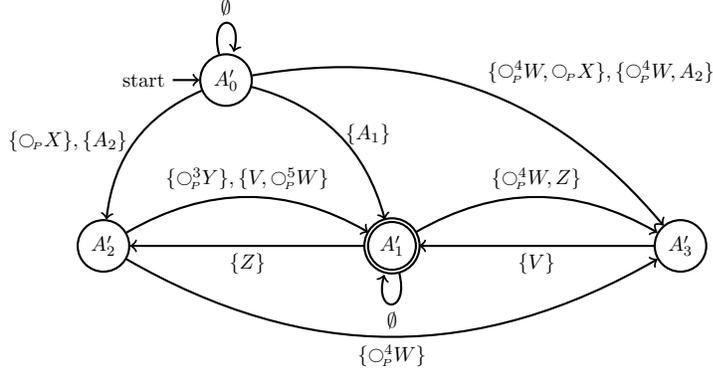
Theorem 7. *Deciding $FO(<)$ - or $FO(<, \equiv_{\mathbb{N}})$ -rewritability of linear $LTL_{horn}^{\circ P}$ OMAQs over Ξ -ABoxes is PSPACE-complete; NC^1 -completeness is in PSPACE.*

Proof. One can reduce \mathcal{L} -rewritability of linear specific OMAQs to \mathcal{L} -rewritability of linear Boolean OMAQs. Let $\mathbf{q} = (\mathcal{O}, A_1)$ be a linear OMAQ. We transform \mathbf{q} to $\mathbf{q}' = (\mathcal{O}', A'_1)$ such that \mathbf{q}' is \mathcal{L} -rewritable over Ξ -ABoxes iff \mathbf{q} is and $A \in \text{idb}(\mathcal{O}')$ only occurs in axioms of the form $\bigcirc_P^{\ell_1} C_1 \wedge \dots \wedge \bigcirc_P^{\ell_k} A_k \wedge \bigcirc_P A \rightarrow B$. For example, $\mathcal{O} = \{\bigcirc_P X \rightarrow A_2, \bigcirc_P^3 Y \wedge \bigcirc_P A_2 \rightarrow A_1, Z \wedge \bigcirc_P A_1 \rightarrow A_2, \bigcirc_P^4 W \wedge A_2 \rightarrow A_3, V \wedge \bigcirc_P A_3 \rightarrow \perp\}$ is transformed to an ontology \mathcal{O}' with the following axioms:

$$\begin{aligned} A_1 \rightarrow A'_1, A_2 \rightarrow A'_2, A_3 \rightarrow A'_3, \bigcirc_P X \rightarrow A'_2, \bigcirc_P^4 W \wedge \bigcirc_P X \rightarrow A'_3, \bigcirc_P^4 W \wedge A_2 \rightarrow A'_3, \\ \bigcirc_P^3 Y \wedge \bigcirc_P A'_2 \rightarrow A'_1, Z \wedge \bigcirc_P A'_1 \rightarrow A'_2, \bigcirc_P^4 W \wedge \bigcirc_P A'_2 \rightarrow A'_3, V \wedge \bigcirc_P A'_3 \rightarrow \perp, \\ \bigcirc_P^4 W \wedge Z \wedge \bigcirc_P A'_1 \rightarrow A'_3, V \wedge \bigcirc_P^5 W \wedge \bigcirc_P A'_2 \rightarrow \perp. \end{aligned}$$

Let $\text{idb}(\mathcal{O}) = \text{sig}(\mathbf{q}) \setminus \text{idb}(\mathcal{O})$ and let $\text{ext}(\mathcal{O})$ be the set of (maximal) basic concepts $\bigcirc_P^\ell A$ with $A \in \text{idb}(\mathcal{O})$ that occur on the left-hand side of an axiom in \mathcal{O} . Thus, $\text{ext}(\mathcal{O}') = \{\bigcirc_P X, \bigcirc_P^3 Y, Z, \bigcirc_P^4 W, V, \bigcirc_P^5 W, A_1, A_2, A_3\}$ in the example above.

Let $\text{ext}_\Xi(\mathcal{O}') = \text{ext}(\mathcal{O}') \upharpoonright \Xi$. Define an NFA $\mathfrak{B}_{\mathbf{q}'}$ over $\Gamma_{\mathbf{q}'} = 2^{\text{ext}_\Xi(\mathcal{O}')}$, which we illustrate below for the OMAQ $\mathbf{q}' = (\mathcal{O}', A'_1)$ in our example, assuming that $\Xi = \{X, Y, Z, W, V, A_1, A_2, A_3\}$ and $S \rightarrow_e S'$ implies $S \rightarrow_{e'} S'$ for all $e' \supseteq e$:



We show that $L(q')$ is \mathcal{L} -definable over Ξ -ABoxes iff $L(\mathfrak{B}_{q'})$ is \mathcal{L} -definable. The proof uses an $\text{FO}(<)$ -reduction mapping $\mathbf{a} \in \Xi^*$ to $\mathbf{e} \in \Gamma_{q'}^*$ with $\mathbf{a} \in L(q')$ iff $\mathbf{e} \in L(\mathfrak{B}_{q'})$, and the other way round. To show that deciding $\text{FO}(<)$ -, $\text{FO}(<, \equiv_{\mathbb{N}})$ -definability, or NC^1 -completeness of $L(\mathfrak{B}_{q'})$ can be done in PSPACE is not immediate as neither q' nor $\mathfrak{B}_{q'}$ is polynomial in $|q|$. However, the number of states in $\mathfrak{B}_{q'}$ is polynomial in q and one can check whether $q \rightarrow_e q'$ by a PSPACE algorithm, which allows us to use Theorems 2 and 3 for $\mathfrak{B}_{q'}$ without explicitly constructing it. The lower bounds are proved by reduction of $\text{FO}(<)$ - and $\text{FO}(<, \equiv_{\mathbb{N}})$ -definability for regular languages.

Theorem 8. *Deciding $\text{FO}(<)$ -rewritability of LTL_{krom}° OMAQs $q = (\mathcal{O}, A)$ over Ξ -ABoxes is CONP -complete.*

Proof. Let $q' = (\mathcal{O}', Y)$ with $\mathcal{O}' = \mathcal{O} \cup \{A \rightarrow \perp\}$ and fresh Y , and $q'' = (\mathcal{O}'', Y)$ with $\mathcal{O}'' = \mathcal{O} \cup \{X \wedge A \rightarrow \perp\}$ and fresh X, Y . For any $(X, Y$ -free) ABox \mathcal{A} , $(\mathcal{O}, \mathcal{A}) \models \exists x A(x)$ iff $(\mathcal{O}', \mathcal{A}) \models \exists x Y(x)$ iff \mathcal{A} is inconsistent with \mathcal{O}' ; similarly, $(\mathcal{O}, \mathcal{A}) \models A(k)$ iff $(\mathcal{O}'', \mathcal{A} \cup \{X(k)\}) \models \exists x A(x)$ iff $\mathcal{A} \cup \{X(k)\}$ is inconsistent with \mathcal{O}'' , for $k \in \text{tem}(\mathcal{A})$. So we only need to consider Boolean OMAQs $q = (\mathcal{O}, A)$ with the yes-answer only for ABoxes inconsistent with \mathcal{O} .

As \mathcal{O} is *krom*, \mathcal{A} is inconsistent with \mathcal{O} iff (i) there are $A(i), B(i) \in \mathcal{A}$ with $\mathcal{O} \models B \wedge A \rightarrow \perp$, or (ii) there exist $k_1 \leq k_2$, $B(k_1) \in \mathcal{A}$ and $C(k_2) \in \mathcal{A}$ with $\mathcal{O} \models B \rightarrow \bigcirc_F^{k_2 - k_1} \neg C$; cf. [4]. So if all $L_{BC} = \{\emptyset^n \mid \mathcal{O} \models B \rightarrow \bigcirc_F^{n+1} \neg C\}$ are $\text{FO}(<)$ -definable for any $B, C \in \Xi$, then $L_{\Xi}(q)$ is $\text{FO}(<)$ -definable and q is $\text{FO}(<)$ -rewritable over Ξ -ABoxes. For any B, C , we construct an NFA \mathfrak{A}_{BC} over the alphabet $\{\emptyset\}$ of size $O(|q|)$ that accepts L_{BC} [4]. Using [22, Theorem 6.1], we show that deciding $\text{FO}(<)$ -rewritability of the language of a unary NFA is CONP -complete, obtaining the required upper bound. To show the matching lower bound, for any unary NFA $\mathfrak{A} = (Q, \{a\}, \delta, I, F)$, we define an LTL_{core}° ontology $\mathcal{O}_{\mathfrak{A}}$ with the axioms $X \rightarrow I$, $q \wedge Y \rightarrow \perp$ and $p \rightarrow \bigcirc_F r$, for all $q \in F$ and transitions $p \xrightarrow{a} r$ in δ . The OMAQ $(\mathcal{O}_{\mathfrak{A}}, A)$ is $\text{FO}(<)$ -rewritable over $\{X, Y\}$ -ABoxes iff $L(\mathfrak{A})$ is $\text{FO}(<)$ -definable, as the answer to the OMAQ over an $\{X, Y\}$ -ABox \mathcal{A} can only be yes iff there are $X(i), Y(j) \in \mathcal{A}$ with $a^{j-i-1} \in L(\mathfrak{A})$.

Theorem 9. *Deciding $\text{FO}(<)$ -rewritability of LTL_{core}° OMPEQs $q = (\mathcal{O}, \varkappa)$ over Ξ -ABoxes is Π_2^P -complete.*

Proof. We assume that \mathcal{O} does not contain disjointness axioms $B \wedge C \rightarrow \perp$ as they can be removed from \mathcal{O} and \varkappa replaced by $\varkappa \vee \bigvee_{\mathcal{O} \models B \wedge C \rightarrow \perp} \diamond_F \diamond_P (B \wedge C)$, giving an equivalent OMQ. We further assume that all of the other rules have the following forms: $A \rightarrow B$, $A \rightarrow \circ_F B$, or $A \rightarrow \circ_P B$. As in the proof of Theorem 7, rewritability of specific OMQs can be reduced to rewritability of Boolean OMQs.

Given \mathcal{O} , \mathcal{A} , B and k , one can check in polytime whether $\mathcal{O}, \mathcal{A} \models B(k)$, which, by structural induction, implies that checking $\mathcal{O}, \mathcal{A} \models \exists x \varkappa(x)$ is in NP.

Let $\mathcal{B} = \{w_1 \dots w_k \in \Sigma_{\Xi}^* \mid \forall i |w(i)| > 0, \sum_i |w(i)| < |\varkappa|\}$. With every $w \in \mathcal{B}$ we associate the language $L_w = \mathbf{L}(\emptyset^* w_1 \emptyset^* \dots \emptyset^* w_k \emptyset^*) \cap \mathbf{L}_{\Xi}(\mathbf{q})$. For Σ_{Ξ}^* -words v and v' , we write $v' \leq v$ if they are of the same length and $v'_i \subseteq v_i$, for all i .

As \mathbf{q} is an LTL_{core}° OMPEQ, we have $\mathcal{O}, \mathcal{A} \models \mathbf{q}$ iff $\mathcal{O}, \mathcal{A}' \models \mathbf{q}$, for some $\mathcal{A}' \subseteq \mathcal{A}$, $|\mathcal{A}'| \leq |\varkappa|$. Then, for every $v \in \Sigma_{\Xi}^*$, we have $v \in \mathbf{L}_{\Xi}(\mathbf{q})$ iff there is $v' \leq v$ such that $v' \in L_w$ for some $w \in \mathcal{B}$. It follows that the language $\mathbf{L}_{\Xi}(\mathbf{q})$ is FO($<$)-definable iff L_w is FO($<$)-definable, for every $w \in \mathcal{B}$. For $w = w_1 \dots w_k \in \mathcal{B}$ and $I = (i_0, \dots, i_k)$, let $v_{w,I} = \emptyset^{i_0} w_1 \emptyset^{i_1} \dots w_k \emptyset^{i_k}$. For $c \in \mathbb{N}$, let $I_{\leq c}$ be I with all $i_j > c$ replaced with c . If L_w is FO($<$)-definable, there is c with $v_{w,I} \in L_w$ iff $v_{w,I_{\leq c}} \in L_w$. By the properties of the canonical models [4], there is a suitable c with $c < 2^{|\text{sig}(\mathbf{q})| + |\varkappa|} + 1$.

Now, \mathbf{q} is not FO($<$)-rewritable iff we can guess $w \in \mathcal{B}$ and I such that $\max(I) < 2c$ and only one of $v_{w,I}$ and $v_{w,I_{\leq c}}$ is in L_w . We can check membership in L_w using an NP-oracle, so FO($<$)-rewritability is in $\text{coNP}^{\text{NP}} = \Pi_2^p$. The matching lower bound is shown by reduction of $\forall \exists 3\text{CNF}$ [21]. Given a QBF $\forall X \exists Y \varphi$ with a 3CNF φ , $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_m\}$, we construct an LTL_{core}° OMPEQ $\mathbf{q}_{\varphi} = (\mathcal{O}_{\varphi}, \varkappa_{\varphi})$ that is FO($<$)-rewritable iff $\forall X \exists Y \varphi(X, Y)$ is true. We use atomic concepts x_i^0 and x_i^1 for $x_i \in X$, y_i^j for $y_i \in Y$ and $0 \leq j < p_i$, where p_i is the i -th prime number, A and B . \mathcal{O}_{φ} has the axioms

$$A \rightarrow y_i^0, \quad y_i^j \rightarrow \circ_F y_i^{(j+1) \bmod p_i}, \quad x_i^0 \rightarrow \circ_F x_i^0, \quad x_i^1 \rightarrow \circ_F x_i^1, \quad B \rightarrow \circ_F \circ_F B.$$

Let φ' result from φ by replacing all x_i with x_i^1 , all $\neg x_i$ with x_i^0 , and similarly for the y_j . We set $\varkappa_{\varphi} = A \wedge \bigwedge_{i=0}^n (x_i^0 \vee x_i^1) \wedge (B \vee \diamond_F \varphi')$. Suppose $\forall X \exists Y \varphi(X, Y)$ is true. Consider an ABox \mathcal{A} , for which there is t with $\mathcal{O}_{\varphi}, \mathcal{A} \models \varkappa_{\varphi}(t)$. Then $A(t) \in \mathcal{A}$ and $\mathcal{O}_{\varphi}, \mathcal{A} \models \bigwedge_{i=0}^n (x_i^0 \vee x_i^1)(t)$, and so, for every i , there is $x_i^0(s)$ or $x_i^1(s)$ in \mathcal{A} , for some $s \leq t$. There is an assignment $\mathbf{a}_1 \in 2^X$ such that $\mathcal{O}_{\varphi}, \mathcal{A} \models x_i^{\mathbf{a}_1(i)}(s)$ for all $s > t$. Take a corresponding assignment $\mathbf{a}_2 \in 2^Y$ that makes φ true. There exists a number r such that $r \bmod p_i = \mathbf{a}_2(i)$ for all $i \leq m$. Therefore $\mathcal{O}_{\varphi}, \mathcal{A} \models \varphi'(t+r)$, and so $\mathcal{O}_{\varphi}, \mathcal{A} \models \diamond_F \varphi'(t)$. Thus, the sentence

$$\exists t \left[A(t) \wedge \bigwedge_{i=0}^n \exists s \left((s \leq t) \wedge (x_i^0(s) \vee x_i^1(s)) \right) \right]$$

is a rewriting of \mathbf{q}_{φ} . If $\forall X \exists Y \varphi(X, Y)$ is false, then there is $\mathbf{a} \in 2^X$ such that φ is false for any assignment to Y . For $w = \{B\} X_{\mathbf{a}}$, $X_{\mathbf{a}} = \{A, x_1^{\mathbf{a}(0)}, \dots, x_n^{\mathbf{a}(n)}\}$, the language L_w is $\mathbf{L}(\emptyset^* \{B\} (\emptyset \emptyset)^* X_{\mathbf{a}} \emptyset^*)$, and so \mathbf{q}_{φ} cannot be FO($<$)-rewritable.

Acknowledgements This work was supported by EPSRC U.K. EP/S032282.

References

1. Artale, A., Calvanese, D., Kontchakov, R., Zakharyashev, M.: The DL-Lite family and relations. *Journal of Artificial Intelligence Research (JAIR)* 36, 1–69 (2009)
2. Artale, A., Kontchakov, R., Kovtunova, A., Ryzhikov, V., Wolter, F., Zakharyashev, M.: First-order rewritability of temporal ontology-mediated queries. In: Yang, Q., Wooldridge, M. (eds.) *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*. pp. 2706–2712. AAAI Press (2015), <http://ijcai.org/Abstract/15/383>
3. Artale, A., Kontchakov, R., Kovtunova, A., Ryzhikov, V., Wolter, F., Zakharyashev, M.: Ontology-mediated query answering over temporal data: A survey (invited talk). In: Schewe, S., Schneider, T., Wijsen, J. (eds.) *24th International Symposium on Temporal Representation and Reasoning, TIME 2017, October 16-18, 2017, Mons, Belgium*. LIPIcs, vol. 90, pp. 1:1–1:37. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2017), <https://doi.org/10.4230/LIPIcs.TIME.2017.1>
4. Artale, A., Kontchakov, R., Kovtunova, A., Ryzhikov, V., Wolter, F., Zakharyashev, M.: First-order rewritability of ontology-mediated queries in linear temporal logic. *CoRR abs/2004.07221* (2020), <https://arxiv.org/abs/2004.07221>
5. Barrington, D.A.M., Compton, K.J., Straubing, H., Thérien, D.: Regular languages in NC^1 . *J. Comput. Syst. Sci.* 44(3), 478–499 (1992), [https://doi.org/10.1016/0022-0000\(92\)90014-A](https://doi.org/10.1016/0022-0000(92)90014-A)
6. Bernátsky, L.: Regular expression star-freeness is PSPACE-complete. *Acta Cybern.* 13(1), 1–21 (1997), http://www.inf.u-szeged.hu/actacybernetica/edb/vol13n1/Bernatsky_1997_ActaCybernetica.xml
7. Bienvenu, M., ten Cate, B., Lutz, C., Wolter, F.: Ontology-based data access: A study through disjunctive datalog, CSP, and MMSNP. *ACM Transactions on Database Systems* 39(4), 33:1–44 (2014)
8. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: the *DL-Lite* family. *Journal of Automated Reasoning* 39(3), 385–429 (2007)
9. Cho, S., Huynh, D.T.: Finite-automaton aperiodicity is PSPACE-complete. *Theor. Comp. Sci.* 88(1), 99–116 (1991), <https://core.ac.uk/download/pdf/82662203.pdf>
10. Compton, K.J., Laflamme, C.: An algebra and a logic for NC^1 . *Inf. Comput.* 87(1/2), 240–262 (1990)
11. Cosmadakis, S.S., Gaifman, H., Kanellakis, P.C., Vardi, M.Y.: Decidable optimization problems for database logic programs (preliminary report). In: *STOC*. pp. 477–490 (1988)
12. Demri, S., Goranko, V., Lange, M.: *Temporal Logics in Computer Science*. Cambridge Tracts in Theoretical Computer Science, Cambridge University Press (2016)
13. Furst, M.L., Saxe, J.B., Sipser, M.: Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory* 17(1), 13–27 (1984)
14. Gabbay, D., Kurucz, A., Wolter, F., Zakharyashev, M.: *Many-Dimensional Modal Logics: Theory and Applications*, *Studies in Logic*, vol. 148. Elsevier (2003)
15. Gerasimova, O., Kikot, S., Kurucz, A., Podolskii, V., Zakharyashev, M.: A data complexity and rewritability tetrachotomy of ontology-mediated queries with a covering axiom. In: *Proceedings of KR* (2020)

16. Hernich, A., Lutz, C., Ozaki, A., Wolter, F.: Schema.org as a description logic. In: Yang, Q., Wooldridge, M.J. (eds.) Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015. pp. 3048–3054. AAAI Press (2015), <http://ijcai.org/Abstract/15/430>
17. Lutz, C., Wolter, F.: Non-uniform data complexity of query answering in description logics. In: Proc. of the 13th Int. Conference on Principles of Knowledge Representation and Reasoning (KR 2012). pp. 297–307. AAAI (2012)
18. Lutz, C., Sabellek, L.: Ontology-mediated querying with the description logic EL: trichotomy and linear datalog rewritability. In: Sierra, C. (ed.) Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017. pp. 1181–1187. ijcai.org (2017), <https://doi.org/10.24963/ijcai.2017/164>
19. Marcinkowski, J.: DATALOG sirups uniform boundedness is undecidable. In: Proceedings, 11th Annual IEEE Symposium on Logic in Computer Science, New Brunswick, New Jersey, USA, July 27-30, 1996. pp. 13–24. IEEE Computer Society (1996), <https://doi.org/10.1109/LICS.1996.561299>
20. Poggi, A., Lembo, D., Calvanese, D., De Giacomo, G., Lenzerini, M., Rosati, R.: Linking data to ontologies. *Journal on Data Semantics X*, 133–173 (2008)
21. Stockmeyer, L.J.: The polynomial-time hierarchy. *Theor. Comput. Sci.* 3(1), 1–22 (1976), [https://doi.org/10.1016/0304-3975\(76\)90061-X](https://doi.org/10.1016/0304-3975(76)90061-X)
22. Stockmeyer, L.J., Meyer, A.R.: Word problems requiring exponential time: Preliminary report. In: Aho, A.V., Borodin, A., Constable, R.L., Floyd, R.W., Harrison, M.A., Karp, R.M., Strong, H.R. (eds.) Proceedings of the 5th Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1973, Austin, Texas, USA. pp. 1–9. ACM (1973), <https://doi.org/10.1145/800125.804029>
23. Straubing, H.: *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhauser Verlag (1994)
24. Thompson, J.G.: Nonsolvable finite groups all of whose local subgroups are solvable. *Bull. Amer. Math. Soc.* 74(3), 383–437 (05 1968), <https://projecteuclid.org:443/euclid.bams/1183529612>
25. Ullman, J.D., Gelder, A.V.: Parallel complexity of logical query programs. *Algorithmica* 3, 5–42 (1988), <https://doi.org/10.1007/BF01762108>
26. Vardi, M.Y.: Decidability and undecidability results for boundedness of linear recursive queries. In: Edmondson-Yurkanan, C., Yannakakis, M. (eds.) Proceedings of the Seventh ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, March 21-23, 1988, Austin, Texas, USA. pp. 341–351. ACM (1988), <http://doi.acm.org/10.1145/308386.308470>