Temporalising *OWL 2 QL*[⋆]

A. Artale[1], R. Kontchakov[2], F. Wolter[3] and M. Zakharyaschev[2]

[1]Faculty of Computer Science, Free University of Bozen-Bolzano, Italy
[2]Department of Computer Science and Information Systems
Birkbeck, University of London, U.K.
[3]Department of Computer Science, University of Liverpool, U.K.

Abstract. We design a temporal description logic, *TQL*, that extends the standard ontology language *OWL 2 QL*, provides basic means for temporal conceptual modelling and ensures first-order rewritability of conjunctive queries for suitably defined data instances with validity time.

1 Introduction

In this paper, we investigate the possibility of extending the current W3C standard language *OWL 2 QL* for ontology-based data access (OBDA) with temporal operators in a way preserving first-order rewritability of conjunctive queries. Our ultimate aim is to understand the feasibility of OBDA for temporal data. In applications, instance data is often time-dependent: employment contracts come to an end, parliaments are elected, children are born. Temporal data can be modelled by pairs consisting of facts and their validity time; for example, $\mathit{givesBirth}(\mathit{diana}, \mathit{william}, 1982)$. To query data with validity time, it would be useful to employ an ontology that provides a conceptual model for both static and temporal aspects of the domain of interest. Thus, when querying the fact above, one could use the knowledge that, if x gives birth to y, then x becomes a mother of y from that moment on:

$$\Diamond_P\, \mathit{givesBirth} \sqsubseteq \mathit{motherOf}, \tag{1}$$

where $\Diamond_P$ reads 'sometime in the past.' *OWL 2 QL* does not support temporal conceptual modelling and, rather surprisingly, no attempt has yet been made to lift OBDA based on query rewriting to temporal ontologies and data.

Temporal extensions of DLs have been investigated since 1993; see [9, 2, 17] for surveys and [8, 6, 14, 5] for more recent developments. Moreover, temporalised *DL-Lite* logics (the logical underpinning of *OWL 2 QL*) have been constructed for temporal conceptual data modelling [3]. But unfortunately, none of the existing temporal DLs supports first-order rewritability.

The aim of this paper is to design a temporal DL that contains *OWL 2 QL*, provides basic means for temporal conceptual modelling and, at the same time,

---

⋆ This paper is an abridged version of the paper accepted for IJCAI 2013; omitted proofs can be found in [4].

ensures first-order rewritability of conjunctive queries (for suitably defined data instances with validity time). The temporal extension *TQL* of *OWL 2 QL* we present here is interpreted over sequences $\mathcal{I}(n)$, $n \in \mathbb{Z}$, of standard DL structures reflecting possible evolutions of data. TBox axioms are interpreted globally, that is, are assumed to hold in all of the $\mathcal{I}(n)$, but the concepts and roles they contain can vary in time. ABox assertions (temporal data) are time-stamped unary (for concepts) and binary (for roles) predicates that hold at the specified moments of time. Concept (role) inclusions of *TQL* generalise *OWL 2 QL* inclusions by allowing intersections of basic concepts (roles) in the left-hand side, possibly prefixed with temporal operators $\Diamond_P$ (sometime in the past) or $\Diamond_F$ (sometime in the future). Among other things, one can express in *TQL* that a concept/role name is rigid (or time-independent), persistent in the past/future or instantaneous. For example, $\Diamond_F \Diamond_P \mathit{Person} \sqsubseteq \mathit{Person}$ states that the concept *Person* is rigid, $\Diamond_P \mathit{hasName} \sqsubseteq \mathit{hasName}$ says that the role *hasName* is persistent in the future, while $\mathit{givesBirth} \sqcap \Diamond_P \mathit{givesBirth} \sqsubseteq \bot$ implies that *givesBirth* is instantaneous. Inclusions such as $\Diamond_P \mathit{Lecturer} \sqcap \Diamond_F \mathit{Lecturer} \sqsubseteq \mathit{Lecturer}$ represent convexity (or existential rigidity) of concepts or roles. However, in contrast to most existing temporal DLs, we cannot use temporal operators in the right-hand side of inclusions (e.g., to say that every student will eventually graduate: $\mathit{Student} \sqsubseteq \Diamond_F \mathit{Graduate}$).

In conjunctive queries (CQs) over *TQL* knowledge bases, we allow time-stamped predicates together with atoms of the form $(\tau < \tau')$ or $(\tau = \tau')$, where τ, τ' are temporal constants denoting integers or variables ranging over integers.

Our main result is that, given a *TQL* TBox $\mathcal{T}$ and a CQ $\boldsymbol{q}$, one can construct a union $\boldsymbol{q}'$ of CQs such that the answers to $\boldsymbol{q}$ over $\mathcal{T}$ and any temporal ABox $\mathcal{A}$ can be computed by evaluating $\boldsymbol{q}'$ over $\mathcal{A}$ extended with the temporal precedence relation $<$ between the moments of time in $\mathcal{A}$. For example, the query $\mathit{motherOf}(x, y, t)$ over (1) can be rewritten as

$$\mathit{motherOf}(x, y, t) \lor \exists t' \big((t' < t) \land \mathit{givesBirth}(x, y, t')\big).$$

Note that the addition of the transitive relation $<$ to the ABox is unavoidable: without it, there is no first-order rewriting even for the simple example above [16].

From a technical viewpoint, one of the challenges we are facing is that, in contrast to known OBDA languages with CQ rewritability (including fragments of datalog$^{\pm}$ [7]), witnesses for existential quantifiers outside the ABox are not independent from each other but interact via the temporal precedence relation. For this reason, a reduction to known languages seems to be impossible and a novel approach to rewriting has to be found. Note that straightforward temporal extensions of *TQL* lose first-order rewritability. For example, query answering over the ontology $\{\mathit{Student} \sqsubseteq \Diamond_F \mathit{Graduate}\}$ is shown to be non-tractable.

Related Work. In addition to research on temporals DLs, the Semantic Web community has developed a variety of extensions of RDF/S and OWL with validity time [19, 20, 13]. The focus of this line of research is on representing and querying time-stamped RDF triples or OWL axioms. In contrast, in our language only instance data are time-stamped, while the ontology is extended with constraints that describe how concepts and roles can change over time. In the temporal DL literature, a similar distinction has been discussed as the difference between temporalised axioms and temporalised concepts/roles; the expressive power of the respective languages is incomparable [9, 6]. To show rewritability we will use the notion of boundedness of recursion. This connection between first-order definability and boundedness is well known from the datalog and logic literature, where boundedness has been investigated extensively [10, 18, 15]. Boundedness for datalog programs on linear orders was investigated in [12]; the results are different from ours since the linear order is the only predicate symbol of the datalog programs considered and no further restrictions (comparable to ours) are imposed.

2 *TQL*: a Temporal Extension of *OWL 2 QL*

Roles S and *concepts* C of *TQL* are defined by the grammar:

$$R ::= \bot \mid P_i \mid P_i^-, \qquad S ::= R \mid S_1 \sqcap S_2 \mid \Diamond_P S \mid \Diamond_F S,$$
$$B ::= \bot \mid A_i \mid \exists R, \qquad C ::= B \mid C_1 \sqcap C_2 \mid \Diamond_P C \mid \Diamond_F C,$$

where P_i is a *role name*, A_i a *concept name* ($i \geq 0$), and $\Diamond_P$ and $\Diamond_F$ are temporal operators 'sometime in the past' and 'sometime in the future,' respectively. We call roles and concepts of the form R and B *basic*. A *TQL TBox*, $\mathcal{T}$, is a finite set of *concept* and *role inclusions* of the form $C \sqsubseteq B$, $S \sqsubseteq R$ which are assumed to hold globally (over the whole timeline). Note that the $\Diamond_{F/P}$-free fragment of *TQL* is an extension of the description logic *DL-Lite*$^{\mathcal{H}}_{horn}$ [1] with role inclusions of the form $R_1 \sqcap \cdots \sqcap R_n \sqsubseteq R$; it properly contains *OWL 2 QL* (the missing role constraints can be safely added to the language). A *TQL ABox*, $\mathcal{A}$, is a (finite) set of atoms $P_i(a, b, n)$ and $A_i(a, n)$, where a, b are *individual constants* and $n \in \mathbb{Z}$ a *temporal constant*. The set of individual constants in $\mathcal{A}$ is denoted by $\mathsf{ind}(\mathcal{A})$, and the set of temporal constants by $\mathsf{tem}(\mathcal{A})$. A *TQL knowledge base* (*KB*) is a pair $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, where $\mathcal{T}$ is a TBox and $\mathcal{A}$ an ABox.

A *temporal interpretation*, $\mathcal{I}$, is given by the ordered set $(\mathbb{Z}, <)$ of *time points* and standard (atemporal) interpretations $\mathcal{I}(n) = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}(n)})$, for each $n \in \mathbb{Z}$. Thus, $\Delta^{\mathcal{I}} \neq \emptyset$ is the common domain of all $\mathcal{I}(n)$, $a_i^{\mathcal{I}(n)} \in \Delta^{\mathcal{I}}$, $A_i^{\mathcal{I}(n)} \subseteq \Delta^{\mathcal{I}}$ and $P_i^{\mathcal{I}(n)} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. We assume that $a_i^{\mathcal{I}(n)} = a_i^{\mathcal{I}(0)}$, for all $n \in \mathbb{Z}$. To simplify presentation, we adopt the *unique name assumption*, that is, $a_i^{\mathcal{I}(n)} \neq a_j^{\mathcal{I}(n)}$ for $i \neq j$ (although the obtained results hold without it as the language has no number restrictions). The temporal constructs are interpreted in $\mathcal{I}$ as follows, where $n \in \mathbb{Z}$:

$$(\Diamond_P C)^{\mathcal{I}(n)} = \{x \mid x \in C^{\mathcal{I}(m)}, \text{ for some } m < n\},$$
$$(\Diamond_F C)^{\mathcal{I}(n)} = \{x \mid x \in C^{\mathcal{I}(m)}, \text{ for some } m > n\},$$
$$(\Diamond_P S)^{\mathcal{I}(n)} = \{(x, y) \mid (x, y) \in S^{\mathcal{I}(m)}, \text{ for some } m < n\},$$
$$(\Diamond_F S)^{\mathcal{I}(n)} = \{(x, y) \mid (x, y) \in S^{\mathcal{I}(m)}, \text{ for some } m > n\}.$$

The *satisfaction relation* $\models$ is defined as usual. If all inclusions in $\mathcal{T}$ and atoms in $\mathcal{A}$ are satisfied in $\mathcal{I}$, we call $\mathcal{I}$ a *model* of $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and write $\mathcal{I} \models \mathcal{K}$.

A *conjunctive query* (*CQ*) is a (two-sorted) first-order formula $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s}) = \exists \boldsymbol{y}, \boldsymbol{t}\, \varphi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}, \boldsymbol{t})$, where $\varphi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}, \boldsymbol{t})$ is a conjunction of atoms of the form $A_i(\xi, \tau)$, $P_i(\xi, \zeta, \tau)$, $(\tau = \sigma)$ and $(\tau < \sigma)$, with ξ, ζ being *individual terms*—individual constants or variables in $\boldsymbol{x}$, $\boldsymbol{y}$—and τ, σ *temporal terms*—temporal constants or variables in $\boldsymbol{t}$, $\boldsymbol{s}$. In a *positive existential query* (PEQ) $\boldsymbol{q}$, the formula φ can also contain $\vee$. A *union of CQs* (UCQ) is a disjunction of CQs (so every PEQ is equivalent to an exponentially larger UCQ).

Given a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and a CQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$, we call tuples $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathsf{tem}(\mathcal{A})$ a *certain answer* to $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ over $\mathcal{K}$ and write $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$, if $\mathcal{I} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$ for any model $\mathcal{I}$ of $\mathcal{K}$ (understood as a two-sorted first-order model).

Example 1. Suppose Bob was a lecturer at UCL between times n_1 and n_2, after which he was appointed professor on a permanent contract. To model this situation, we use individual names, e_1 and e_2, to represent the two events of Bob's employment. The ABox will contain $n_1 < n_2$ and the atoms $\mathit{lect}(\mathit{bob}, e_1, n_1)$, $\mathit{lect}(\mathit{bob}, e_1, n_2)$, $\mathit{prof}(\mathit{bob}, e_2, n_2 + 1)$. In the TBox, we make sure that everybody is holding the corresponding post over the duration of the contract, and include other knowledge about the university life:

$$
\begin{array}{ll}
\Diamond_P \mathit{lect} \sqcap \Diamond_F \mathit{lect} \sqsubseteq \mathit{lect}, & \Diamond_P \mathit{prof} \sqsubseteq \mathit{prof}, \\
\exists \mathit{lect} \sqsubseteq \mathit{Lecturer}, & \exists \mathit{prof} \sqsubseteq \mathit{Professor}, \\
\mathit{Professor} \sqsubseteq \exists \mathit{supervisesPhD}, & \mathit{Professor} \sqsubseteq \mathit{Staff}, \\
\Diamond_P \mathit{supervisesPhD} \sqcap \Diamond_F \mathit{supervisesPhD} \sqsubseteq \mathit{supervisesPhD}. &
\end{array}
$$

We can now obtain staff who supervised PhDs between times k_1 and k_2 by posing the following CQ: $\exists y, t\, \big((k_1 < t < k_2) \wedge \mathit{Staff}(x, t) \wedge \mathit{supervisesPhD}(x, y, t)\big)$.

The key idea of OBDA is to reduce answering CQs over KBs to evaluating FO-queries over relational databases. To obtain such a reduction for *TQL* KBs, we employ a very basic type of temporal databases. With every *TQL* ABox $\mathcal{A}$, we associate a data instance $[\mathcal{A}]$ that contains all atoms from $\mathcal{A}$ as well as the atoms $(n_1 < n_2)$ such that $n_1, n_2 \in \mathbb{Z}$ with $\min \mathsf{tem}(\mathcal{A}) \leq n_1, n_2 \leq \max \mathsf{tem}(\mathcal{A})$ and $n_1 < n_2$. Thus, in addition to $\mathcal{A}$, we explicitly include in $[\mathcal{A}]$ the temporal precedence relation over the *convex closure* of the time points that occur in $\mathcal{A}$. (Note that, in standard temporal databases, the order over timestamps is built-in.) The main result of this paper is the following:

Theorem 1. *Suppose $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ is a CQ and $\mathcal{T}$ a TQL TBox. Then one can construct a UCQ $\boldsymbol{q}'(\boldsymbol{x}, \boldsymbol{s})$ such that, for any consistent KB $(\mathcal{T}, \mathcal{A})$ such that $\mathcal{A}$ contains all temporal constants from $\boldsymbol{q}$, any $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and any $\boldsymbol{n} \subseteq \mathsf{tem}(\mathcal{A})$, we have $(\mathcal{T}, \mathcal{A}) \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$ iff $[\mathcal{A}] \models \boldsymbol{q}'(\boldsymbol{a}, \boldsymbol{n})$.*

Such a UCQ $\boldsymbol{q}'(\boldsymbol{x}, \boldsymbol{s})$ is called a *rewriting* for $\boldsymbol{q}$ and $\mathcal{T}$. Note that consistency checking can easily be reduced to CQ-answering. Indeed, let F be a fresh role name. Denote by $\mathcal{T}^{\perp}$ the result of replacing $\perp$ with F in all role inclusions of $\mathcal{T}$

and with $\exists F$ in all concept inclusions. Clearly, $(\mathcal{T}^\perp, \mathcal{A})$ is consistent for any ABox $\mathcal{A}$, and $(\mathcal{T}, \mathcal{A})$ is inconsistent iff $(\mathcal{T}^\perp, \mathcal{A}) \models \boldsymbol{q}^\perp$, where $\boldsymbol{q}^\perp = \exists x, y, t\, F(x, y, t)$.

For an ABox $\mathcal{A}$, we denote by $\mathcal{A}^\mathbb{Z}$ the *infinite* data instance which contains the atoms in $\mathcal{A}$ as well as all $(n_1 < n_2)$ such that $n_1, n_2 \in \mathbb{Z}$ and $n_1 < n_2$. It will be convenient to regard CQs $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ as *sets* of atoms, so that we can write, e.g., $A(\xi, \tau) \in \boldsymbol{q}$. We say that $\boldsymbol{q}$ is *totally ordered* if, for any temporal terms τ, τ' in $\boldsymbol{q}$, at least one of the constraints $\tau < \tau'$, $\tau = \tau'$ or $\tau' < \tau$ is in $\boldsymbol{q}$ and the set of such constraints is consistent (in the sense that it can be satisfied in $\mathbb{Z}$). Every CQ is equivalent to a union of totally ordered CQs (the empty union is $\perp$).

Lemma 1. *For every UCQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$, one can compute a UCQ $\boldsymbol{q}'(\boldsymbol{x}, \boldsymbol{s})$ such that, for any ABox $\mathcal{A}$ containing all temporal constants from $\boldsymbol{q}$, any $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathsf{tem}(\mathcal{A})$, we have $\mathcal{A}^\mathbb{Z} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$ iff $[\mathcal{A}] \models \boldsymbol{q}'(\boldsymbol{a}, \boldsymbol{n})$.*

Example 2. Let $\mathcal{T} = \{\Diamond_{\!F} C \sqsubseteq A,\ \Diamond_{\!P} A \sqsubseteq B\}$ and $\boldsymbol{q}(x, s) = B(x, s)$. Then, for any $\mathcal{A}$, $a \in \mathsf{ind}(\mathcal{A})$, $n \in \mathsf{tem}(\mathcal{A})$, we have $(\mathcal{T}, \mathcal{A}) \models \boldsymbol{q}(a, n)$ iff $\mathcal{A}^\mathbb{Z} \models \boldsymbol{q}'(a, n)$, where $\boldsymbol{q}'(x, s) = B(x, s) \lor \exists t\, \big((t < s) \land A(x, t)\big) \lor \exists t, r\, \big((t < s) \land (t < r) \land C(x, r)\big)$. Note, however, that $\boldsymbol{q}'$ is *not* a rewriting for $\boldsymbol{q}$ and $\mathcal{T}$. Take, for example, $\mathcal{A} = \{C(a, 0)\}$. Then $(\mathcal{T}, \mathcal{A}) \models B(a, 0)$ but $[\mathcal{A}] \not\models \boldsymbol{q}'(a, 0)$. A correct rewriting is obtained by replacing the last disjunct in $\boldsymbol{q}'$ with $\exists r\, C(x, r)$; it can be computed by applying Lemma 1 to $\boldsymbol{q}'$ and slightly simplifying the result.

In view of Lemma 1, from now on we will only focus on rewritings over $\mathcal{A}^\mathbb{Z}$.

The problem of finding rewritings for CQs and *TQL* TBoxes can be reduced to the case where the TBoxes only contain inclusions of the following form:

$$B_1 \sqcap B_2 \sqsubseteq B,\ \Diamond_{\!F} B_1 \sqsubseteq B_2,\ \Diamond_{\!P} B_1 \sqsubseteq B_2,\ R_1 \sqcap R_2 \sqsubseteq R,\ \Diamond_{\!F} R_1 \sqsubseteq R_2,\ \Diamond_{\!P} R_1 \sqsubseteq R_2.$$

We say that such TBoxes are in *normal form.*

Theorem 2. *For every TQL TBox $\mathcal{T}$, one can construct in polynomial time a TQL TBox $\mathcal{T}'$ in normal form (possibly containing additional concept and role names) such that $\mathcal{T}' \models \mathcal{T}$ and, for every model $\mathcal{I}$ of $\mathcal{T}$, there exists a model of $\mathcal{T}'$ that coincides with $\mathcal{I}$ on all concept and role names in $\mathcal{T}$.*

Suppose now that we have a UCQ rewriting $\boldsymbol{q}'$ for a CQ $\boldsymbol{q}$ and the TBox $\mathcal{T}'$ in Theorem 2. We obtain a rewriting for $\boldsymbol{q}$ and $\mathcal{T}$ simply by removing from $\boldsymbol{q}'$ those CQs that contain symbols occurring in $\mathcal{T}'$ but not in $\mathcal{T}$. From now on, we assume that *all TQL TBoxes are in normal form.* The set of role names in $\mathcal{T}$ and with their inverses is denoted by $\mathsf{R}_\mathcal{T}$, while $|\mathcal{T}|$ is the number of concept and role names in $\mathcal{T}$.

We begin the construction of rewritings by considering the case when all concept inclusions are of the form $C \sqsubseteq A_i$, so existential quantification $\exists R$ does not occur in the right-hand side. *TQL* TBoxes of this form will be called *flat.* Note that RDFS statements can be expressed by means of flat TBoxes.

3 UCQ Rewriting for Flat TBoxes

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB with a flat TBox $\mathcal{T}$ (in normal form). Our first aim is to construct a model $\mathcal{C}_\mathcal{K}$ of $\mathcal{K}$, called the *canonical model*, for which the following theorem holds:

Theorem 3. *For any consistent $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ with flat $\mathcal{T}$ and any CQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$, we have $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$ iff $\mathcal{C}_\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$, for all tuples $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathbb{Z}$.*

The construction uses a closure operator, cl, which applies the rules **(ex)**, **(c1)**–**(c3)**, **(r1)**–**(r3)** below to a set, $\mathcal{S}$, of atoms of the form $R(u, v, n)$, $A(u, n)$, $\exists R(u, n)$ or $(n < n')$:

(ex) If $R(u, v, n) \in \mathcal{S}$ then add $\exists R(u, n)$, $\exists R^-(v, n)$ to $\mathcal{S}$;
(c1) if $(B_1 \sqcap B_2 \sqsubseteq B) \in \mathcal{T}$ and $B_1(u, n)$, $B_2(u, n) \in \mathcal{S}$, then add $B(u, n)$ to $\mathcal{S}$;
(c2) if $(\Diamond_P B \sqsubseteq B') \in \mathcal{T}$, $B(u, m) \in \mathcal{S}$ for some $m < n$ and n occurs in $\mathcal{S}$, then add $B'(u, n)$ to $\mathcal{S}$;
(c3) if $(\Diamond_F B \sqsubseteq B') \in \mathcal{T}$, $B(u, m) \in \mathcal{S}$ for some $m > n$ and n occurs in $\mathcal{S}$, then add $B'(u, n)$ to $\mathcal{S}$;
(r1) if $(R_1 \sqcap R_2 \sqsubseteq R) \in \mathcal{T}$ and $R_1(u, v, n), R_2(u, v, n) \in \mathcal{S}$, then add $R(u, v, n)$ to $\mathcal{S}$;
(r2) if $(\Diamond_P R \sqsubseteq R') \in \mathcal{T}$, $R(u, v, m) \in \mathcal{S}$ for some $m < n$ and n occurs in $\mathcal{S}$, then add $R'(u, v, n)$ to $\mathcal{S}$;
(r3) if $(\Diamond_F R \sqsubseteq R') \in \mathcal{T}$, $R(u, v, m) \in \mathcal{S}$ for some $m > n$ and n occurs in $\mathcal{S}$, then add $R'(u, v, n)$ to $\mathcal{S}$.

We set

$$\mathsf{cl}^0(\mathcal{S}) = \mathcal{S}, \qquad \mathsf{cl}^{i+1}(\mathcal{S}) = \mathsf{cl}(\mathsf{cl}^i(\mathcal{S})), \qquad \mathsf{cl}^\infty(\mathcal{S}) = \bigcup\nolimits_{i \geq 0} \mathsf{cl}^i(\mathcal{S}).$$

Note first that $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is inconsistent iff $\bot \in \mathsf{cl}^\infty(\mathcal{A}^\mathbb{Z})$. If $\mathcal{K}$ is consistent, we define the *canonical model* $\mathcal{C}_\mathcal{K}$ of $\mathcal{K}$ by taking $\Delta^{\mathcal{C}_\mathcal{K}} = \mathsf{ind}(\mathcal{A})$, $a \in A^{\mathcal{C}_\mathcal{K}(n)}$ iff $A(a, n) \in \mathsf{cl}^\infty(\mathcal{A}^\mathbb{Z})$, and $(a, b) \in P^{\mathcal{C}_\mathcal{K}(n)}$ iff $P(a, b, n) \in \mathsf{cl}^\infty(\mathcal{A}^\mathbb{Z})$, for $n \in \mathbb{Z}$. (As $\mathcal{T}$ is flat, atoms of the form $\exists R(u, n)$ can only be added by (**ex**).) This gives us Theorem 3. The following lemma shows that to construct $\mathcal{C}_\mathcal{K}$ we actually need only a bounded number of applications of cl that does not depend on $\mathcal{A}$:

Lemma 2. *Let $\mathcal{T}$ be a flat TBox and $n_\mathcal{T} = (4 \cdot |\mathcal{T}|)^4$. Then $\mathsf{cl}^\infty(\mathcal{A}^\mathbb{Z}) = \mathsf{cl}^{n_\mathcal{T}}(\mathcal{A}^\mathbb{Z})$, for any ABox $\mathcal{A}$.*

We now use Lemma 2 to construct a rewriting for any flat TBox $\mathcal{T}$ and CQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$. For a concept C and a role S, denote by $C^\sharp$ and $S^\sharp$ their standard FO-translations: for example, $(\Diamond_F A)^\sharp(\xi, \tau) = \exists t\, ((\tau < t) \land A(\xi, t))$ and $(\exists R)^\sharp(\xi, \tau) = \exists y\, R(\xi, y, \tau)$. Now, given a PEQ φ, we set $\varphi^{0\downarrow} = \varphi$ and define, inductively, $\varphi^{(n+1)\downarrow}$ as the result of replacing every

- $A(\xi, \tau)$ with $A(\xi, \tau) \lor \bigvee_{(C \sqsubseteq A) \in \mathcal{T}} (C^\sharp(\xi, \tau))^{n\downarrow}$,
- $P(\xi, \zeta, \tau)$ with $P(\xi, \zeta, \tau) \lor \bigvee_{(S \sqsubseteq P) \in \mathcal{T}} (S^\sharp(\xi, \zeta, \tau))^{n\downarrow}$.

Finally, we set: $\mathsf{ext}^{\mathcal{T}}_{\boldsymbol{q}}(\boldsymbol{x},\boldsymbol{s}) = (\boldsymbol{q}(\boldsymbol{x},\boldsymbol{s}))^{n_{\mathcal{T}}\downarrow}$. Clearly, $\mathsf{ext}^{\mathcal{T}}_{\boldsymbol{q}}(\boldsymbol{x},\boldsymbol{s})$ is a PEQ, and so can be equivalently transformed into a UCQ. By Theorem 3, Lemma 2 and Lemma 1, we obtain a rewriting for $\boldsymbol{q}$ and $\mathcal{T}$:

Theorem 4. *Let $\mathcal{T}$ be a flat TBox and $\boldsymbol{q}(\boldsymbol{x},\boldsymbol{s})$ a CQ. Then, for any consistent KB $(\mathcal{T},\mathcal{A})$, any $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathbb{Z}$, $(\mathcal{T},\mathcal{A}) \models \boldsymbol{q}(\boldsymbol{a},\boldsymbol{n})$ iff $\mathcal{A}^{\mathbb{Z}} \models \mathsf{ext}^{\mathcal{T}}_{\boldsymbol{q}}(\boldsymbol{a},\boldsymbol{n})$.*

4 Canonical Models for Arbitrary TBoxes

Canonical models for consistent KBs $\mathcal{K} = (\mathcal{T},\mathcal{A})$ with not necessarily flat TBoxes $\mathcal{T}$ (in normal form) can be constructed starting from $\mathcal{A}^{\mathbb{Z}}$ and using the rules given in the previous section together with the following one:

($\rightsquigarrow$) if $\exists R(u,n) \in \mathcal{S}$ and $R(u,v,n) \notin \mathcal{S}$ for any v, then add $R(u,v,n)$ to $\mathcal{S}$, for some fresh individual name v; in this case we write $u \rightsquigarrow^n_R v$.

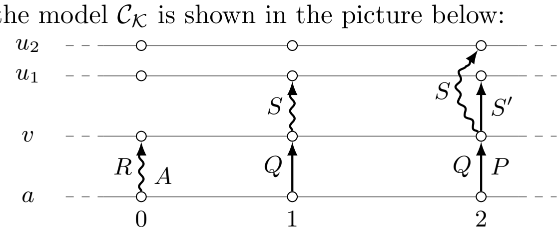
Denote by cl_1 the closure operator under the resulting 8 rules. Again, $\mathcal{K}$ is inconsistent iff $\bot \in \mathsf{cl}^{\infty}_1(\mathcal{A}^{\mathbb{Z}})$. If $\mathcal{K}$ is consistent, we define the *canonical model* $\mathcal{C}_{\mathcal{K}}$ for $\mathcal{K}$ by the set $\mathsf{cl}^{\infty}_1(\mathcal{A}^{\mathbb{Z}})$ in the same way as in Section 3 but taking the domain $\Delta^{\mathcal{C}_{\mathcal{K}}}$ to contain all the individual names in $\mathsf{cl}^{\infty}_1(\mathcal{A}^{\mathbb{Z}})$.

Theorem 5. *For every consistent $\mathcal{K} = (\mathcal{T},\mathcal{A})$ and every CQ $\boldsymbol{q}(\boldsymbol{x},\boldsymbol{s})$, we have $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a},\boldsymbol{n})$ iff $\mathcal{C}_{\mathcal{K}} \models \boldsymbol{q}(\boldsymbol{a},\boldsymbol{n})$, for any tuples $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathbb{Z}$.*

Example 3. Let $\mathcal{K} = (\mathcal{T},\mathcal{A})$ with $\mathcal{A} = \{A(a,0)\}$ and

$$\mathcal{T} \quad = \quad \{\ A \sqsubseteq \exists R,\quad \Diamond_P R \sqsubseteq Q,\quad \exists Q^- \sqsubseteq \exists S,\quad \Diamond_P Q \sqsubseteq P,\quad \Diamond_P S \sqsubseteq S'\ \}.$$

A fragment of the model $\mathcal{C}_{\mathcal{K}}$ is shown in the picture below:



We say that the individuals $a \in \mathsf{ind}(\mathcal{A})$ are of *depth* 0 in $\mathcal{C}_{\mathcal{K}}$; now, if u is of depth d in $\mathcal{C}_{\mathcal{K}}$ and $u \rightsquigarrow^n_R v$, for some $n \in \mathbb{Z}$ and R, then v is of *depth* $d+1$ in $\mathcal{C}_{\mathcal{K}}$. Thus, both u_1 and u_2 in Example 3 are of depth 2 and v is of depth 1. The restriction of $\mathcal{C}_{\mathcal{K}}$, treated as a set of atoms, to the individual names of depth $\leq d$ is denoted by $\mathcal{C}^d_{\mathcal{K}}$. Note that this set is not necessarily closed under the rule ($\rightsquigarrow$).

In the remainder of this section, we describe the structure of $\mathcal{C}_{\mathcal{K}}$, which is required for the rewriting in the next section. We split $\mathcal{C}_{\mathcal{K}}$ into two parts: one consists of the elements in $\mathsf{ind}(\mathcal{A})$, while the other contains the fresh individuals introduced by ($\rightsquigarrow$). As this rule always uses *fresh* individuals, to understand the structure of the latter part it is enough to consider KBs of the form $\mathcal{K}_{\mathcal{T},R} = (\mathcal{T} \cup \{A \sqsubseteq \exists R\}, \{A(a,0)\})$ with fresh A. We begin by analysing the behaviour of the atoms $R'(a,u,n)$ entailed by $R(a,u,0)$, where $a \rightsquigarrow^0_R u$.

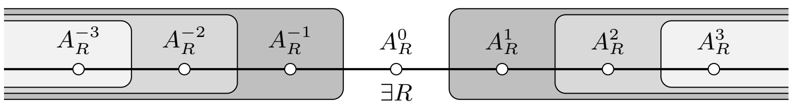
Lemma 3. *Let* $a \rightsquigarrow^0_R u$ *in* $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$. *If either* $m < n < 0$ *or* $0 < n < m$*, then* $R'(a,u,n) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ *implies* $R'(a,u,m) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$*; moreover, if* $n < m = -|\mathsf{R}_\mathcal{T}|$ *or* $|\mathsf{R}_\mathcal{T}| = m < n$*, then* $R'(a,u,m) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ *iff* $R'(a,u,n) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$.

The atoms $R'(a,u,n)$ entailed by $R(a,u,0)$ in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ via **(r1)**–**(r3)**, also have an impact, via **(ex)**, on the atoms of the form $B(a,n)$ and $B(u,n)$ in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$. Thus, in Example 3, $R(a,v,0)$ entails $\exists Q(a,n)$, for $n > 0$. To analyse the behaviour of such atoms, it is helpful to assume that $\mathcal{T}$ is in *concept normal form* (CoNF) in the following sense: for every role $R \in \mathsf{R}_\mathcal{T}$, the TBox $\mathcal{T}$ contains

$$\exists R \sqsubseteq A^0_R,\ \Diamond_F \exists R \sqsubseteq A^{-1}_R,\ \Diamond_F A^{-m}_R \sqsubseteq A^{-m-1}_R,\ \Diamond_P \exists R \sqsubseteq A^1_R,\ \Diamond_P A^m_R \sqsubseteq A^{m+1}_R,$$

for $0 \leq m \leq |\mathsf{R}_\mathcal{T}|$ and some concepts A^i_R, and

$$A^m_R \sqsubseteq \exists R', \qquad \text{for all } |m| \leq |\mathsf{R}_\mathcal{T}| \text{ and } R'(a,v,m) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}.$$



(In Example 3, $\mathcal{C}_\mathcal{K}$ will contain the atoms $A^1_R(a,n)$ and $A^2_R(a,n+1)$, for $n \geq 1$.) By Lemma 3, if $\mathcal{T}$ is in CoNF, then we can compute the atoms $B(a,n)$ and $B(u,n)$ in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ without using the rules **(r1)**–**(r3)**. Lemma 3 also implies that we can add the inclusions above (with fresh A^i_R) to $\mathcal{T}$ if required, thereby obtaining a conservative extension of $\mathcal{T}$; so from now on we always assume $\mathcal{T}$ *to be in CoNF*. These observations enable the proof of the following two lemmas. The first one characterises the atoms $B(u,n)$ in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$:

Lemma 4. *Let* $a \rightsquigarrow^0_R u$ *in* $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$. *If either* $m < n < 0$ *or* $0 < n < m$*, then* $B(u,n) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ *implies* $B(u,m) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$*; moreover, if either* $n < m = -|\mathcal{T}|$ *or* $|\mathcal{T}| = m < n$*, then* $B(u,m) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ *iff* $B(u,n) \in \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$.

The second lemma characterises the ABox part of $\mathcal{C}_\mathcal{K}$ and is a straightforward generalisation of Lemma 2:

Lemma 5. *For any KB* $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ *and any atom* α *of the form* $A(a,n)$, $\exists R(a,n)$ *or* $R(a,b,n)$*, where* $a, b \in \mathsf{ind}(\mathcal{A})$ *and* $n \in \mathbb{Z}$*, we have* $\alpha \in \mathcal{C}_\mathcal{K}$ *iff* $\alpha \in \mathsf{cl}^{n_\mathcal{T}}(\mathcal{A}^\mathbb{Z})$.

An obvious extension of the rewriting of Theorem 4 provides, for every CQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$, a UCQ $\mathsf{ext}^\mathcal{T}_{\boldsymbol{q}}(\boldsymbol{x}, \boldsymbol{s})$ of the appropriate length such that

$$\mathcal{C}^0_\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n}) \quad \text{iff} \quad \mathcal{A}^\mathbb{Z} \models \mathsf{ext}^\mathcal{T}_{\boldsymbol{q}}(\boldsymbol{a}, \boldsymbol{n}), \qquad \text{for all } \boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A}), \boldsymbol{n} \subseteq \mathbb{Z}. \tag{2}$$

In particular, for every basic concept of the form $\exists R$, we have $\exists R(a,n) \in \mathcal{C}_\mathcal{K}$ iff $\mathcal{A}^\mathbb{Z} \models \mathsf{ext}^\mathcal{T}_{\exists R}(a,n)$, for all $a \in \mathsf{ind}(\mathcal{A})$ and $n \in \mathbb{Z}$.

We now use the obtained results to show that one can find all answers to a CQ $\boldsymbol{q}$ over a *TQL* KB $\mathcal{K}$ by only considering a fragment of $\mathcal{C}_\mathcal{K}$ whose size is polynomial in $|\mathcal{T}|$ and $|\boldsymbol{q}|$. This property is called the *polynomial witness* *property* [11]. Denote by $\mathcal{C}_{\mathcal{K}}^{d,\ell}$, for $d, \ell \geq 0$, the restriction of $\mathcal{C}_{\mathcal{K}}^{d}$ to the moments of time in the interval $[\min \mathsf{tem}(\mathcal{A}) - \ell, \max \mathsf{tem}(\mathcal{A}) + \ell]$.

Let $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ be a CQ. Tuples $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathsf{tem}(\mathcal{A})$ give a certain answer to $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ over $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ iff there exists a *homomorphism* h from $\boldsymbol{q}$ to $\mathcal{C}_{\mathcal{K}}$, which maps individual (temporal) terms of $\boldsymbol{q}$ to individual (respectively, temporal) terms of $\mathcal{C}_{\mathcal{K}}$ in such a way that the following conditions hold: $h(\boldsymbol{x}) = \boldsymbol{a}$ and $h(\boldsymbol{s}) = \boldsymbol{n}$; $h(b) = b$ and $h(m) = m$, for any individual and temporal constants b and m; and $h(\boldsymbol{q}) \subseteq \mathcal{C}_{\mathcal{K}}$, where $h(\boldsymbol{q})$ is the set of atoms obtained by replacing every term in $\boldsymbol{q}$ with its h-image, e.g., $P(\xi, \zeta, \tau)$ with $P(h(\xi), h(\zeta), h(\tau))$ and $(\tau_1 < \tau_2)$ with $h(\tau_1) < h(\tau_2)$. Now, using the monotonicity lemmas for the temporal dimension and the fact that the atoms of depth $> |\mathsf{R}_{\mathcal{T}}|$ in the canonical models duplicate atoms of smaller depth, we obtain

Theorem 6. *There are polynomials f_1 and f_2 such that, for any consistent TQL KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, any CQ $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ and any $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathsf{tem}(\mathcal{A})$, we have $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a}, \boldsymbol{n})$ iff there is a homomorphism $h \colon \boldsymbol{q} \to \mathcal{C}_{\mathcal{K}}$ such that $h(\boldsymbol{q}) \subseteq \mathcal{C}_{\mathcal{K}}^{d,\ell}$, where $d = f_1(|\mathcal{T}|, |\boldsymbol{q}|)$ and $\ell = f_2(|\mathcal{T}|, |\boldsymbol{q}|)$.*

5 UCQ Rewriting

We now define a rewriting for any given CQ and *TQL* TBox. Suppose $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{s})$ is a CQ and $\mathcal{T}$ a *TQL* TBox (in CoNF). Without loss of generality we assume $\boldsymbol{q}$ to be totally ordered. By a *sub-query* of $\boldsymbol{q}$ we understand any subset $\boldsymbol{q}' \subseteq \boldsymbol{q}$ containing all temporal constraints $(\tau < \tau')$ and $(\tau = \tau')$ that occur in $\boldsymbol{q}$. In the rewriting for $\boldsymbol{q}$ and $\mathcal{T}$ given below, we consider all possible splittings of $\boldsymbol{q}$ into two sub-queries (sharing the same temporal terms). One is to be mapped to the ABox part of the canonical model $\mathcal{C}_{(\mathcal{T},\mathcal{A})}$, and so we can rewrite it using (2). The other sub-query is to be mapped to the non-ABox part of $\mathcal{C}_{(\mathcal{T},\mathcal{A})}$ and requires a different rewriting.

For every $R \in \mathsf{R}_{\mathcal{T}}$, we construct the set $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}^{d,\ell}$, where d and ℓ are provided by Theorem 6. Let h be a map from a sub-query $\boldsymbol{q}_h \subseteq \boldsymbol{q}$ to $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}^{d,\ell}$ such that $h(\boldsymbol{q}_h) \subseteq \mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}^{d,\ell}$. Denote by $\mathcal{X}_h$ the set of individual terms ξ in $\boldsymbol{q}_h$ with $h(\xi) = a$, and let $\mathcal{Y}_h$ be the remaining set of individual terms in $\boldsymbol{q}_h$. We call h a *witness for R* if

- $\mathcal{X}_h$ contains at most one individual constant;
- every term in $\mathcal{Y}_h$ is an existentially quantified variable in $\boldsymbol{q}$;
- $\boldsymbol{q}_h$ contains all atoms in $\boldsymbol{q}$ with a variable from $\mathcal{Y}_h$.

Let h be a witness for R. Denote by $\leadsto$ the union of all $\leadsto_{R'}^{n}$ in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}^{d,\ell}$. Clearly, $\leadsto$ is a tree order on the individuals in $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}^{d,\ell}$, with root a. Let T_h be its minimal sub-tree containing a and the h-images of all the individual terms in $\boldsymbol{q}_h$. For each $v \in T_h \setminus \{a\}$, we take the (unique) moment $\mathfrak{g}(v)$ with $u \leadsto_{R}^{\mathfrak{g}(v)} v$, for some u and R, and set $\mathfrak{g}(a) = 0$. For $A(y, \tau) \in \boldsymbol{q}_h$, we say that $h(y)$ *realises* $A(y, \tau)$. For any $P(\xi, \xi', \tau) \in \boldsymbol{q}_h$, there are $u, u' \in T_h$ with $u \leadsto u'$ and $\{u, u'\} = \{h(\xi), h(\xi')\}$; we say that u' *realises* $P(\xi, \xi', \tau)$. Let $\boldsymbol{r}$ be a list of fresh temporal variables r_u, for

$u \in T_h \setminus \{a\}$. Consider the following formula, whose free variables are r_a and the temporal variables of $\boldsymbol{q}_h$:

$$\mathsf{t}_h = \exists \boldsymbol{r} \,\big(\bigwedge_{u \rightsquigarrow v} \delta^{\mathfrak{g}(v)-\mathfrak{g}(u)}(r_u, r_v) \wedge \bigwedge_{u \text{ realises } \alpha(\boldsymbol{\xi},\tau)} \delta^{h(\tau)-\mathfrak{g}(u)}(r_u, \tau)\big),$$

where the formulas $\delta^n(t,s)$ say that t is at least n moments before s: that is, $\delta^0(t,s)$ is $(t=s)$ and $\delta^n(t,s)$ is

$$\begin{array}{ll} \exists s_1,\dots,s_{n-1}(t < s_1 < \dots < s_{n-1} < s), & \text{if } n>0,\\ \exists s_1,\dots,s_{|n|-1}(t > s_1 > \dots > s_{|n|-1} > s), & \text{if } n<0. \end{array}$$

Take a fresh variable x_h and associate with h the formula

$$\mathsf{w}_h = \exists r_a \exists x_h \,\big[\mathsf{ext}^{\mathcal{T}}_{\exists R}(x_h, r_a) \;\wedge\; \bigwedge_{h(\xi)=a} (\xi = x_h) \;\wedge\; \mathsf{t}_h\big].$$

To give the intuition behind w_h, suppose that $\mathcal{C}_{(\mathcal{T},\mathcal{A})} \models^g \mathsf{w}_h$, for some assignment g. Then g maps all terms in $\mathcal{X}_h$ to $g(x_h) \in \mathsf{ind}(\mathcal{A})$ such that $\exists R(g(x_h), g(r_a)) \in \mathcal{C}_{(\mathcal{T},\mathcal{A})}$, so $(g(x_h), g(r_a))$ is the root of a substructure of $\mathcal{C}_{(\mathcal{T},\mathcal{A})}$ isomorphic to $\mathcal{C}_{\mathcal{K}_{\mathcal{T},R}}$ in which the variables from $\mathcal{Y}_h$ can be mapped according to h. For temporal terms, the formula t_h cannot specify the values prescribed by h: without $\neg$ in UCQs, we can only say that τ is at least (not exactly) n moments before τ'. However, by Lemmas 3 and 4, this is still enough to ensure that g and h give a homomorphism from $\boldsymbol{q}_h$ to $\mathcal{C}_{(\mathcal{T},\mathcal{A})}$.

Example 4. Let $\mathcal{T}$ be the same as in Example 3 and let

$$\boldsymbol{q}(x,t) = \exists y, z, t' \,\big((t < t') \wedge Q(x,y,t) \wedge S'(y,z,t')\big).$$

The map $h = \{x \mapsto a, y \mapsto v, z \mapsto u_1, t \mapsto 1, t' \mapsto 2\}$ is a witness for R, with $\boldsymbol{q}_h = \boldsymbol{q}$ and w_h is the following formula

$$\exists r_a \exists x_h \,\big(\mathsf{ext}^{\mathcal{T}}_{\exists R}(x_h, r_a) \wedge (x_h = x) \wedge \\ \exists r_v \exists r_{u_1} \big(\delta^0(r_a, r_v) \wedge \delta^1(r_v, r_{u_1}) \wedge \delta^1(r_v, t) \wedge \delta^1(r_{u_1}, t')\big)\big).$$

We now define a rewriting for $\boldsymbol{q}(\boldsymbol{x},\boldsymbol{s})$ and $\mathcal{T}$. Let $\mathfrak{T}$ be the set of all witnesses for $\boldsymbol{q}$ and $\mathcal{T}$. We call $\mathfrak{S} \subseteq \mathfrak{T}$ *consistent* if $(\mathcal{X}_{h_1} \cup \mathcal{Y}_{h_1}) \cap (\mathcal{X}_{h_1} \cup \mathcal{Y}_{h_2}) \subseteq \mathcal{X}_{h_1} \cap \mathcal{X}_{h_2}$, for any distinct $h_1, h_2 \in \mathfrak{S}$. Assuming that $\boldsymbol{y}$ and $\boldsymbol{t}$ contain all the existentially quantified individual and temporal variables in $\boldsymbol{q}$ and $\boldsymbol{q} \setminus \mathfrak{S}$ is the sub-query of $\boldsymbol{q}$ obtained by removing the atoms in $\boldsymbol{q}_h$, $h \in \mathfrak{S}$, other than $(\tau < \tau')$ and $(\tau = \tau')$, we set:

$$\boldsymbol{q}^*(\boldsymbol{x},\boldsymbol{s}) \;=\; \exists \boldsymbol{y}, \boldsymbol{t} \bigvee_{\substack{\mathfrak{S} \subseteq \mathfrak{T}\\ \mathfrak{S} \text{ consistent}}} \big(\bigwedge_{h \in \mathfrak{S}} \mathsf{w}_h \;\wedge\; \mathsf{ext}^{\mathcal{T}}_{\boldsymbol{q} \setminus \mathfrak{S}}\big).$$

Theorem 7. *Let $\mathcal{T}$ be a TQL TBox in CoNF and $\boldsymbol{q}(\boldsymbol{x},\boldsymbol{s})$ a totally ordered CQ. Then, for any consistent KB $(\mathcal{T},\mathcal{A})$ and any tuples $\boldsymbol{a} \subseteq \mathsf{ind}(\mathcal{A})$ and $\boldsymbol{n} \subseteq \mathbb{Z}$, $(\mathcal{T},\mathcal{A}) \models \boldsymbol{q}(\boldsymbol{a},\boldsymbol{n})$ iff $\mathcal{A}^{\mathbb{Z}} \models \boldsymbol{q}^*(\boldsymbol{a},\boldsymbol{n})$.*

Theorem 1 now follows by Lemma 1.

6 Non-Rewritability

In this section, we show that the language *TQL* is nearly optimal as far as rewritability of CQs and ontologies is concerned.

We note first, that the syntax of *TQL* allows concept inclusions and role inclusions; 'mixed' axioms such as the datalog rule $A(x,t) \wedge R(x,y,t) \to B(x,t)$ are not expressible. The reason is that mixed rules often lead to non-rewritability, as is well known from $\mathcal{EL}$. For example, for $\mathcal{T} = \{A(y,t) \wedge R(x,y,t) \to A(x,t)\}$, there is no FO-query $\boldsymbol{q}(x,t)$ such that $(\mathcal{T},\mathcal{A}) \models A(a,n)$ iff $\mathcal{A}^{\mathbb{Z}} \models \boldsymbol{q}(a,n)$ since such a query has to express that at time-point t there is an R-path from x to some y with $A(y,t)$.

Second, it would seem to be natural to extend *TQL* with the temporal next/previous-time operators as concept or role constructs. However, again this would lead to non-rewritability: any FO-rewriting for $A(x,t)$ and the TBox $\{\bigcirc_{P} A \sqsubseteq B, \bigcirc_{P} B \sqsubseteq A\}$ has to express that there exists $n \geq 0$ such that $A(x,t-2n)$ or $B(x,t-(2n+1))$, which is impossible [16].

Another natural extension would be inclusions of the form $A \sqsubseteq \Diamond_{F} B$. (Note that inclusions of the form $A \sqsubseteq \exists R.B$ are expressible in *OWL 2 QL*.) But again such an extension would ruin rewritability. The reason is that temporal precedence $<$ is a total order, and so one can construct an ABox $\mathcal{A}$ and a UCQ $\boldsymbol{q}(x) = \boldsymbol{q}_1 \vee \boldsymbol{q}_2$ such that $(\mathcal{T},\mathcal{A}) \models \boldsymbol{q}(a)$ but $(\mathcal{T},\mathcal{A}) \not\models \boldsymbol{q}_i(a)$, $i = 1,2$, for $\mathcal{T} = \{A \sqsubseteq \Diamond_{F} B\}$. Indeed, we can take $\mathcal{A} = \{A(a,0), C(a,1)\}$ and

$$\begin{aligned}
\boldsymbol{q}_1(x) &= \exists t\,(C(x,t) \wedge B(x,t)),\\
\boldsymbol{q}_2(x) &= \exists t,t'\,((t < t') \wedge C(x,t) \wedge B(x,t')).
\end{aligned}$$

In fact, by reduction of 2+2-SAT [21], we prove the following:

Theorem 8. *Answering CQs over the TBox $\{A \sqsubseteq \Diamond_{F} B\}$ is* coNP*-hard for data complexity.*

7 Conclusion

In this paper, we have proved UCQ rewritability for conjunctive queries and *TQL* ontologies over data instances with validity time. Our focus was solely on the existence of rewritings, and we did not consider efficiency issues such as finding shortest rewritings, using temporal intervals in the data representation or mappings between temporal databases and ontologies. We only note here that these issues are of practical importance and will be addressed in future work. It would also be of interest to investigate the possibilities to increase the expressive power of both ontology and query language. For example, we believe that the extension of *TQL* with the next/previous time operators, which can only occur in TBox axioms not involved in cycles, will still enjoy rewritability. We can also increase the expressivity of conjunctive queries by allowing the arithmetic operations $+$ and $\times$ over temporal terms, which would make the CQ $A(x,t)$ and the TBox $\{\bigcirc_{P} A \sqsubseteq B, \bigcirc_{P} B \sqsubseteq A\}$ rewritable in the extended language.

References

1. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The DL-Lite family and relations. J. of Artifical Intelligence Research 36, 1–69 (2009)
2. Artale, A., Franconi, E.: Temporal description logics. In: Handbook of Temporal Reasoning in Artificial Intelligence, pp. 375–388. Foundations of Artificial Intelligence, Elsevier (2005)
3. Artale, A., Kontchakov, R., Ryzhikov, V., Zakharyaschev, M.: Complexity of reasoning over temporal data models. In: Proc. of ER-2010. LNCS, vol. 6412, pp. 174–187. Springer (2010)
4. Artale, A., Kontchakov, R., Wolter, F., Zakharyaschev, M.: Temporal Description Logic for Ontology-Based Data Access (Extended Version). CoRR abs/1304.5185 (2013)
5. Baader, F., Borgwardt, S., Lippmann, M.: Temporalizing ontology-based data access. In: Proc. of CADE-24. LNAI, vol. 7898, pp. 330–344. Springer (2013)
6. Baader, F., Ghilardi, S., Lutz, C.: LTL over description logic axioms. ACM Trans. Comput. Log. 13(3), 21 (2012)
7. Calì, A., Gottlob, G., Lukasiewicz, T.: A general datalog-based framework for tractable query answering over ontologies. J. Web Sem. 14, 57–83 (2012)
8. Franconi, E., Toman, D.: Fixpoints in temporal description logics. In: Proc. of IJCAI 2011. pp. 875–880 (2011)
9. Gabbay, D., Kurucz, A., Wolter, F., Zakharyaschev, M.: Many-dimensional modal logics: theory and applications. Studies in Logic. Elsevier (2003)
10. Gaifman, H., Mairson, H.G., Sagiv, Y., Vardi, M.Y.: Undecidable optimization problems for database logic programs. In: Proc. of LICS 87. pp. 106–115 (1987)
11. Gottlob, G., Schwentick, T.: Rewriting ontological queries into small nonrecursive datalog programs. In: Proc. of DL 2011. CEUR-WS, vol. 745 (2011)
12. Grohe, M., Schwandtner, G.: The complexity of datalog on linear orders. Logical Methods in Computer Science 5(1) (2009)
13. Gutierrez, C., Hurtado, C.A., Vaisman, A.A.: Introducing time into RDF. IEEE Trans. Knowl. Data Eng. 19(2), 207–218 (2007)
14. Gutiérrez-Basulto, V., Klarman, S.: Towards a unifying approach to representing and querying temporal data in description logics. In: Proc. of RR 2012. LNCS, vol. 7497, pp. 90–105. Springer (2012)
15. Kreutzer, S., Otto, M., Schweikardt, N.: Boundedness of monadic FO over acyclic structures. In: Proc. of ICALP 2007. pp. 571–582 (2007)
16. Libkin, L.: Elements Of Finite Model Theory. Springer (2004)
17. Lutz, C., Wolter, F., Zakharyaschev, M.: Temporal description logics: A survey. In: Proc. of TIME 2008. pp. 3–14. IEEE Computer Society (2008)
18. van der Meyden, R.: Predicate boundedness of linear monadic datalog is in PSPACE. Int. J. Found. Comput. Sci. 11(4), 591–612 (2000)
19. Motik, B.: Representing and querying validity time in RDF and OWL: A logic-based approach. J. Web Sem. 12, 3–21 (2012)
20. Pugliese, A., Udrea, O., Subrahmanian, V.S.: Scaling RDF with time. In: Proc. of WWW 2008. pp. 605–614 (2008)
21. Schaerf, A.: On the complexity of the instance checking problem in concept languages with existential quantification. J. of Intel. Inf. Systems 2, 265–278 (1993)