Temporalized $\mathcal{EL}$ Ontologies for Accessing Temporal Data: Complexity of Atomic Queries

Víctor Gutiérrez-Basulto$^1$ and Jean Christoph Jung$^1$ and Roman Kontchakov$^2$

$^1$ Department of Computer Science, Universität Bremen, Germany
$^2$ Department of Computer Science and Information Systems, Birkbeck, University of London, UK

Abstract

We study access to temporal data with $\mathcal{T}\mathcal{EL}$, a temporal extension of the tractable description logic $\mathcal{EL}$. Our aim is to establish a clear computational complexity landscape for the atomic query answering problem, in terms of both data and combined complexity. Atomic queries in full $\mathcal{T}\mathcal{EL}$ turn out to be undecidable even in data complexity. Motivated by the negative result, we identify well-behaved yet expressive fragments of $\mathcal{T}\mathcal{EL}$. Our main contributions are a semantic and sufficient syntactic conditions for decidability and three orthogonal tractable fragments, which are based on restricted use of rigid roles, temporal operators, and novel acyclicity conditions on the ontologies.

1 Introduction

In recent years, the use of ontologies to enrich plain data with a semantic layer has become one of the outstanding applications of description logic (DLs) technologies in the Semantic Web. The ontology-based data access (OBDA) setting provides information systems with various advantages, e.g., a friendlier vocabulary for accessing heterogeneous data is given by the ontology, and means of querying potentially incomplete data are provided by taking account of the implicit knowledge derived from the data and the ontology. Due to the increasing need to account for the temporal dimension of data available on the Web [Roth and Tan, 2013; Dong and Tan, 2015], the DL community has recently investigated extensions of the OBDA paradigm for temporal data. The initial efforts concentrated on temporal query languages with atemporal ontologies [Gutiérrez-Basulto and Klarmann, 2012; Klarmann and Meyer, 2014; Baader, Borgwardt, and Lippmann, 2015; Borgwardt, Lippmann, and Thost, 2015; Borgwardt and Thost, 2015]. On the other hand, temporal ontology languages can enhance conceptual modelling with temporal aspects [Artale et al., 2015], which are required, e.g., in applications managing data from sensor networks. In this line, the research has focused on temporal extensions of $\mathcal{DL}$-Lite that support rewratability of temporal queries into the monadic second-order logic with order or into two-sorted first-order logic with $<$ and $+$ [Artale et al., 2013b; 2015]. Since standard relational database management systems have such built-in predicates, they can in principle evaluate the $\text{FO}(<,+)$-rewritings. However, no temporal extensions of other classical DLs have been investigated yet in the context of OBDA, which is partly because of the intractability and often even undecidability of the standard reasoning tasks (e.g., subsumption) [Artale et al., 2007; Gutiérrez-Basulto, Jung, and Lutz, 2012; Gutiérrez-Basulto, Jung, and Schneider, 2014]. On the other hand, temporal data has also been studied in classical database theory [Chomicki and Toman, 2005]. In their seminal paper, Chomicki and Imielski [1988] identified $\text{DATALOG}_{1S}$ as a decidable extension of $\text{DATALOG}$ with one successor function. Here we make the first (to the best of our knowledge) attempt to link temporal OBDA with temporal deductive databases [Chomicki, 1990; Baudinet, Chomicki, and Wolper, 1993].

In this paper, we study $\mathcal{T}\mathcal{EL}$, a temporal extension of $\mathcal{EL}$ [Baader, Brandt, and Lutz, 2005]. The underlying DL component, $\mathcal{EL}$, underpins the OWL 2 EL profile of OWL 2 and the medical ontology SNOMED CT, which provides the vocabulary for electronic health records (EHRs). Indeed, applications managing EHRs must be able to provide information, e.g., on when and for how long some drug has been prescribed to a patient, so that drugs that interact adversely are not prescribed at the same time. Clinical trials [Shankar et al., 2008; O’Connor et al., 2009] also require a unified conceptual model for specifying temporal constraints of protocol entities such as ‘a viable participant should have had a vaccination with live virus 5 days ago’ or ‘blood tests of a patient should be run every 3 days’. These statements can be encoded in $\mathcal{T}\mathcal{EL}$:

\[
\text{Patient} \sqcap \Box_{x,t} \exists \text{vaccinated. LiveVirus} \sqsubseteq \text{ViableParticip}, \quad (1)
\]
\[
\text{Patient} \sqcap \Box_{x,t} \text{ReqBloodTest} \sqsubseteq \text{ReqBloodTest}. \quad (2)
\]

Our main objective is to establish the limits of decidability and tractability of the query answering problem over $\mathcal{T}\mathcal{EL}$ ontologies, in terms of both data and combined complexity. In order to set the foundations, we focus on temporal atomic queries. On the one hand, an atomic query $\text{ViableParticip}(x,t)$ with the temporal concept inclusion (1) effectively encodes a tree-shaped temporal conjunctive query. On the other hand, using (1) to extend the vocabulary with a concept $\text{ViableParticip}$ is closer to the spirit of the OBDA paradigm than repeating the same conjunction in all similar user queries. Moreover, a recurrent pattern $\text{ReqBloodTest}$ is expressible as an atomic query.
We begin by introducing TEL (in the set concept definitions classical DL dimension) are inspired by the 'traditional' notion of acyclic-two novel acyclicity conditions (each constraining only one EL highly parallelizable). This fragment contains many acyclic not contain (2). (EL establish PS construct a polynomial rewriting into D time'). In fragments of surprising (and challenging) results are obtained. But not expressible as a query without temporal concept inclusion ( ReqBloodTest(x, t) with the temporal concept inclusion (2) but not expressible as a query without temporal concept inclusions like (2). As we shall see, even for atomic queries rather surprising (and challenging) results are obtained.

Our main contributions are complexity bounds, algorithms and rewritability into DATALOG[S] for atomic query answering in fragments of TEL. Since query answering over unrestricted TEL turns out to be undecidable (in data complexity), we investigate its fragments to attain decidability and tractability. First, for TEL^O, which allows only the 'next' O_r and 'previous-time' O_r operators, we identify ultimate periodicity as a natural semantic condition ensuring decidability; more precisely, PSPACE data complexity (the question of decidability of the full TEL^O is left open for future work). Then, we identify a number of fragments with better computational properties. (i) For the fragment of TEL^O without rigid (not changing over time) roles on the right-hand side of concept inclusions, we construct a polynomial rewriting into DATALOG[S], and so, establish PSPACE-completeness for data complexity. This fragment contains all EL ontologies as well as both (1) and (2). (ii) Over temporally acyclic TEL^O-ontologies (with rigid roles and concepts), query answering is PTIME-complete in both data and combined complexity. This tractable fragment contains (1) and fully captures all atemporal EL ontologies and may prove particularly useful in applications; it, however, does not contain (2). (iii) Query answering over DL-acyclic TEL^O ontologies is NC^1-complete for data complexity (in principle, highly parallelizable). This fragment contains many acyclic EL ontologies as well as both (1) and (2) (note that large parts of SNOMED CT are in fact acyclic). We remark that our two novel acyclicity conditions (each constraining only one dimension) are inspired by the 'traditional' notion of acyclicity in (temporal extensions of) DLs [Haase and Lutz, 2008; Gutiérrez-Basulto, Jung, and Schneider, 2015]. Finally, (iv) we show that the language with only O_r and O_r (sometime in the past/future) on the left-hand side of concept inclusions also enjoys PTIME query answering.

2 Preliminaries

We begin by introducing TEL, a temporal extension of the classical DL EL. Let N_C, N_R, N_I be countably infinite sets of concept, role and individual names, respectively. We assume that N_R is partitioned into two infinite sets, N_R^f and N_R^c, of rigid and local role names, respectively. TEL concepts are defined by the following grammar:

\[
C, D :: \; A \mid C \circledcirc D \mid \exists_r C \mid O_r C \mid O_r C,
\]

where A ∈ N_C, r ∈ N_R, and s ∈ \{F, P\}. A TEL-TBox (ontology) T is a finite set of concept inclusions (CI) C ⊆ D and concept definitions (CDs) C ⊆ D for TEL concepts C, D. Data is given in terms of temporal ABoxes A, which are finite sets of assertions of the form A(a, n) and r(a, b, n), where A ∈ N_C, r ∈ N_R, a, b ∈ N_I, and n ∈ Z. We denote by ind(A) the set of individuals naming in A, and by tem(A) the set \{n ∈ Z \mid min[A] ≤ n ≤ max[A]\}, where min[A] and max[A] are, respectively, the minimal and maximal time points in A. The size, |T| and |A|, of T and A is the number of symbols required to write T and A, respectively, with time points n ∈ Z encoded in unary. A temporal knowledge base (KB) K is a pair (T, A).

An interpretation I is a structure (Δ^3, (I_a)_a∈Z), where each I_a is a classical DL interpretation with domain Δ^3; we have A^I_a ⊆ Δ^3 and r^I_a ⊆ Δ^3 x Δ^3. Rigid roles r ∈ N^f_R do not change their interpretation in time: r^I_n = r^I_0 for all n ∈ Z. We usually write A^I_n and r^I_n instead of A^I and r^I, respectively, and the mapping A^I_n is extended to complex TEL-concepts as follows:

\[(C \circledcirc D)^I_n = C^I_n \circledcirc D^I_n,\]

\[(\exists_r C)^I_n = \{ d \mid \text{there is } e \in C^I_n \text{ with } (d, e) \in r^I_n \},\]

\[(O_r C)^I_n = C^I_n \circledcirc r^I_n,\]

\[(O_r C)^I_n = \{ d \mid d \in C^I_n \circledcirc r^I_n \},\]

where op_a stands for − if s = P and for + if s = F. We use strict ⊆, (k > 0) but our results do not depend on the choice.

An interpretation I is said to be a model of C ⊆ D, written I |= C ⊆ D, if C^I_n ⊆ D^I_n, for all n ∈ Z; and a model of C ⊆ D if C^I_n = D^I_n, for all n ∈ Z. We call I a model of a TBox T, written I |= T, if I |= α for all α ∈ T. Note that TBoxes are interpreted globally in the sense that all CIs and CDs must be satisfied at every time point. A concept D subsumes a concept C with respect to T, written T |= C ⊆ D, if I |= C ⊆ D for all models I of T.

For ABoxes A we adopt the standard name assumption: a^I_n = a for all a ∈ ind(A), n ∈ Z (and thus ind(A) ⊆ Δ^3).

The relation |= is extended to ABoxes by taking I |= A(a, n) iff a ∈ A^I_n and I |= r(a, b, n) iff (a, b) ∈ r^I_n; I is a model of A iff I |= α for all α ∈ A. An interpretation I is a model of a KB (T, A), written I |= (T, A), iff I |= T and I |= A.

Finally, K |= A(a, n) if I |= A(a, n) for every model I of K.

As the query language, we consider temporal atomic queries (TAQs) of the form A(x, t) with A ∈ N_C, x an individual variable and t a temporal variable. Given K = (T, A), a certain answer to A(x, t) over K is a pair (a, n) ∈ ind(A) × tem(A) with K |= A(a, n). We study the complexity of the query answering problem over temporal knowledge bases:

<table>
<thead>
<tr>
<th>TAQ answering</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> TBox T, ABox A, TAQ A(x, t), a pair (a, n).</td>
</tr>
<tr>
<td><strong>Question:</strong> Is (a, n) a certain answer to A(x, t) over (T, A)?</td>
</tr>
</tbody>
</table>

Our results concern both the combined and data complexity of the problem: for data complexity, the TBox is fixed. As usual, for a complexity class C and a class X of TBoxes, we say that TAQ answering over X is C-hard in data complexity if there is some T ∈ X such that answering TAQs over T is C-hard. Conversely, TAQ answering over X is in C in data complexity if, for all T ∈ X, answering TAQs over T is in C.

As classes X, we will in particular look at full TEL and its fragments TEL^O and TEL^O in which, respectively, only the temporal operators O_r and O_r are allowed. Note that O_r on the left-hand side and ⊆, (with the usual semantics) on the right-hand side of CIs can be expressed in TEL^O, e.g., instead of O_r A ⊆ X or, equivalently, A ⊆ O_r X, take A ⊆ A' and O_r A' ⊆ A' ⊆ X, for a fresh A'. Thus, rigid concepts, which do not change their interpretation in time, can be expressed in these two fragments using O_r O_r on the left-hand side of CIs.
3 Query Answering in \textit{TEL}: Undecidability

We first pinpoint different sources of complexity for the query answering problem in \textit{TEL} in order to identify computationally well-behaved fragments in the sequel.

We begin by showing that TAQ answering over \textit{TEL} is undecidable. The known undecidability of subsumption in \textit{TEL} [Artale et al., 2007] translates only into the combined complexity of TAQ answering. We strengthen the result to obtain undecidability in data complexity by reducing the halting problem for the universal Turing machine. We exploit the crucial observation that disjunction, although not in the syntax, can be simulated with \(\sqcap\) [Artale et al., 2007].

**Theorem 1** TAQ answering over \textit{TEL} is undecidable in data complexity.

The proof can also be adapted to the non-strict semantics of \(\sqcap\), using the chessboard technique [Gabbay et al., 2003].

\textit{TEL}, unlike \textit{TEL}, is not capable of expressing disjunction.

**Theorem 2** TAQ answering over \textit{TEL} is non-elementary in combined complexity and PSPACE-hard in data complexity.

The proof of PSPACE-hardness is close in spirit to that for \textsc{Datalog}_{LS} [Chomicki and Imieliński, 1988]; note that the lower bound holds even in the restriction of \textit{TEL} without \(\exists r.C\) on the right-hand side of CIs. For the non-elementary lower bound, we take inspiration in the construction for the product modal logic LTL\(\times\)K [Gabbay et al., 2003, Theorem 6.34]. Our proof requires a careful implementation of the yardstick technique [Stockmeyer, 1974] with only Horn formulas.

Decidability of TAQ answering in full \textit{TEL} is left open as interesting and challenging future work; more insights on the difficulty of the problem are given in Sec. 4. Nevertheless, we show that extending \textit{TEL} with certain DL constructs that are harmless for data complexity of atemporal query answering [Krisnadhi and Lutz, 2007] immediately leads to undecidability. Let \textit{TEL} and \textit{TEL} be the extensions of \textit{TEL} with inverse roles \(r^*\) and functionality axioms \(\text{func}(r)\), respectively.\(^1\) For both languages, we reduce the halting problem for the universal Turing machine to prove:

**Theorem 3** TAQ answering over \textit{TEL} and \textit{TEL} is undecidable in data complexity.

In the rest of the paper, we study decidability and complexity of TAQ answering in various fragments of \textit{TEL} and \textit{TEL}.

4 Foundations of Query Answering in \textit{TEL}

In this section, we lay the groundwork for the development of algorithms for query answering in fragments of \textit{TEL} by introducing canonical quasimodels, which are succinct abstract representations of the universal model of the KB, see also [Artale et al., 2013b; 2015]. They can also be viewed as a generalization of the canonical structures used for query answering in pure \(\mathcal{EL}\) [Lutz, Toman, and Wolter, 2009].

In the sequel, we assume that \textit{TEL}-TBoxes are in normal form, that is, they consist of CIs of the form

\[
A \sqcap A' \sqsubseteq B, \quad A \sqsubseteq \exists r.B, \quad X \sqsubseteq A,
\]

with the usual semantics: \((r^*)^n = \{(e,d) \mid (d,e) \in r^n\}\); \(\models \text{func}(r)\) iff \(e_1 = e_2\), for all \((d,e_1), (d,e_2) \in r^n\), and \(n \in \mathbb{Z}\).

where \(A, A'\) and \(B\) are concept names and \(X\) is a basic concept of the form \(A, \sqcap A, \sqcup A, \exists r.A\), for a concept name \(A\). Observe that, without loss of generality, \(\sqcup\), is restricted to the left-hand side of CIs: for instance, \(A \sqcup C \sqsubseteq B\) is equivalent to \(\sqcup A \sqsubseteq B\).

It is routine to show that every \textit{TEL}-TBox can be transformed into the normal form by introducing fresh concept names; see, e.g., [Baader, Brandt, and Lutz, 2005].

Fix now a KB \((T, \alpha)\) with a \textit{TEL} TBox \(T\) in normal form and let \(CN\) be the set of concept names in \((T, \alpha)\). A map \(\pi: \mathbb{Z} \rightarrow 2^{CN}\) is a trace for \(T\) if it satisfies the following:

\((t1)\) if \(A \sqcap A' \subseteq B \in T\) and \(A, A' \in \pi(n)\), then \(B \in \pi(n)\);

\((t2)\) if \(\exists r.A \sqsubseteq B \in T\) and \(A \in \pi(n)\), then \(B \in \pi(n op, 1)\).

These are the building blocks of quasimodels: they represent the temporal evolution of individual domain elements. For traces, \(\pi\) such that \(\pi(i) = \{B\}\) for odd \(i\) and \(\{C\}\) for even \(i\) is a trace for \(T\).

In order to describe interactions of domain elements, we require more notation. Let \(\pi\) be a trace for \(T\). For a rigid role \(r \in \mathbb{N}_r\), the r-projection of \(\pi\) is a map \(\pi_r\) \(\mathbb{Z} \rightarrow 2^{CN}\) that sends each \(i \in \mathbb{Z}\) to \(\{A \mid \exists r.B \sqsubseteq A \in T, B \in \pi(i)\}\); for a local role \(r \in \mathbb{N}_r\), \(\pi_r\) is defined in the same way on 0 but is \(\emptyset\) for all other \(i \in \mathbb{Z}\). Given a map \(\pi: \mathbb{Z} \rightarrow 2^{CN}\) and \(n \in \mathbb{Z}\), we say that \(\pi\) contains the n-shift of \(\gamma\) and write \(\gamma \subseteq^{n} \pi\) if \(\gamma(i - n) \subseteq \pi(i)\), for all \(i \in \mathbb{Z}\). For example, let \(T = \{\exists r.B \sqsubseteq B'\}\) with rigid role \(r\). In the picture below, the trace \(\pi_a\) contains the 1-shift of the r-projection of \(\pi_B\):

\[
\begin{array}{cccccccc}
& & & & B' & & & \\
\pi_B & & & & B & & & \\
\pi_a & & & & B & & & \\
\end{array}
\]

If \(r\) is local then \(\pi_a\) has to contain \(B'\) only at 1 (but not at 3, etc.). We are now fully equipped to define quasimodels.

Let \(D = \text{ind}(A) \cup CN\) henceforth. A quasimodel \(\Omega\) for \((T, \alpha)\) is a set of traces \(\pi_d, d \in D, \) for \(T\) such that

\((q1)\) \(A \in \pi_a(n)\), for all \(A(a, n) \in A\);

\((q2)\) \(B \in \pi_B(0)\), for all \(B \in CN\);

\((q3)\) \(\tau_{r,r}\) \(\subseteq^{0} \pi_a\), for all \(r(a,b, n) \in A\);

\((q4)\) if \(A \in \pi_d(n)\) then \(\tau_{r,r} \subseteq^{n} \pi_a\), for all \(d \in D, Z \in \mathbb{Z}\) and \(A \sqsubseteq \exists r.B \in T\).

Intuitively, quasimodels represent models of \((T, \alpha)\): each \(\pi_a\) stands for the ABox individual \(a\); each \(\pi_B\), on the other hand, represents all individuals that witness \(B\) for CIs \(A \sqsubseteq B\in T\). The latter is, in fact, the crucial abstraction underlying quasimodels. Note that traces \(\pi_a\) are normalized: \(B\) occurs at time point 0, which is compensated by the shift operation in \((q4)\). For example, in the picture above, if \(A \sqsubseteq B \in T\) then, in any model, \(a\) has an r-successor that belongs to \(B\) at moment 1. Such a successor can be obtained as a ‘copy’ of trace \(\pi_B\) shifted by 1 so that its origin, 0, matches moment 1 for \(a\). Then, by \((q4)\), \(a\) belongs to \(B'\) at all odd moments.

For the purposes of query answering we need to identify canonical (minimal) quasimodels. We define the canonical quasimodel as the limit of the following saturation (chase-like) procedure. Start with initially empty maps \(\pi_a\), for \(d \in D\), and apply \((t1)\)–\((t2)\), \((q1)\)–\((q4)\) as rules: \((q3)\), for example, says ‘if
r(a, b, n) ∈ A and A ∈ proj_p(π_d)(i), then add A to π_a(i). Then we have the following characterization:

**Theorem 4** Let Ω = {π_d | d ∈ D} be the canonical quasi-model of (T, A) with T a TEL^O-TBox. Then, for any A ∈ CN, (T, A) |= A(a, i) iff A ∈ π_a(i), for a ∈ ind(A), i ∈ Z.

The procedure for constructing the canonical quasi-model deals with infinite data structures (traces) and is generally not terminating. So, although Theorem 4 provides a criterion for certain answers, it does not immediately yield a decision algorithm for full TEL^O. We remark that known techniques for dealing with such infinite structures cannot be easily applied: for example, MSO (over Z), a standard tool for decidability proofs in temporal DLs [Gabbay et al., 2003], is not sufficient to encode the canonical quasi-model directly because (q4) requires +. In fact, the key to showing decidability for (fragments of) TEL^O is finding a finite representation of traces.

The starting point of the rest of the paper is a semantic condition on the canonical quasi-model, ultimate periodicity, which ensures decidability in data complexity. Let T be a TEL^O-TBox and Ω the canonical quasi-model for (T, Ω). We say that T is ultimately periodic, if there is p ∈ N such that all π_B, B ∈ CN, in Ω are ultimately p-periodic, that is, for each B ∈ CN, there are positive integers m_p, p_p, m_p, p_p ≤ p satisfying the following conditions:

\[
π_B(n - m_p) = π_B(n), \text{ for all } n ≤ -m_p,
\]

\[
π_B(n + p_p) = π_B(n), \text{ for all } n ≥ m_p.
\]

Intuitively, an ultimately p-periodic trace has repeating sections on the left and on the right:

<table>
<thead>
<tr>
<th>Section</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>-m_p - 2p_p</td>
<td>(π_B(n - m_p) )</td>
</tr>
<tr>
<td>-m_p</td>
<td>(π_B(n) )</td>
</tr>
<tr>
<td>m_p</td>
<td>(π_B(n + p_p) )</td>
</tr>
<tr>
<td>m_p + 2p_p</td>
<td>(π_B(n) )</td>
</tr>
</tbody>
</table>

The condition of ultimate periodicity is rather natural. On the practical side, it is motivated by applications with recurrent patterns such as health care support [Shankar et al., 2008], see CIs (1) and (2) in Section 1. From the theoretical point of view, any satisfiable LTL formula has an ultimately periodic model [Manna and Wolper, 1984].

We next show that ultimate periodicity is indeed sufficient for decidability in data complexity.

**Theorem 5** TAT answering over ultimately periodic TEL^O-TBoxes is PSPACE-complete in data complexity.

PSPACE-hardness follows from the proof of PSPACE-hardness for entailment in Horn-LTL [Chen and Lin, 1993] and PTIME-hardness from atomic query answering in EC. For the upper bounds, let (T, A) be a KB with T a TEL^O-TBox and Ω its canonical quasi-model. We take a propositional variable P_A,d for each A ∈ CN and d ∈ D and construct a Horn-LTL formula \(ϕ_T,A\) whose minimal model is isomorphic to Ω: variable P_A,d is true in the model at moment n iff A ∈ π_d(n).

We take the conjunction of the following formulas, for d ∈ D:

\[\Box(P_A,d ∧ P_{A'},d → P_{B},d),\] for A ∩ A' ⊆ B ∈ T,

\[\Box(ϕ_T,A → P_{B},d),\] for A, A ⊆ B ∈ T,

\[[o^n P_A,a],\] for A(a, n) ∈ A,

\[P_{B},B,\] for B ∈ CN,

\[o^n P_{B,b} → [o^n P_A,a],\] for r(a, b, n) ∈ A, ∃r. B ⊆ A ∈ T,

\[P_{B}',B → [o^n P_A,d → P_{A'},d],\] for A ⊆ ∃r. B, B' ⊆ A' ∈ T,

where \(o^n\) is Op^n if n ≥ 0 and Op^-n if n < 0 and \(\Box\) is the 'globally' operator. It is readily verified that \(ϕ_T,A\) is as required. Crucially, (q4) for local roles boils down to the last formula above. Since entailment in LTL is in PSPACE [Sistla and Clarke, 1985] and \(ϕ_T,A\) is polynomial in the size of (T, A), we obtain membership in PSPACE for combined complexity.
To obtain the PTIME data complexity, observe that traces $\pi_T$, $B \in \CN$, are ultimately $2^{|T|}$-periodic because they are
to the canonical quasimodel for $(T, \emptyset)$; so, they can be
maintained in constant space. Next, traces $\pi_T, a \in \text{ind}(A)$, are
ultimately $2^{|T|+|A|}$-periodic, but a closer inspection reveals
that the middle irregular section, $m_0 + m_c$, is bounded by
$|A| + 2^{|T|}$, while both periods, $p_a$ and $p_e$, by $2^{|T|}$; see, e.g.,
Lemma 3 [Artale et al. 2013a]. Thus, $\Omega$ can be stored
in space bounded by a polynomial in $|A|$. Since each rule application
extends the traces, the saturation procedure for constructing
$\Omega$ terminates in polynomial time in the size of $A$.

Since ontologies without rigid roles at all may be too re-
strictive for applications, we consider $\TEC_0$-TBoxes where
rigid roles are allowed only in CIs of the form $\exists r.B \sqsubseteq A$.

**Theorem 7** TAQ answering over $\TEC_0$-is PSPACE-complete
in data complexity and in EXPTIME in combined complexity.

PSPACE-hardness in data complexity follows from the proof
of Theorem 2. For the upper bounds, we construct rewritings
into DATALOG$_{1, S}$, similarly to $\Pi_T$ in Section 4 (Theorem 5).

**6 Acyclicity Conditions**

It is known that acyclicity conditions can lead to better com-
plexity. In particular, acyclic TBoxes are a way of obtaining
CTL-based temporal extensions of $\mathcal{EL}$ that have rigid roles
and enjoy PTIME subsumption [Gutiérrez-Basulto, Jung, and
Schneider, 2015]. In DATALOG$_{1, S}$, a restriction on recursion
has also been used to attain tractability [Chomicki, 1990].
From the application point of view, large parts of SNOMED CT
and GO [Gene Ontology Cons., 2001] are indeed acyclic. So,
we believe that the fragments we consider below are well-
suited for temporal extensions of such ontologies.

**Acyclic TBoxes** are finite sets of CDs $A \equiv C, A \in \CN$, such
that no two CDs have the same left-hand side, and there are
no CDs $A_1 \equiv C_1, \ldots, A_k \equiv C_k$ in $T$ such that $A_{k+1}$ occurs
in $C_i$, for all $1 \leq i \leq k$ (where $A_{k+1} := A_1$). We say $A$ is
defined in $T$ if $A \equiv C \in T$ and primitive otherwise.

**Theorem 8** TAQ answering over acyclic $\TEC^0$ is in LOG-
TIME-uniform $\mathcal{AC}^0$ in data complexity and in PTIME in com-
bined complexity.

The LOGTIME-uniform $\mathcal{AC}^0$ upper bound is established by
rewriting into $\FO(+)$: for a given TAQ $A(x,t)$ and TBox $T$,
we construct a two-sorted first-order formula $\varphi_{T, A}(x,t)$
with functions $+1$ and $-1$ on temporal terms such that
$(T, A) \models A(a, i)$ iff $A$ (viewed as an interpretation) is a
model of $\varphi_{T, A}(a, i)$, for all ABoxes $A$, $a \in \text{ind}(A)$, $i \in \mathbb{Z}$.
We adapt the technique developed for atemporal $\mathcal{EL}$ [Bienvenu,
Lutz, and Wolter, 2012]:

$$\varphi_{T, A}(x, t) = S_A(x, t), \quad \text{if $A$ is primitive},$$
$$\varphi_{T, A}(x, t) = S_A(x, t) \vee \varphi_{T, C}(x, t), \quad \text{if $A \equiv C \in T$},$$
$$\varphi_{T, \exists r.B}(x, t) = \exists y (R_r(x, y, t) \land \varphi_{T, B}(y, t)),$$
$$\varphi_{T, \forall r.B}(x, t) = \varphi_{T, B}(x, t \ op^1_1),$$

where $S_A(x, t)$ is a disjunction of all $B(x, t)$ for concept
names $B$ with $T \models B \subseteq A$, and $R_r(x, y, t) = r(x, y, t)$ for
$r \in \mathbb{N}^R_{\text{REG}}$ and $\exists f r(x, y, t)$ for $r \in \mathbb{N}^R_{\text{REG}}$.

Note that $\varphi_{T, A}$ is an FO$^2$-rewriting in the terminology of Ar-
tale et al. [2013b; 2015] because the temporal terms range
over $\mathbb{Z}$. However, the infinite interpretation of $A$ is empty after
at most $|T|$ steps from the ABox and so, $\varphi_{T, A}$ can be con-
verted into an FO-rewriting whose temporal terms range over
term$(A)$ only; see [Artale et al., 2015].

We next introduce novel notions of acyclicity that restrict
only one dimension, DL or temporal.

**DL Acyclicity**

First, we introduce DL-acyclic $\TEC^0$-TBoxes, which are well-
suited as temporal extensions of, say, biomedical ontologies
that may require recurrent patterns but have an acyclic DL
component. A $\TEC^0$-TBox $T$ with concept names $\CN$ is called
DL-acyclic if there is a mapping $\ell_{\CN} : \CN \to \mathbb{N}$ such that:

(i) $A \sqsubseteq \exists r.B$ or $\exists r.B \sqsubseteq A \in T$ implies $\ell_{\CN}(A) > \ell_{\CN}(B)$;
(ii) $\bigcirc, A \sqsubseteq B$ implies $\ell_{\CN}(A) = \ell_{\CN}(B)$;
(iii) $A \sqcap A' \sqsubseteq B \in T$ implies $\ell_{\CN}(A) = \ell_{\CN}(A') = \ell_{\CN}(B)$.

A DL-acyclic TBox is of depth $k$ if $k$ is minimal such that a
witnessing mapping $\ell_{\CN}$ satisfies $\ell_{\CN}(B) \le k$ for all $B \in \CN$.

**Theorem 9** TAQ answering over DL-acyclic $\TEC^0$-Boxes of
depth $k \geq 1$ is $k$-EXPSPACE-complete in combined complexity
and NC$^1$-complete in data complexity.

A closer inspection of the non-elementary lower bound proof
in Theorem 2 reveals that the TBox used is DL-acyclic and
TAQ answering over TBoxes of depth $k$ is $k$-EXPSPACE-hard.
NC$^1$-hardness in data complexity follows [Artale et al., 2015]
by reduction of the word problem of NFAs to TAQ answering,
even without the DL dimension.

For the matching upper bounds, fix $(T, A)$ with $T$ of
depth $k$. We devise a completion procedure, which is based
on special LTL-formulas and implies ultimate periodicity of
all traces in the canonical quasimodel of $(T, A)$; cf. Section 5.
Given any $A$, let the slice $A_i$ consist of all $A(a, i) \in A$, all
$r(a, b, i) \in A$ and all $r(a, b, i)$ with $r(a, b, j) \in A$, for some $j$, and
$r \in \mathbb{N}^R_{\text{REG}}$. The algorithm separates consequences of the
role structure of $A$ and local temporal consequences of $T$. In
particular, it exhaustively extends $A$ by all $(a, i)$ with either

$$(T, A_i) \models A(a, i) \text{ or } B(a, i \ op^1_1) \in \CN, \bigcirc, B \sqsubseteq A \in T.$$ (6)

It turns out that $A_i$ in (6) can be replaced by its suitably defined
quotient $B_i$. Intuitively, the logic can only distinguish distinct
trees of depth $k$, whose number depends on $|T|$ only; so, the
size of $B_i$ is independent of $|A|$. By induction on depth $k$, we
define LTL-formulas $\varphi_{a,i}$ of $k$-fold-exponential size char-
terizing all $A \in \CN$ with $(T, B_i) \models A(a, i)$: we begin from formulas as in Theorem 6; the induction step takes account of
the structure of $B_i$ and incurs an exponential blowup.

Now, for combined complexity, observe that each of the
polynomially many $\varphi_{a,i}$ can be analyzed in $k$-EXPSPACE.
For data complexity, observe that checking $(T, B_i) \models A(a, i)$
can be done in constant time. The second option in (6), however,
cannot be implemented directly as the number of steps depends
on $|A|$. Instead, we construct a Büchi automaton that accepts
precisely the traces for $T$ and cast the second option in (6)
as the question of whether all traces extending $A$ have $A$ at
position $i$, which is a regular property and so, is in NC$^1$. 
Temporal Acyclicity

We next relax acyclicity by admitting recursion in the DL dimension (but not in temporal); thus, temporally acyclic TBoxes include general $\mathcal{EL}$-TBoxes. A $\mathcal{TEL}^\omega$-TBox $T$ with concept names $CN$ is temporally acyclic if there is $\ell_0: CN \to \mathbb{N}$ such that:

(i) $\lor F A \sqsubseteq B$ or $\lor F B \sqsubseteq A \in T$ implies $\ell_0(B) = \ell_0(A) + 1$;
(ii) $\exists \tau B \sqsubseteq A \lor A \sqsubseteq \exists \tau B \in T$ implies $\ell_0(A) = \ell_0(B)$;
(iii) $A \sqcap A' \sqsubseteq B \in T$ implies $\ell_0(A) = \ell_0(A') = \ell_0(B)$.

Temporally acyclic TBoxes cannot, unlike $DL$ acyclic ones, express rigid concepts. Still, we can partition concept names $NC$ into local $l^{\text{nc}}_\omega$ and rigid $l^{\text{rig}}_\omega$ and obtain the following:

Theorem 10 TAQ answering over temporally acyclic $\mathcal{TEL}^\omega$ (with rigid concepts) is $\text{PTIME}$-complete in data and combined complexity.

The lower bounds are from $\mathcal{EL}$. For the upper bounds, we show a small quasimodel property: traces of the canonical quasimodel of any $(T, A)$ with such a TBox $T$ satisfy

\[
\pi_a(j) = \pi_a(j'), \quad \text{if } j, j' > u + |T| \text{ or } j, j' < l - |T|, \\
\pi_B(j) = \pi_B(j'), \quad \text{if } j, j' > |T| \text{ or } j, j' < -|T|,
\]

where $u = \max A$ and $l = \min A$. Intuitively, the canonical quasimodel has a restricted temporal extension that stretches only $|T|$ time points beyond $A$. By the small quasimodel property, the procedure for constructing the canonical quasimodel can be implemented in polynomial time: traces $\pi_a$ require only polynomial space, and rules (q1)-(q4) extend the traces.

Inflationary $\mathcal{TEL}^\omega$

Next, we follow an approach suggested by Artale et al. [2013b] (in the context of temporal $DL$-Lite) and restrict $\mathcal{TEL}^\omega$ by allowing $\diamond$, only on the left-hand side of CIs. This fragment is denoted by $\mathcal{TEL}^\omega_{\text{inf}}$, for inflationary $\mathcal{TEL}$ (which is related to inflationary $\mathcal{DATALOG}_{1S}$ [Chomicki, 1990]). Note that $\mathcal{TEL}^\omega_{\text{inf}}$ extends general $\mathcal{EL}$-TBoxes. Yet, the complexity is the same:

Theorem 11 TAQ answering over $\mathcal{TEL}^\omega_{\text{inf}}$ is $\text{PTIME}$-complete in both data and combined complexity.

We need to show only the upper bounds. Observe that $\mathcal{TEL}^\omega_{\text{inf}}$ can still be viewed as a fragment of $\mathcal{TEL}^\omega$; see Section 2. In fact, one can show an analogue of Theorem 4 with the following replacement of (t2):

\[(t^{'2}) \text{ if } \square a, A \sqsubseteq B \in T \text{ and } A \in \pi_d(n), \text{ then } B \in \pi_d(n') \text{ for all } n' > n \text{ if } * = P \text{ and for all } n' < n \text{ if } * = F.\]

We establish a special shape of the traces in the canonical model of any $(T, A)$. Let $\varrho: \mathbb{Z} \to \mathbb{Z}^{CN}$ be a map and let $t, u \in \mathbb{Z}$ with $t \leq u$. We say that $\varrho$ is an $[t, u]$-bow tie if

- for all $i > u$, we have $\varrho(i + 1) \geq \varrho(i)$, and if $\varrho(i + 1) = \varrho(i)$ then all $\varrho(i')$, for $i' \geq i$, coincide;
- symmetrically, for all $i < t$, we have $\varrho(i - 1) \geq \varrho(i)$, and if $\varrho(i - 1) = \varrho(i)$ then all $\varrho(i')$, for $i' \leq i$, coincide.

These properties mean that $\varrho$ grows monotonically to the right of $u$ and to the left of $l$; in other words, $\varrho$ has inflationary behaviour. We prove that the traces $\pi_d$ in the canonical quasimodel $\Omega$ of $(T, A)$, for any $A$, enjoy the following properties:

- $\pi_a$ is a $[\min A, \max A]$-bow tie, for each $a \in \text{ind}(A)$;
- $\pi_B$ is a $[0, 0]$-bow tie, for each $B \in CN$.

Thus, the traces in $\Omega$ can be represented in polynomial space because only the middle section and at most $|CN|$ steps at both ends need to be stored. Since the traces are extended with every rule application, the procedure terminates polynomially many steps; Theorem 11 follows.

7 Discussion and Future Work

We summarize the fragments of $\mathcal{TEL}$, their relationships and the obtained complexity results in the following diagram:

\[
\begin{array}{c}
\text{undecidable} \\
\text{PSpace} \text{ \quad $\mathcal{TEL}^\omega$ \quad non-elem} \\
\text{PTime} \text{ \quad $\mathcal{TEL}^\omega_{\text{inf}}$ \quad in PTime} \\
\text{\quad $\mathcal{TEL}^\omega_{\text{int}}$ \quad in PTime and PACE} \\
\text{\quad $\mathcal{TEL}^\omega_{\text{acyc}}$ \quad in PTime and PSpace} \\
\text{\quad $\mathcal{TEL}^\omega_{\text{acyc}}$ \quad in PT and PSpace} \\
\text{\quad $\mathcal{TEL}^\omega_{\text{acyc}}$ \quad in PTime} \\
\end{array}
\]

where the solid lines are inclusions of DLs, the dashed line is a reduction that preserves answers to all queries (model conservative extension). The data complexity is indicated by shading and the combined complexity is below the language.

Our data-tractability results show theoretical adequacy of the identified fragments of $\mathcal{TEL}$ for data-intensive applications. Our two novel forms of acyclicity, DL- and temporal, are somewhat close in spirit to multi-separability [Chomicki, 1990]; the latter, however, puts a weaker restriction on recursion but a stricter one on the interaction between the temporal and data component. DL-acyclic $\mathcal{TEL}^\omega$ is the first (to the best of our knowledge) DL shown to have $\text{NC}^1$-complete query answering (the large gap between data and combined complexity is also remarkable). On the practical side, there is evidence that such data-tractable fragments should be sufficient for many biomedical applications. Following the principles of OBDA, our framework provides a means of defining temporal concepts in the ontology for these applications: temporal concepts capture both (restricted) tree-shaped temporal conjunctive queries (CQs) and recurring temporal patterns.

As our immediate future work, we will address decidability of (full) $\mathcal{TEL}^\omega$ and then consider CQs with the $+$ operation on temporal terms. We expect that our positive results can be lifted to CQs using the combined approach [Lutz, Toman, and Wolter, 2009], which utilizes a structure similar to our canonical quasimodel. We will also study succinct and expressive representations of temporal data. For example, the only known algorithm for $\mathcal{DATALOG}_{1S}$ with binary encoding of timestamps in the data runs in $\text{ExpTime}$ in the size of the data [Chomicki and Imieliński, 1988]. We, however, presume that careful materialization should be sufficient to deal with the issue. We will also consider interval encoding of temporal ABoxes, e.g., $A(a, [n_1, n_2])$, and settings capturing infinite temporal periodic data as introduced by Kabanza, Stévenne, and Wolper [1990] and Chomicki and Imieliński [1993].
A Additional Preliminaries

**DATALOG$_1S$.** We introduce DATALOG$_1S$ along the lines of [Chomicki, 1990; Chomicki and Imieliński, 1988]. Temporal terms are expressions of the form $i$, $t \pm i$ where $t$ is a temporal variable and $i$ is a non-negative integer encoded in unary. A non-temporal term is a constant or variable name. A temporal rule is a formula of the form

$$B(\bar{x}, t) \leftarrow A_1(\bar{x}_1, t_1), \ldots, A_k(\bar{x}_k, t_k),$$

where $B, A_1, \ldots, A_k$ are predicate symbols, $\bar{x}_i$ are (tuples of) non-temporal terms and $t, t_1, \ldots, t_k$ are temporal terms. A DATALOG$_1S$ program is a finite set of temporal rules. A temporal database is a finite set of ground atoms $B(\bar{a}, i)$. Note that temporal databases are restricted to predicates of arity two or three are, syntactically, nothing else than temporal ABoxes; thus, we will make no difference between them unless it could cause confusion. Given a set of temporal rules $\Pi$, a temporal database $D$, and a query $\mathcal{Q}(x, t)$, we say that a pair $(\bar{a}, i)$ of a constant name $a$ and a time point $i$ is an answer and write $(\Pi, A) \models A(\bar{a}, i)$ if $A(\bar{a}, i)$ is true in the minimal model $\mathcal{M}$ of $\Pi$ and $D$, see [Chomicki, 1990]. (The least Herbrand model $\Pi(\bar{a})$ can alternatively be defined as the result of exhaustive application of rules in $\Pi$ to $\bar{a}$.) It is known that query answering in DATALOG$_1S$ is PSPACE-complete in data complexity and EXPTIME-complete in combined complexity [Chomicki and Imieliński, 1988].

**Horn-LTL.** We introduce syntax and semantics of (a sufficiently expressive version of) propositional Horn-LTL, see also [Chen and Lin, 1993]. Let $P$ be a countably infinite set of propositional variables. Propositional Horn-LTL formulas $\varphi$ are built according to the following grammar:

$$\psi, \psi' ::= p \mid \bigcirc \psi \mid \psi \land \psi',$$

$$\varphi, \varphi' ::= \psi \mid \psi \rightarrow \psi' \mid \varphi \land \varphi' \mid \Box \varphi',$$

where $p \in P$. We often write $\bigcirc^n \psi$ for $\bigcirc^n \psi_n$ if $n \geq 0$ and $\bigcirc^n \psi$ for $n < 0$.

A temporal interpretation for a Horn-LTL formulas is a structure $\mathfrak{M} = (W_p)_{p \in P}$, where $W_p \subseteq \mathbb{Z}$. The relation $i \models p$, if $i \in W_p$;

$$(\mathfrak{M}, i) \models \bigcirc \psi \quad \text{if} \quad (\mathfrak{M}, i \oplus 1) \models \psi;$$

$$(\mathfrak{M}, i) \models \varphi_1 \land \varphi_2 \quad \text{if} \quad (\mathfrak{M}, i) \models \varphi_1 \quad \text{and} \quad (\mathfrak{M}, i) \models \varphi_2;$$

$$(\mathfrak{M}, i) \models \varphi_1 \rightarrow \varphi_2 \quad \text{if} \quad (\mathfrak{M}, i) \models \varphi_1 \quad \text{implies} \quad (\mathfrak{M}, i) \models \varphi_2;$$

$$(\mathfrak{M}, i) \models \Box \varphi \quad \text{if} \quad (\mathfrak{M}, j) \models \varphi \quad \text{for all} \quad j \in \mathbb{Z}.$$}
- $Q_i \sqsubset \Diamond D^r \sqcap \Diamond D^r$ to make sure that every state has the two markers around it;
- concept inclusions to identify defects in the encoding of each configuration (by marking them with concept $F$):
  \[
  A_a \sqcap A_{a'} \subseteq F, \quad \text{for all } a \neq a',
  \]
  \[
  Q_i \sqcap Q_j \subseteq F, \quad \text{for all } i \neq j,
  \]
  \[
  \overline{Q} \sqcap Q_i \subseteq F, \quad \text{for all } i,
  \]
  \[
  Q_i \sqcap \Diamond D^r \subseteq F, \quad \text{for all } i,
  \]
  \[
  Q_i \sqcap \Diamond D^r \subseteq F, \quad \text{for all } i
  \]
  (the last two CIs ensure that the $D^r$ and $D^f$ markers can only occur immediately next to a $Q_i$ marker);
- concept inclusions to identify defects of transitions between successive configurations (again by marking them with $F$):
  - non-active cells do not change:
    \[
    \overline{Q} \sqcap A_a \sqcap \exists_r.A_{a'} \subseteq F, \quad \text{for all } a \neq a';
    \]
  - if $\delta(q_i, a) = (q_k, b, \ell)$ then
    \[
    Q_i \sqcap A_a \sqcap \exists_r.A_{a'} \subseteq F, \quad \text{for } c \neq b,
    \]
    \[
    D^f \sqcap \Diamond P (Q_i \sqcap A_a) \sqcap \exists_r.Q_j \subseteq F, \quad \text{for } j \neq k,
    \]
    \[
    D^f \sqcap \Diamond P (Q_i \sqcap A_a) \sqcap \exists_r.Q \subseteq F;
    \]
  - similarly, if $\delta(q_i, a) = (q_k, b, r)$ then
    \[
    Q_i \sqcap A_a \sqcap \exists_r.A_{a'} \subseteq F, \quad \text{for } c \neq b,
    \]
    \[
    D^r \sqcap \Diamond P (Q_i \sqcap A_a) \sqcap \exists_r.Q_j \subseteq F, \quad \text{for } j \neq k,
    \]
    \[
    D^r \sqcap \Diamond P (Q_i \sqcap A_a) \sqcap \exists_r.Q \subseteq F;
    \]
- concept inclusions detecting that a non-blank symbol appears to the right of the input word in the initial configuration:
  \[
  I \sqcap \Diamond P A_a \subseteq F, \quad \text{for all } a \in \Gamma \setminus \{b\};
  \]
- concept inclusions to propagate the defects:
  \[
  \exists_r.F \subseteq F \quad \text{and} \quad \Diamond r.F \subseteq F.
  \]

An input word $w = a_1 \cdots a_n$ is encoded in the ABox $A_w$ with the following assertions:

- $A(a, 0), Q_0(a, 1), A_{a_1}(a, 1), \ldots, A_{a_n}(a, n), I(a, n)$.

Intuitively, any “defect-free” model of $(\mathcal{T}_\mathcal{A}, A_w)$ describes an infinite computation of $\mathcal{A}$ on input $w$; more precisely, we can show that

$(\mathcal{T}_\mathcal{A}, A_w) \models F(a, 0)$ if and only if $\mathcal{A}$ halts on input $w$.

This completes the proof of Theorem 1. $\square$

We defer the proof of Theorem 2 until the end of the section and present first proofs that are more similar in the structure to the proof of Theorem 1.

**Lemma 12** There is a $\mathcal{T}_{L^C}^0$-TBox $\mathcal{T}$ using only a single role name $r \in N^R_{\mathbb{R}}$ and no $\exists_r.B$ on the right-hand side of CIs such that $\mathcal{T}_{AQ}$ answering over $\mathcal{T}$ is PSPACE-hard in data complexity.

**Proof.** We reduce the word problem for deterministic linear space-bounded Turing machines (LBAs), similarly to [Chomicki, 1990]. Let $\mathcal{A} = (Q, \Sigma, \Gamma, 0, \delta, q_0, F)$ be an LBA with a PSPACE-hard word problem. The TBox, $\mathcal{T}_{\mathcal{A}}$, uses the following concept names (cf. the proof of Theorem 1):

- $A_a$, for each $a \in \Gamma$, to encode the content of a tape cell;
- $Q_i$, for $q_i \in Q$, to represent the current state;
- $\overline{Q}$ for marking the no-head cells; in addition, we use $\overline{Q}^f$ and $\overline{Q}^r$ for the non-head cells on the left and right of the active cell, respectively;
- $G$ to propagate an accepting state.

We also use a single rigid role name $r \in N^R_{\mathbb{R}}$. For the sake of readability, we introduce two auxiliary concept names, $P$ and $S$, and abbreviate the complex concepts $\exists_r(P \sqcap \exists_r.C)$ and $\exists_r(S \sqcap \exists_r.C)$ by $\exists\text{pred}.C$ and $\exists\text{succ}.C$, respectively. A configuration of $\mathcal{A}$ is encoded in the DL dimension: neighbouring cells are connected back-and-forth via succ and pred, respectively, as shown in the following picture:

![Diagram showing the connection between cells and roles](attachment:diagram.png)

The TBox $\mathcal{T}_{\mathcal{A}}$ consists now of the following set of concept inclusions. For a transition $\delta(q_i, a) = (q_k, b, r)$, we take

$\exists\text{pred}.(Q_i \sqcap A_a) \sqcap \Diamond O_Q q_k$,

$Q_i \sqcap A_a \sqcap \Diamond r.A_b$.

For a transition $\delta(q_i, a) = (q_k, b, \ell)$, we take

$\exists\text{succ}.(Q_i \sqcap A_a) \sqcap \Diamond O_Q q_k$,

$Q_i \sqcap A_a \sqcap \Diamond r.A_b$.

The content of non-active cells is preserved by concept inclusions, for $a \in \Gamma$,

$\overline{Q} \sqcap A_a \sqcap \Diamond r.A_a$

and the following concept inclusions that propagate the no-head markers along the tape:

$\exists\text{pred}.Q_i \sqsubseteq \overline{Q}^r$, $\exists\text{succ}.Q_i \sqsubseteq \overline{Q}^r$, for $q_i \in Q$,

$\exists\text{pred}.\overline{Q}^f \sqsubseteq \overline{Q}^f$, $\exists\text{succ}.\overline{Q}^f \sqsubseteq \overline{Q}^f$.

The auxiliary concept names $S$ and $P$ are simply propagated into the future:

$S \sqsubseteq \Diamond r.S$, $P \sqsubseteq \Diamond r.P$.  

When an accepting state is reached, we activate the marker \( G \) and propagate it back to the left-most cell at moment 0:

\[ Q_i \subseteq G, \quad \text{for } q_i \in F, \]
\[ \exists \text{succ.} G \subseteq G. \]

An input word \( w = a_1 \cdots a_n \) is encoded as an ABox \( A_w \) with \( \text{ind}(A_w) = \{c_i \mid 1 \leq i \leq n\} \cup \{p_i, s_i \mid 1 \leq i \leq n\} \) and the following assertions that initialize the tape and fix the succ and pred relations:

\[
Q_0(c_1, 0), \ A_{a_i}(c_i, 0), \quad \text{for } 1 \leq i \leq n,
\]
\[
r(c_i, s_i, 0), S(s_i, 0), r(s_i, c_{i+1}, 0), \quad \text{for } 1 \leq i < n,
\]
\[
r(c_{i+1}, p_i, 0), P(p_i, 0), r(p_i, c_i, 0), \quad \text{for } 1 \leq i < n.
\]

It is routine to verify that \( \mathfrak{A} \) accepts input \( w \) just in case \( \mathcal{T}_A(A_w) = G(c_1, 0) \).

\[ \square \]

**Theorem 3** TAQ answering over \( TELC^C \) and \( TELF \) is undecidable in data complexity.

**Proof.** We give a proof only for \( TELC^C \), a proof for \( TELF \) is a straightforward adaptation, which is indicated below.

We reduce the halting problem of the universal Turing machine \( \mathfrak{A} = (Q, q_0, \Sigma, \Gamma, \delta, F) \), which, similarly to the proof of Theorem 1, is assumed to be deterministic. Again, we assume that \( \mathfrak{A} \) works on a one-end infinite tape and that the left-most cell is labelled with a special marker \( b \) (which is never overwritten) and the right end and everything beyond is labelled with another special symbol, \( a \). We use the following concept names, with the intuition similar to the proof of Lemma 12:

- \( A_a \), for \( a \in \Gamma \), to represent contents of the tape;
- \( Q_i \), for \( q_i \in Q \), to represent the current state;
- \( \overline{Q} \) for marking the no-head cells; \( \overline{Q} \) and \( \overline{Q} \) are used for the non-head cells on the left and right of the active cell, respectively;
- \( G \) for propagating an accepting states.

(Note, however, that unlike in the proof of Lemma 12, we use the temporal dimension to represent configurations and the DL dimension for transitions between the configurations.)

As in the proof of Theorem 1, a configuration of \( \mathfrak{A} \) is represented along the temporal dimension, that is, by the temporal evolution of a certain individual over time. The computation of \( \mathfrak{A} \) is done in the DL dimension along an infinite \( r \)-chain, for a rigid role name \( r \).

We start by enforcing that everything right of \( b \) is labeled with \( b \) using concept inclusion

\[ A_{b'} \subseteq \text{op}_b A_{b'}. \]

We next introduce the constraints needed to synchronise two successive configurations. To this end, we enforce every cell left and right of the state to be labelled by \( \overline{Q} \) by adding

\[
\text{op}_b Q_i \subseteq \overline{Q}, \quad \text{op}_b Q_i \subseteq \overline{Q}, \quad \text{for all } q_i \in Q,
\]
\[
\overline{Q} \subseteq \overline{Q}, \quad \overline{Q} \subseteq \overline{Q}, \quad \overline{Q} \subseteq \overline{Q}.
\]

For transitions between successive configuration, we include the following concept inclusions:

\[
Q_i \cap A_a \subseteq \exists \text{r.}(A_b \cap \overline{Q} Q_k), \quad \text{if } \delta(q_i, a) = (q_k, b, r),
\]
\[
Q_i \cap A_a \subseteq \exists \text{r.}(A_b \cap \overline{Q} Q_k), \quad \text{if } \delta(q_i, a) = (q_k, b, \ell).
\]

Non-active cells (those labelled by \( \overline{Q} \)) are not changed, so we add, for all \( a \in \Gamma \):

\[
\exists r^-.(\overline{Q} \cap A_a) \subseteq A_a.
\]

(Note that this is the only concept inclusion that uses inverse roles. If \( r \) is functional, we can use \( \overline{Q} \cap A_a \subseteq \exists r. A_a \) instead.) Finally, for identifying an accepting configuration, we include

\[
Q_i \subseteq G, \quad \text{for } q_i \in F,
\]
\[
\exists \text{succ.} G \subseteq G.
\]

The ABox \( A_w \) encodes an input word \( w = a_1 \cdots a_n \) as the following assertions:

\[
A(a, 0), Q_0(a, 1), A_{a_1}(a, 1), \ldots, A_{a_n}(a, n), \quad A_{b'}(a, n + 1).
\]

It is now routine to verify that \( \mathcal{T}_A(A_w) \models G(a, 0) \) iff \( \mathfrak{A} \) accepts input \( w \).

\[ \square \]

**Theorem 2** TAQ answering over \( TELC^C \) is non-elementary in combined complexity and \( \text{PSPACE-hard} \) in data complexity.

**Proof.** \( \text{PSPACE-hardness} \) is by Lemma 12. For the non-elementary lower bound we take inspiration in the construction for the product modal logic LTL \( \times K \) [Gabbay et al., 2003, Theorem 6.34]. Our proof requires a careful implementation of the yardstick technique [Stockmeyer, 1974] using only Horn formulas. The proof consists of two steps: first, we show how to encode arbitrarily large elementary numbers; second, we prove \( k \)-\( \text{EXPSPACE-hard} \) for every \( k \geq 1 \). We concentrate on the first step, because, given the encoding of the numbers, it is straightforward to establish the respective lower bounds.

We define function \( \exp_k \) by taking

\[ \exp_0(n) = n, \quad \exp_{k+1}(n) = \exp_k(n) \cdot 2^{\exp_k(n)}. \]

**Lemma 13** For each \( k, n \geq 1 \), there is a DL-acyclic \( TELC^C \)-TBox \( \mathcal{T}_k,n \), and concept names \( \text{Init}_k, \text{Zero}_k, \overline{\text{Zero}}_k \) such that

\[
\mathcal{T}_k,n \models \text{Init}_k \subseteq \text{op}_k^m \overline{\text{Zero}}_k \iff m = 0 \mod \exp_k(n), \quad (7)
\]
\[
\mathcal{T}_k,n \models \text{Init}_k \subseteq \text{op}_k^m \text{Zero}_k \iff m \neq 0 \mod \exp_k(n). \quad (8)
\]

**Proof.** Fix \( n \geq 1 \). We use single rigid role \( r \) and throughout the proof, concepts of the form \( \overline{A} \) will be intended as the complements of respective \( A \).

The proof is by induction on \( k \). For the basis of induction, \( k = 1 \), we describe a \( TELC^C \)-TBox \( \mathcal{T}_1,n \), that models a binary counter with \( n \) bits along the temporal dimension (these will be counters at level 1). The TBox is based on the following concept names:

- \( S_1 \) and \( \overline{S}_1 \) to indicate the start of the encoding of the counter,
– $B_1$ and $\overline{B}_1$ to represent the bits of the binary counter;
– $C_1$ and $\overline{C}_1$ for the value of the carry bit;
– $\text{Init}_1$, $\text{Zero}_1$ and $\text{Zero}_0$ for the claim of the lemma.

First, $\mathcal{T}_{1,n}$ sets out delimiters each $n$ moments of time:

\begin{align*}
\text{Init}_n & \sqsubseteq S_1, \\
S_1 & \sqsubseteq \bigcup_0^n S_1, \\
S_1 & \sqsubseteq \bigcup_0^n \overline{S}_1, \text{ for } 1 \leq i < n.
\end{align*}

Then, we write the representation of 0 as the first value of the counter:

$$\text{Init}_1 \sqsubseteq \overline{B}_1 \sqcap \bigcup_0^1 \overline{B}_1 \sqcap \ldots \sqcap \bigcup_0^n \overline{B}_1.$$  \hfill (12)

The following CIs increment the value of the $n$-bit counter (with the least significant bit at the smallest time point):

\begin{align*}
S_1 \sqcap \bigcup_0^n B_1 & \sqsubseteq \overline{B}_1 \sqcap \bigcup_0 C_1, \\
S_1 \sqcap \bigcup_0^n \overline{B}_1 & \sqsubseteq B_1 \sqcap \bigcup_0 \overline{C}_1, \\
S_1 \sqcap C_1 \sqcap \bigcup_0^n B_1 & \sqsubseteq \overline{B}_1 \sqcap \bigcup_0 C_1, \\
S_1 \sqcap C_1 \sqcap \bigcup_0^n \overline{B}_1 & \sqsubseteq B_1 \sqcap \bigcup_0 \overline{C}_1, \\
S_1 \sqcap \overline{C}_1 \sqcap \bigcup_0^n B_1 & \sqsubseteq \overline{B}_1 \sqcap \bigcup_0 \overline{C}_1, \\
S_1 \sqcap \overline{C}_1 \sqcap \bigcup_0^n \overline{B}_1 & \sqsubseteq B_1 \sqcap \bigcup_0 \overline{C}_1. 
\end{align*}

\hfill (13)–(16)

Intuitively, $S_1$ marks the positions for the least significant bit in the binary counter. Then, (13)–(16) flip all the bits until 0 is reached. After that, (18) and (17) copy the remaining bits (until the delimiter, $S_1$, is reached). Concept names $\text{Zero}_0$ and $\text{Zero}_1$ are then related to the binary representation of the counter by means of

\begin{align*}
S_1 \sqcap \bigcup_0^n B_1 & \sqsubseteq \text{Zero}_1, \\
S_1 \sqcap \bigcup_0^n \overline{B}_1 & \sqsubseteq \text{Zero}_1, \text{ for } 0 \leq i < n.
\end{align*}

\hfill (19)–(20)

It should be clear that $\mathcal{T}_{1,n}$ satisfies (7) and (8). We only remark that, due to the structure of the CIs above, if $d \in \text{Init}_1 \sqcap \mathcal{T}_{1,n}$, then at each moment $m \geq t$, the element $d$ belongs either to $A$ or $\overline{A}$, for all concepts $A$ of the form $B_1$, $C_1$ and $S_1$. On the other hand, there is always a model of $\mathcal{T}_{1,n}$ in which, for all concepts names $A$ as above, none of the domain elements belongs to both $A$ and $\overline{A}$.

For the inductive step, let $k > 1$. By the induction hypothesis, there is a TBox $\mathcal{T}_{k-1,n}$, satisfying (7) and (8) for $k - 1$. So, $\mathcal{T}_{k-1,n}$ is the first ingredient of $\mathcal{T}_{k,n}$; let $\mathcal{T}_{k,n}$ contain all concept inclusions of $\mathcal{T}_{k-1,n}$. As before, a counter value is represented by a sequence of $\exp_{k-1}(n)$ consecutive time points but the main difficulty is now to relate two consecutive counter values. For that purpose, we use the following symbols (in addition to the signature of $\mathcal{T}_{k-1,n}$):

– $S_k$ and $\overline{S}_k$ to indicate the start of the encoding of the counter value at level $k$,
– $Z_k$ to write down the initial counter value of 0,
– $B_k$ and $\overline{B}_k$ to represent the bits of the binary counter,
– $C_k$ and $\overline{C}_k$ and for the value of the carry bit,

– $Y_k$ and $\overline{Y}_k$ to record the value of the bit and communicate it between two consecutive numbers,
– $I_k$ and $\overline{I}_k$ to detect sequences of set bits of length $\exp_{k-1}(n)$,
– $\text{Delim}_k$, $\text{Keep}_k$, $\text{Read}_k$ to distinguish different uses of the counters.

– $\text{Init}_k$, $\text{Zero}_k$ and $\text{Zero}_0$ as required in claim of the lemma.

Intuitively, the level $k$ counter is realized by first enforcing a sequence of $\exp_{k-1}(n)$ zeros, that is, $\overline{B}_k$, along the temporal dimension, and then enforcing that, if there is a sequence of length $\exp_{k-1}(n)$ encoding some number, say $M$, then the subsequent sequence of the same length encodes $M + 1$. We mark the beginnings of a number using concepts $S_k$ and $\overline{S}_k$ and a special counter of type $\text{Delim}_k$ of level $k - 1$ as follows:

$$\text{Init}_k \sqsubseteq \exists r. (\text{Delim}_k \sqcap \text{Init}_{k-1}),$$

\hfill (21)

$$\exists r. (\text{Delim}_k \sqcap \text{Zero}_{k-1}) \sqsubseteq S_k,$$

\hfill (22)

$$\exists r. (\text{Delim}_k \sqcap \text{Zero}_0) \sqsubseteq \overline{S}_k;$$

\hfill (23)

cf. (9)–(11). Next, we write down 0 as the first value of the counter, that is, we enforce the first $\exp_{k-1}(n)$ time points belong to $\overline{B}_k$:

$$\text{Init}_k \sqsubseteq Z_k,$$

\hfill (24)

$$S_k \sqcap \bigcup_0 Z_k \sqsubseteq Z_k,$$

\hfill (25)

$$Z_k \sqsubseteq \overline{B}_k;$$

\hfill (26)

cf. (12). In order to correctly increment the level $k$ counter, we have to communicate between time point elements having a distance of $\exp_{k-1}(n)$. The following concept inclusions create a individual and store the value of the current bit there:

$$B_k \sqsubseteq \exists r. (\text{Keep}_k \sqcap \text{Init}_{k-1} \sqcap Y_k),$$

\hfill (27)

$$\overline{B}_k \sqsubseteq \exists r. (\text{Keep}_k \sqcap \text{Init}_{k-1} \sqcap \overline{Y}_k),$$

\hfill (28)

$$\text{Keep}_k \sqcap \text{Init}_{k-1} \sqsubseteq \bigcup_0 Y_k,$$

\hfill (29)

$$\text{Keep}_k \sqcap \text{Init}_{k-1} \sqsubseteq \bigcup_0 \overline{Y}_k.$$

\hfill (30)

In order to transfer the stored value, $Y_k$ or $\overline{Y}_k$, precisely $\exp_{k-1}(n)$ time points, the relevant time point is marked with $\text{Read}_k$:

$$\text{Zero}_{k-1} \sqcap \bigcup_0 \text{Keep}_k \sqsubseteq \text{Keep}_k,$$

\hfill (31)

$$\text{Zero}_{k-1} \sqcap \bigcup_0 \text{Keep}_k \sqsubseteq \text{Read}_k.$$
\[
\overline{S}_k \cap \overline{C}_k \cap \exists r. (\text{Read}_k \cap Y_k) \subseteq B_k \cap \overline{O}_r \overline{C}_k.
\]

These CIs work in precisely the same way as (13)–(18) for the case of \( k = 1 \) except that now we use \( \exists r. (\text{Read}_1 \cap Y_1) \) and \( \exists r. (\text{Read}_k \cap Y_k) \) in place of \( \overline{O}_r B_1 \) and \( \overline{O}_r B_k \).

To obtain correct valuations for concepts \( \text{Zero}_k \) and \( \overline{\text{Zero}}_k \), we require an auxiliary concept \( I_k \) defined by the following set of CIs:

\[
S_k \cap B_k \subseteq I_k,
\]
\[
B_k \subseteq \tau_k,
\]
\[
\overline{S}_k \cap \overline{O}_r I_k \cap B_k \subseteq I_k,
\]
\[
\overline{S}_k \cap \overline{O}_r \overline{I}_k \subseteq \tau_k.
\]

Informally, \( I_k \) identifies long runs of set bits that begin at one of the delimiters. With \( I_k \) and \( \tau_k \) at hand, we add the following CIs to define \( \text{Zero}_k \) and \( \overline{\text{Zero}}_k \):

\[
\text{Init}_k \subseteq \text{Zero}_k,
\]
\[
S_k \cap \overline{O}_r I_k \cap \text{Zero}_k,
\]
\[
S_k \cap \overline{O}_r \overline{I}_k \cap \overline{\text{Zero}}_k,
\]
\[
\overline{S}_k \cap \overline{\text{Zero}}_k.
\]

This completes the construction of \( T_{k,n} \) and it remains to show (7) and (8).

Let \( \mathcal{O} \) be a model of \( T_{k,n} \) and \( d \in \text{Init}_k^2 \). Since the arguments are independent of \( t \), we will assume \( t = 0 \) from now on. Let \( N = \exp_{k-1}(n) \). By (21), there is some \( e \in \Delta^2 \) such that \( (d,e) \in r^3 \) and \( e \in \text{Init}_k^1 \cap \text{Delim}_k^m \). By (22), we have \( e \in \text{Delim}_k^2 \) for all \( i \geq 0 \); by (7) and (8) of the induction hypothesis, we have, for all \( i \geq 0 \),

\[
e \in \text{Zero}_k^1 \quad \text{if} \quad i = 0 \mod N,
\]
\[
e \in \overline{\text{Zero}}_k^1 \quad \text{if} \quad i \neq 0 \mod N.
\]

By (23) and (24), we obtain

\[
d \in S_k^2 \quad \text{if} \quad i = 0 \mod N,
\]
\[
d \in \overline{S}_k^2 \quad \text{if} \quad i \neq 0 \mod N.
\]

For each \( i \geq 0 \), define \( \text{bit}(i) \) as follows:

\[
\text{bit}(i) = \begin{cases} 
\{0\} & \text{if } d \in B_k^2 \setminus \overline{B}_k^2, \\
\{1\} & \text{if } d \in B_k^2 \setminus \overline{B}_k^2, \\
\{0,1\} & \text{otherwise}.
\end{cases}
\]

Starting from that, we define for each \( m \geq 0 \) a function \( \text{val}(m) \) reading off the \( m \)-th counter value:

\[
\text{val}(m) = \left\{ \sum_{i=0}^{N-1} 2^i b_i \mid b_i \in \text{bit}(mN + i) \right\}.
\]

Clearly, \( \text{val}(m) \subseteq \{0, \ldots, 2^N - 1\} \).

Claim 1. If \( m \mod 2^N \in \text{val}(m) \), for all \( m \geq 0 \).

Proof of Claim 1. The proof is by induction on \( m \). For the induction base, let \( m = 0 \). Clearly, it suffices to show \( 0 \in \text{bit}(i) \), for every \( 0 \leq i < N \). By (25)–(27) and (49), \( d \in B_k^2 \) for all \( 0 \leq i < N \).

For the induction step, take \( m > 0 \). By induction, we know that \( m - 1 \) mod \( 2^N \in \text{val}(m-1) \). We show the following characterization of a single incrementation operation via an inductive argument for all \( mN \leq i < (m+1)N \):

\begin{enumerate}[label=(P\arabic*)]
\item If \( 1 \in \text{bit}(j) \) for all \( mN \leq j < i \), then \( d \in C_k^2 \), otherwise, \( d \in C_k^2 \).
\item If \( d \in C_k^2 \), then \( \{1 - v \mid v \in \text{bit}(i-N)\} \subseteq \text{bit}(i) \).
\item If \( d \in C_k^2 \), then \( \text{bit}(i-N) \subseteq \text{bit}(i) \).
\end{enumerate}

By (34) and (48), \( d \in C_k^{2,mN} \). If \( 1 \in \text{bit}((m-1)N) \), then \( d \in B_k^2((m-1)N) \). By (28), there is some \( e \) with \( (d,e) \in r^3 \) and \( e \in (\text{Init}_k^1 \cap \text{Delim}_k^m)^2 \). By (30), \( e \in Y_k^{2,mN} \).

Then, (32)–(33) together with (7) and (8) of the induction hypothesis imply \( e \in \text{Read}_k^{2,mN} \), whence \( d \in \exists r. (\text{Read}_k \cap Y_k)^3 \).

By (36), \( d \in B_k^{2,mN} \), and hence \( 0 \in \text{bit}(mN) \). The carry bit is also propagated to \( mN + 1 \). The transfer of \( Y_k \) in case \( 0 \in \text{bit}((m-1)N) \) is analogous, that is, (37) implies \( 0 \in \text{bit}(mN) \) and \( d \in C_k^{2,mN+1} \).

The remaining cases are treated in a similar way. Then, Properties (P2) and (P3) imply that

\[
\{v + 1 \mod 2^N \mid v \in \text{val}(m-1)\} \subseteq \text{val}(m),
\]

which finishes the proof of Claim 1.

We next show that the maximal value, \( 2^N - 1 \), is recognized correctly.

Claim 2. For all \( m \geq 0 \),

\begin{enumerate}[label=- if \( k \in \text{val}(m) \), for \( k \neq 2^N - 1 \), then \( d \in T_k^{2,(m+1)N-1} \).
\end{enumerate}

Proof of Claim 2. Suppose first that \( 2^N - 1 \in \text{val}(m) \), that is, for each \( 0 \leq i < N \), we have \( 1 \in \text{bit}(mN + i) \) and thus \( d \in B_k^{2,mN+1} \). By (40) and (48), we obtain \( d \in I_k^{2,mN} \). Cl (42) with (49) yield \( d \in I_k^{2,mN+1} \) for all \( 0 \leq i < N \).

Suppose now that \( k \in \text{val}(m) \) with \( k \neq 2^N - 1 \), that is, there is some \( 0 \leq i < N \) with \( 0 \in \text{bit}(mN + i) \) and thus \( d \in B_k^{2,mN+i} \). By (41), \( d \in T_k^{2,mN+i} \). By (43) and (49), \( d \in T_k^{2,mN+i} \), for all \( 0 \leq i < N \). This finishes the proof of Claim 2.

We are now in a position to show (7) and (8) for \( k \). We concentrate on the “if”-directions; the “only if”-directions are consequences of the observation that there is a minimal model of \( T_{k,n} \) and \( \text{Init}_k \) such that, for all concept names \( A \), none of the domain elements belongs to both \( A \) and \( \overline{A} \) (at any time point).

If \( m \neq 0 \mod N \), then \( d \in \overline{S}_k^m \). By (47), \( d \in \overline{\text{Zero}}_k^m \). If \( m = 0 \mod N \), then \( m = m_0N \) for some \( m_0 \). If \( m_0 = 0 \), then \( m = 0 \) and \( d \in \overline{\text{Zero}}_k^m \) by (44). If \( m_0 > 0 \), we distinguish the following two cases.
where

This completes the proof of Theorem 2.

For every configuration of $A$ is routine to do so using the $T$Box corresponding tape cells in two consecutive configurations; it
similarity to the above. The main difficulty is to communicate
the computation of $A$ configuration of $A$.

This finishes the proof of Lemma 13.

$\Box$

2 Foundations of Query Answering in $\text{TEL}^\circ$

In the following lemma we show how to establish the normal form.

Lemma 14 For every $\text{TEL}^\circ$-TBox $T$ with concept names CN, we can construct in polynomial time a $\text{TEL}^\circ$-TBox $T'$ with concept names CN such that $T'$ is a model conservative extension of $T$:

- every model of $T'$ is a model of $T$ and
- every model of $T$ can be extended to a model of $T'$ by giving a suitable interpretation to fresh concept names CN \ \ CN.

Proof. Let sub($T$) be the set of sub-concepts appearing in $T$. For every $C \in \text{sub}(T)$, take a fresh concept name $X_C$ and define concept $C$ as follows:

$$C = \begin{cases} A, & \text{if } C = A \in \text{CN}, \\ \exists r.X_D, & \text{if } C = \exists r.D, \\ \exists r.X_D, & \text{if } C = \exists r.D, \\ X_{D_1} \sqcap X_{D_2}, & \text{if } C = D_1 \sqcap D_2. \end{cases}$$

Now, let $T'$ consist of the following concept inclusions:

- $C \subseteq X_D$, for all $C \subseteq D \in T$;
- $C \subseteq X_C, X_C \subseteq C$, for all $C \in \text{sub}(T)$.

Observe that $T'$ is almost of the required shape: conjunctions on the right-hand side are just abbreviations, and, for instance, $O_p$ on the right-hand side can be expressed via $O_p$ on the left-hand side. Finally, it is routine to prove the two properties mentioned above.

$\Box$

Theorem 4 Let $\Omega = \{ \pi_d \mid d \in D \}$ be the canonical quasi-model of $(T, A)$ with a $\text{TEL}^\circ$-TBox $T$ and $D = \text{ind}(A) \cup \text{CN}$. Then, for every $A \in \text{CN}, a \in \text{ind}(A)$, $i \in \mathbb{Z}$, we have

$$(T, A) \models A(a, i) \iff A \in \pi_a(i).$$

Proof. ($\Rightarrow$) Suppose that $(T, A) \not\models A(a, i)$, that is, there is a model $\mathcal{J} = (\Delta^3, (I_n)_{n \in \mathbb{Z}})$ of $(T, A)$ such that $a \notin A^{3,i}$. We can read off a collection of traces from $\mathcal{J}$ as follows. First assign a trace $\pi_a$ to every $a \in A^3$ by taking:

$$\pi_a(i) = \{ C \in \text{CN} \mid a \in C^{3,i} \},$$

for every $i \in \mathbb{Z}$.

The definition of $\pi_B, B \in \text{CN}$ is slightly more complicated. Intuitively, $\pi_B(i)$ contains a concept name $C$ iff every instance of $B$, say at time point $n$, in $\mathcal{J}$ evolves to an instance of $C$ at time point $n + i$. Formally, we put for $i \in \mathbb{Z}, A \in \text{CN}:

$$\pi_B(i) = \{ C \in \text{CN} \mid \exists B \subseteq \text{CN} \},$$

where $O^i = O^i$ if $i \geq 0$ and $O^i$ if $i < 0$. It remains to verify that the collection $\{ \pi_d \mid d \in D \}$ satisfies properties (t1)–(t2), and (q1)–(q4). Properties (t1), (t2), and (q3) are satisfied since $\mathcal{J} \models T$. Properties (q1) and (q2) are satisfied since $\mathcal{J} \models A$. For verifying (q4), let $C \in \pi_d(n)$, for some $d \in D, n \in \mathbb{Z}$ and $C \subseteq \exists r.B \in T$ for some rigid role $r \in \text{R}^\circ$; the case $r \in \text{R}^\circ$ is similar. Moreover, suppose that $E \in \text{proj}_d(\pi_B(i) - n)$ for some $i \in \mathbb{Z}, E \in \text{CN}$.

We need to show that $E \in \pi_a(d)$. By the definition of $\text{proj}_d$, there is $E' \in \pi_B(i - n)$ such that $\exists r.E' \subseteq E \in T$. We distinguish cases on $d$.

- If $d \in \text{ind}(A)$, then $C \in \pi_d(n)$ implies $d \in C^{3,n}$. Since $\mathcal{J} \models T$ and $C \subseteq \exists r.B \in T$, there is some $d' \in B^{3,n}$ such that $(d, d') \in r^{3,n}$. As $E' \in \pi_B(i - n)$, we know that $\mathcal{J} \models B \subseteq O^{3,n} E'$, and thus $d' \in E^{3,n}$. As $\exists r.E' \subseteq E \in T$, we obtain $d \in E^{3,n}$, whence $E \in \pi_d(d)$.

- If $d \in \text{CN}$, then $C \in \pi_d(n)$ implies $\mathcal{J} \models d \subseteq O^3 \cap \text{CN}$. Since $\mathcal{J} \models T$ and $C \subseteq \exists r.B \in T$, we have $\mathcal{J} \models d \subseteq O^3 \cap \text{CN}$. From $E' \in \pi_B(i - n)$, we obtain $\mathcal{J} \models B \subseteq O^{3,n} E'$, and thus $\mathcal{J} \models d \subseteq O^3 \cap O^{3,n} E'$. As $\exists r.E' \subseteq E \in T$, we obtain $\mathcal{J} \models x \subseteq E^{3,n}$, whence $E \in \pi_d(d)$.

By the initial assumption $a \notin A^{3,i}$ and the definition of $\pi_a$, we have $A \notin \pi_a(i)$. Thus, $A \notin \pi_a(i)$ in the canonical quasi-model, which finishes the proof of the “$\Rightarrow$”-direction.

($\Leftarrow$) Let $\Omega = \{ \pi_d \mid d \in D \}$ be the canonical quasi-model for $(T, A)$. We define the unrolling of $\Omega$ inductively as follows. We start with an interpretation $\mathcal{J}_0 = (\Delta_0, (I_{0n})_{n \in \mathbb{Z}})$ defined by taking:

$$\Delta_0 = \text{ind}(A),$$
\[A^{3,n} = \{a \in \text{ind}(A) \mid A \in \pi_a(n)\},\]
\[r^{3,n} = \{(a, b) \mid r(a, b, n) \in A, r \in R^\text{loc}\} \cup\]
\[\{(a, b) \mid r(a, b, m) \in A, r \in R^\text{reg}\} \cup\]
The interpretation \(J_{i+1}\) is obtained from \(J_i\) by applying the following rule where applicable:

- if \(a \in A^{3,n}\) and \(A \sqsubseteq \exists r.B \in \mathcal{T}\), then extend \(\Delta\) with a fresh element \(b\) and
- if \(r \in R^\text{loc}\), then put \((a, b) \in r^{3,n};\)
- if \(r \in R^\text{reg}\), then put \((a, b) \in r^{3,n,m}\) for all \(m \in \mathbb{Z}\);
- for all \(E \in CN, m \in \mathbb{Z}\), put
  \[b \in E^{3,n,m} \quad \text{iff} \quad E \in \pi_B(m)\].

Obtain \(J\) from \(J_{i+1}\) by taking for all \(n \in \mathbb{N}\), concept names \(A\), and role names \(r\):

\[\Delta = \bigcup_{i \geq 0} \Delta_i, \quad A^{3,n} = \bigcup_{i \geq 0} A^{3,n}_i, \quad r^{3,n} = \bigcup_{i \geq 0} r^{3,n}_i.\]

Let us verify that \(J \models (\mathcal{T}, A)\). The definition of \(J_0\) together with properties (q1) and (q2) yields \(J \models A\). From properties (t1) and (t2) together with the definition of \(A^{3,n}\), we obtain that \(J \models \alpha\) for all \(\alpha \in \mathcal{T}\) of the form \(A_1 \sqcap A_2 \subseteq A\), and \(\Box \alpha \subseteq A\). For the remaining concept inclusions from \(\mathcal{T}\), we provide more detail. Let \(A \sqsubseteq \exists r.B \in \mathcal{T}\) and \(a \in A^{3,n}\) for some \(a \in \Delta\). By the above rule, we have that the freshly introduced element \(b\) satisfies \((a, b) \in r^{3,n}\) and \(b \in B^{3,n}\).

Consider finally a concept inclusion \(\exists r.B \in \mathcal{T}\), domain elements \(a, b \in \Delta^2\) satisfying \((a, b) \in r^{3,n}\) and \(b \in B^{3,n}\). We need to show that \(a \in A^{3,n}\). If \((a, b) \in r^{3,n}\), then we define \(A^{3,n}\) due to the definition of concept names and (q3). If \((a, b)\) was added to \(r^{3,n+1}\) in some rule application, we obtain \(a \in A^{3,n}\) due to the definition of concept names and (q3).

We conclude the proof by noting that if \(a \notin \pi_a(i)\), for some \(A \in CN, \pi_a(i) \in \text{ind}(A), i \in \mathbb{Z}\), then, by definition of \(J\), \(a \notin A^{3,n}\) and hence \((\mathcal{T}, A) \not\models A(a, i)\).

**Theorem 5** TAQ answering over ultimately periodic \(\mathcal{E}\mathcal{C}\)-Boxes is decidable and \(\mathsf{PSPACE}\)-complete in data complexity.

**Proof.** The lower bound is by Lemma 12. The matching upper bound is established in Lemma 15 below for the rewriting \(\Pi\eta\) constructed in Section 4.

**Lemma 15** For all ABoxes \(A, A \in CN, a \in \text{ind}(A), i \in \mathbb{Z}\), we have \((\Pi\eta, A) \models A(a, i)\) iff \((\mathcal{T}, A) \models A(a, i)\).

**Proof.** (\(\Leftarrow\)) Suppose that \((\Pi\eta, A) \not\models A(a, i)\), that is, \(A(a, i) \notin \Pi\eta(A)\), where \(\Pi\eta(A)\) is the minimal Herbrand model of \((\Pi\eta, A)\). We read off a collection of traces \(\Omega = \{\pi_d \mid d \in D\}\) precisely as in the “\(\Leftarrow\)" direction in the proof of Theorem 4, that is, for every \(i \in \mathbb{Z}, a \in \text{ind}(A), B \in CN:\

\[\pi_a(i) = \{C \in CN \mid C(a, i) \in \Pi\eta(A)\},\]
\[\pi_B(i) = \{C \in CN \mid \forall a, n : B(a, n) \in \Pi\eta(A)\} \Rightarrow C(a, n + i) \in \Pi\eta(A)\].

It is routine to verify that \(\Omega\) satisfies (t1), (t2), and (q1)–(q4).

It remains to show that \(\Omega\) satisfies the quasimodel conditions. (t1) is satisfied by formulas (4). For (t2), assume \(A \in \pi_d(n)\) and \(Q \subseteq B \in \mathcal{T}\). The latter and (t2) implies that \(B \in \pi_d(1)\), where \(\{\pi_d \mid d \in D\}\) are the traces in the canonical quasimodel. By construction of \(\Pi\eta\), we have a rule \(B(x, t + 1) \Leftarrow A(x, t)\). We distinguish cases on \(d\):

- if \(d \in \text{ind}(A)\), we get \(A(d, i) \in \Pi\eta(A)\).
- \(\Rightarrow\) By the mentioned rule, also \(B(d, i + 1) \in \Pi\eta(A)\) and hence \(B \in \pi_d(i + 1)\).
- if \(d = X \in CN\), then for all \(b, n\), if \(X(b, n) \in \Pi\eta(A)\), then \(A(b, i + n) \in \Pi\eta(A)\).
- By the above rule, \(X(b, n) \in \Pi\eta(A)\) also implies \(B(b, i + n + 1) \in \Pi\eta(A)\) for all \(b, n\). Hence, \(B \in \pi_d(i + 1)\).

Properties (q1) and (q2) are clear as \(A \in \Pi\eta(A)\) and (q3) follows directly from formulas (3) and (5) included in \(\Pi\eta\). For (q4), assume that \(A \in \pi_d(n)\), \(A \subseteq \exists r.B \in \mathcal{T}\), and \(E \in \proj_i(\pi_B)(i - n)\) for some \(i \in \mathbb{Z}\). We need to show that \(E \in \pi_d(i)\).

By definition of \(\proj_i\), there is some \(E' \in \pi_B(i - n)\) such that \(\exists r.E' \subseteq E \in \mathcal{T}\). Since \(A \subseteq \exists r.B \in \mathcal{T}\) and (q4) applied to the canonical quasimodel, we get \(\proj_i(\pi_B) \subseteq \pi_A\). In particular, \(E \in \pi_A(j)\) whenever \(E' \in \pi_B\) and thus \(E \in \pi_A(i - n)\). The construction of \(\Pi\eta\), in particular, the rules included for \(\pi_A\) implies the rule \(E(x, t + i - n) \Leftarrow A(x, t)\). We distinguish cases on \(d\):

- if \(d \in \text{ind}(A)\), then \(A(d, n) \in \Pi\eta(A)\).
- \(\Rightarrow\) By the mentioned rule, also \(B(d, i) \in \Pi\eta(A)\) and hence \(B \in \pi_d(i)\).
- if \(d = X \in CN\), then for all \(b, m\), if \(X(b, m) \in \Pi\eta(A)\), then \(A(b, m + n) \in \Pi\eta(A)\).
- By the above rule, \(X(b, m) \in \Pi\eta(A)\) also implies \(E(b, m + i) \in \Pi\eta(A)\), for all \(b, m\).
- Hence, \(B \in \pi_d(i)\).

Since \(A(a, i) \notin \Pi\eta(A)\), we have \(A \notin \pi_d(i)\); thus, \(A \notin \pi_d(i)\) in the canonical quasimodel. Theorem 4 then yields \((\mathcal{T}, A) \not\models A(a, i)\).

(\(\Rightarrow\)) Assume that \((\mathcal{T}, A) \not\models A(a, i)\) and thus \(A \notin \pi_d(i)\) in the canonical quasimodel \(\{\pi_d \mid d \in D\}\). We show that \(A(a, i) \notin \Pi\eta(A)\) by noting that the set \(M\), defined as

\[M = \{B(b, n) \mid B \in \pi_n(b), b \in \text{ind}(A)\} \cup\]
\[\{r(a, b, n) \mid r(a, b, m) \in A, r \in R^\text{reg}\} \cup\]
\[\{r(a, b, n) \mid r(a, b, m) \in A, r \in R^\text{loc}\},\]

clearly is a model for \(\Pi\eta\) and \(A\), thus \(M \supseteq \Pi\eta(A)\), but \(A(a, i) \notin M\).

**5 Restricted Use of Rigid Roles**

**Theorem 6** TAQ answering over \(\mathcal{E}\mathcal{C}\)-Boxes is \(\mathsf{PSPACE}\)-complete in combined and \(\mathsf{PTIME}\)-complete in data complexity.

**Proof.** \(\mathsf{PTIME}\)-hardness in data complexity follows from \(\mathsf{PTIME}\)-hardness for atomic query answering in \(\mathcal{E}\mathcal{C}\). \(\mathsf{SPACE}\)-hardness in combined complexity follows from the (proof of) \(\mathsf{PSPACE}\)-hardness for entailment in Horn-LTL [Chen and Lin, 1993], which, in fact, shows \(\mathsf{PSPACE}\)-hardness of checking whether \(\square \varphi \models p \rightarrow q\), where \(\varphi\) is a conjunction of implications involving only \(\bigcirc\), and \(\land\), and \(p\) and \(q\) are propositional
variables. Now, $\varphi$ can be translated in a straightforward way into a TBox $T_\varphi$ (using only concept and no role names).

For the upper bounds, let $\varphi_{T,A}$ be the Horn-LTL formula constructed in the main part of the paper, and let $\mathfrak{M}$ be its minimal model. We prove correctness of $\varphi_{T,A}$ in the following sense:

**Claim.** For all $a \in \text{ind}(A), i \in \text{tem}(A), A \in \text{CN}$, we have $(\mathfrak{M}, i) \models P_{A,a}$ iff $(T, A) \models A(a, i)$.

**Proof of Claim.** Suppose $(T, A) \not\models A(a, i)$. By Theorem 4, we have $A \not\in \pi_a$. Define a temporal LTL interpretation $\mathfrak{M}'$ by taking, for all $i \in \mathbb{Z}, d \in D, B \in \text{CN}$,

$$(\mathfrak{M}', i) \models P_{B,d} \iff B \in \pi_d(i).$$

It is routine to prove that $\mathfrak{M}'$ is a model of $\varphi_{T,A}$; note that the formulas in that order correspond almost literally to (t1)-(t2) and (q1)-(q4). By construction, we have $(\mathfrak{M}', i) \not\models P_{A,a}$ and thus, by the minimality of $\mathfrak{M}$, we obtain $(\mathfrak{M}, i) \not\models P_{A,a}$.

$(\Rightarrow)$ Suppose $(\mathfrak{M}, i) \not\models P_{A,a}$. We define a quasimodel $\Omega = \{\pi_d | d \in D\}$ with $A \not\in \pi_a$. Theorem 4 will then yield $(T, A) \not\models A(a, i)$. Define $\Omega$ by taking

$$\pi_d(i) = \{B \in \text{CN} | (\mathfrak{M}, i) \models P_{B,d}\} \quad \text{for all } i \in \mathbb{Z}, d \in D.$$ 

It is routine to verify that $\Omega$ is indeed a quasimodel. This proves the result of the Claim. Theorem 6 follows from Claim by the runtime arguments given in the proof.

**Theorem 7** TAQ answering over $\mathcal{TEL}_{\mathfrak{N}}^0$ is PSPACE-complete in data complexity and in EXPTIME in combined complexity.

**Proof.** PSpace-hardness in data complexity is by Theorem 2. For the upper bound, let $\Pi_T^\mathcal{Q}$ be the DATALOG$_{1S}$-program constructed as in the proof of Theorem 5 in Section 4 but where the rules for traces $\pi_B$ in $\Pi_T$ are replaced by the set of rules

$$A(x, t) \leftarrow B(x, t), \quad \text{for all } A \in \pi_B(0).$$

It is routine to verify that $\Pi_T^\mathcal{Q}$ can be used for TAQ answering in the sense of Lemma 15. For showing the complexity upper bounds, denote with $T' \subseteq T$ the subset of $T$ not mentioning rigid roles, and let $\{\pi'_d | d \in D\}$ be the canonical quasimodel for $(T', \emptyset)$. It should be clear that, by the restriction posed on the occurrence of rigid roles, we have $A \in \pi'_B(0)$ iff $A \in \pi_B(0)$. The latter can be decided in PTIME data and PSPACE combined complexity, see Theorem 6. This yields the result since query answering in DATALOG$_{1S}$ is in PSPACE data and EXPTIME combined complexity.

**6 Acyclicity Conditions**

**Theorem 8** TAQ answering over acyclic $\mathcal{TEL}^0$ is in LOGTIME-uniform AC$^0$ in data complexity and in PTIME in combined complexity.

**Proof.** We start with data complexity. It suffices to show that $\varphi_{T,A}(x, t)$ is indeed an FO$(+)$-rewriting of $T, A(x, t)$. Denote with $\mathcal{J}_A$ the ABox $A$ viewed as an interpretation.

**Claim.** For all temporal ABoxes $\mathcal{A}, A \in \text{CN}, a \in \text{ind}(A),$ and $i \in \text{tem}(A), (T, A) \models A(a, i)$ iff $\mathcal{J}_A \models \varphi_{T,A}(a, i)$.

**Proof of the Claim.** The proof is by induction on the ‘depth’ of the expansion of $A$ relative to the acyclic TBox $T$. If $A$ is primitive, that is, has depth 0, the statement is direct. If $A \equiv B_1 \sqcap B_2 \in T$, the statement follows directly from the induction hypothesis. For the other cases, the proof uses the universal model $\mathcal{J}_A$ of $(T, A)$, which is the result of unravelling the canonical quasimodel, and which satisfies the following property for all $a \in \text{ind}(A)$ and $i \in \mathbb{Z}$:

$$(T, A) \models A(a, i) \iff \mathcal{J}_A \models A(a, i). \quad (50)$$

Assume first $A \equiv \exists B. B \in T, r \in N^R$ a rigid role, and $(T, A) \not\models A(a, i)$, and, by (50), also $\mathcal{J}_A \not\models A(a, i)$. Moreover, also by (50), we obtain $\mathcal{J}_A \not\models D(a, i)$ for all concept names $D$ with $T \models D \subseteq A$. Assume, for the sake of contradiction, that $\mathcal{J}_A \models \exists y t'(r(a, y, t') \land \varphi_{T,B}(y, i))$. By the induction hypothesis, $\mathcal{J}_A \models \exists y t'(r(a, y, t') \land B(y, i))$, and hence $\mathcal{J}_A \models (\exists B)(a, i)$, a contradiction. Thus, $\mathcal{J}_A \not\models \varphi_{T,A}(a, i)$, whence $\mathcal{J}_A \not\models \varphi_{T,A}(a, i)$.

For the converse direction, assume that $(T, A) \models A(a, i)$ and thus $\mathcal{J}_A \models A(a, i)$. It is straightforward to show that this is the case iff there is some $D(a, i) \in A$ such that $T \models D \subseteq A$ or there are $b \in \text{ind}(A), j \in \text{tem}(A)$ such that $r(a, b, j) \in A$ and $\mathcal{J}_A \models B(b, i)$, that is, $(T, A) \models B(b, i)$, which, by the induction hypothesis, implies $\mathcal{J}_A \models \varphi_{T,B}(b, i)$. In the former case, we obtain $\mathcal{J}_A \models \varphi_{T,A}(a, i)$ as there is a disjunct for $D$. In the latter case, we have $\mathcal{J}_A \models \exists y t'(r(a, y, t') \land \varphi_{T,B}(y, i))$ and hence $\mathcal{J}_A \models \varphi_{T,A}(a, i)$.

The case when $r \in N^R$ is a local role is almost identical (with the quantifier $\exists$ omitted and $t$ replaced by $t'$). The case $A \equiv \emptyset, B \in T$ can be dealt with in the same way. This finishes the proof of the claim and thus shows that $\varphi_{T,A}$ is indeed an FO$(+)$-rewriting over the infinite models of the form $\mathcal{J}_A$.

For PTIME combined complexity it suffices to note that, by acyclicity of $T$, the traces in the canonical quasimodel $\mathcal{Q} = \{\pi_d | d \in D\}$ have very restricted temporal extension. In particular, it is not hard to verify, see Lemma 17 in the proof of Theorem 10, that for all $a \in \text{ind}(A), B \in \text{CN}$ and $j \in \mathbb{Z}$:

$$\pi_a(j) = \emptyset, \quad \text{if } j < \min A - |T| \text{ or } j > \max A + |T|;$$
\[\pi_B(j) = \emptyset, \quad \text{if } j < -|T| \text{ or } j > |T|.

Thus, the algorithm for constructing the canonical quasimodel terminates in polynomial time.

**DL-acyclicity**

**Theorem 9** TAQ answering over DL-acyclic $\mathcal{TEL}^0$-Boxes of depth $k \geq 1$ is k-EXPSpace-complete in combined complexity and NC$^k$-complete in data complexity.

The following Lemma shows how to characterize traces in the canonical quasimodel by Horn-LTL formulas. The applied technique will be useful in the proof of Theorem 9. We define function $\exp^k(n)$ by taking

$$\exp^0(n) = n, \quad \exp^{k+1}(n) = 2^{\exp^k(n)}$$

(with the superscript $k$ rather than subscript in Theorem 2).
Lemma 16 For all $k \geq 0$ and $B \in \text{CN}$ with $\ell_{\text{ta}}(B) = k$, there is a propositional Horn-LTL formula $\varphi_B$ of size $\exp^k(|T|)$ over variables $P_A$, $A \in \text{CN}$, such that, for all $n \in \mathbb{Z}$,
$$A \in \pi_B(n) \iff \varphi_B \models \bigcirc^n P_A. \quad (51)$$

Proof. We prove it by induction on $k$. Let $\varphi_0$ be a conjunction of $\varphi_{\text{ta}}(F_{P_A_1} \land P_{A_2} \to P_A)$, for $A_1 \land A_2 \subseteq A \in T$, and $\Box (\varphi_{\text{ta}}(P_A) \to P_A)$, for $\Box, B \subseteq A \in T$.

For the basis of induction, $\ell_{\text{ta}}(B) = 0$, observe that, by definition of $\ell_{\text{ta}}$ and (49), only CIs without existential restrictions are relevant for the construction of $\pi_B$. Taking $\varphi_B = B \land \varphi_0$ clearly satisfies (51).

For the induction step, take $B$ with $\ell_{\text{ta}}(B) > 0$ and let $B_1, \ldots, B_k$ be the set of all concept names such that
$$A_i \subseteq \exists r_i.B_i \in T \quad \text{and} \quad A_i \in \pi_B(n), \text{for some } n \in \mathbb{Z}.$$ Since $\ell_{\text{ta}}(B_i) < \ell_{\text{ta}}(B)$, by the induction hypothesis, there are formulas $\varphi_{B_i}$, satisfying (51). Starting from $\varphi_B$, it is not hard to construct formulas $\varphi_{B_i}$ that describe $\varphi_{\text{ta}}(\pi_{A_i})$ (in the sense of (51)). Note that $\varphi_{B_i}(\pi_{A_i})$ is ultimately $p$-periodic for some $p \leq \exp^k(2(|T|))$ and let $m_p, m_{tp, t_p, tp, \ldots}$ be the respective constants. From $\varphi_B$, we obtain a conjunction $\chi_{B_i}$ in the “loop-normal form”: $\bigwedge \varphi_{B_i}$ for $0 \leq j < m_p$ with $\varphi_i = \Box_j \varphi_{D_i} \land \bigwedge \varphi_{D_{tp, t_p, tp, \ldots}}(F_{i}) \land \bigwedge \varphi_{D_{tp, t_p, tp, \ldots}}(F_{i})$ for $0 \leq j < m_{tp}$ and symmetric formulas for $m_{tp, t_p, tp, \ldots}$, cf. the DATALOG$_{GS}$ program in the proof of Theorem 5. Note that this incurs an exponential blow-up. It is routine to show that $\varphi_B = B \land \varphi_0 \land \Box A_i(A_i \to \chi_{B_i})$ satisfies (51).

We continue now with the proof of Theorem 9.

Proof. A closer inspection of the non-elementary lower bound for $\text{TLOG}^0$ in Theorem 2 reveals that the TBoxes used are DL-acyclic. In fact, it is shown that subsumption, and so TBox answering, over TBoxes of depth $k$ is $k$-EXPSPACE-hard. NC$^1$-hardness in data complexity follows from the fact that the word problem of NFAs can be reduced to TBox answering, even without the DL dimension, see [Artale et al., 2015, Theorem 9].

In order to show the matching upper bounds, consider a temporal knowledge base $(T, A)$ with $T$ of depth $k$. We devise a completion procedure for TBox answering. For this purpose, we define sets $A_i$ consisting of all assertions $A(a, i)$ and $r(a, b, i)$ in $A$ and, for $i \in \mathbb{N}^k$, assertions $r(a, b, i)$ such that $r(a, b, j) \in A$ for some $j \in \mathbb{N}$. The algorithm exhaustively adds assertions $A(a, i)$ to $A$ if one of the following holds:
(a) $(T, A_i) \models A(a, i)$;
(b) $r(a, b, i) \in A$ and $B \subseteq \exists r_i.B_i \in T$.

Intuitively, Condition (b) captures the immediate temporal implications of the TBox, while Condition (a) takes into account the role structure of the individuals in $A$. Soundness of the algorithm should be clear. For completeness, it is routine to construct the canonical quasi-mold of $(T, A)$ given the result of the algorithm. The crucial observation is that $A_{\text{min}}(A)$ determines the traces $\pi_a$, $a \in \text{ind}(A)$, in the quasi-model on all $i \leq \text{min} A$ (respectively, $i > \text{max} A$).

It turns out that, in Condition (a), $A_i$ can be replaced with a suitably defined $\text{quotient} B_i$ of $A_i$. Intuitively, the considered query language can only distinguish all possible distinct trees of depth $k$, whose number depends on $|T|$ only; thus, the size of $B_i$ is independent of the size of $A$. For making the notion of “quotient” precise, assume we want to check whether $(T, A_i) \models A(a, i)$, for a concept $A$ with $\ell_{\text{ta}}(A) = k$. Let $A'\subseteq A$ be the unravelling of $A_i$, from point $a$ up to depth $k$ such that for all $0 \leq j \leq k$, and all $b$ in distance $j$ of $a$, we have $B(b, j) \in A'\subseteq A$ only if $\ell_{\text{ta}}(A) = k - j$.

Claim 1. $(T, A_i) \models A(a, i) \iff (T, A'_i) \models A(a, i)$.

Next, let $\sim$ be the smallest equivalence relation on $\text{ind}(A_i)$ such that $b \sim b'$ implies
$$B(b, i) \in A'_i \iff B(b', i) \in A'_i,$$ for all $B$ and $i \in \text{tem}(A_i)$, and
$$B(b, c, i) \in A'_i \iff B(b', c', i) \in A'_i,$$ for each $c \in \text{ind}(A)$ and $i \in \text{tem}(A_i)$ with $r(b, c, i) \in A_i$.

Denote by $[b]$ the equivalence class of $b$ with respect to $\sim$ and let $B_i$ consist of the following assertions:
$$A([b], i), \text{ for } A(b, i) \in A'_i, r([b], [b'], i), \text{ for } r(b, b', i) \in A'_i.$$ Clearly, $\text{ind}(B_i) = \{[b] \mid b \in \text{ind}(A'_i)\}$.

Claim 2. $(T, A'_i) \models A(a, i) \iff (T, B_i) \models A(a, i)$.

Following the steps in the proof of Lemma 16, but taking account of the structure of $B_i$, we can construct a formula $\varphi_{a, i}$, of size $\exp^k(|T|)$ such that $\varphi_{a, i} \models \bigcirc^k P_A$ $(T, B_i) \models A(a, i)$ for all $A \in \text{CN}$ with $\ell_{\text{ta}}(A) = k$. The former can be checked in space polynomial in the size of $\varphi_{a, i}$, which together with a polynomial-time algorithm for checking Condition (b), yields the $k$-EXPSPACE upper combined complexity bound.

Note that the implementation sketched for the $k$-EXPSPACE upper combined complexity bound does not yield an NC$^1$ upper data complexity bound. We explain now how Conditions (a) and (b) can be implemented in NC$^1$. We begin with Condition (a) and give an algorithm that determines $B_i$ “on-the-fly”, which requires some notation. For $0 \leq j \leq k$, let
$$\text{CN}_j = \{A \in \text{CN} \mid \ell_{\text{ta}}(A) = k - j\}.$$ Define equivalence relations $\sim_j$, for all $0 \leq j \leq k$ and $i \in \text{tem}(A)$, similarly to $\sim$ above:
$$b \sim_j b' \iff B(b, i) \in A \iff B(b', i) \in A,$$ for all $B \in \text{CN}_0$;
$$r(b, c, i) \in A \iff r(b', c', i) \in A,$$ for each $c \in \text{ind}(A)$ with $r(b, c, i) \in A$, there is $c' \in \text{ind}(A)$ with $r(b', c', i) \in A$ and $c \sim_{j - 1} c'$, and vice versa.

Denote by $[a]^ j$ the equivalence class of $a$ with respect to $\sim_j$. Note that the maximum number of equivalence classes of $\sim_j$ depends only on $T$ and is independent of $A$. Moreover, each $b, i, j$ give rise to a unique tree-shaped ABox $B_{b, i, j}$ defined by induction on $j$ as follows:
Claim 3.\( A(a,i) \) is added to \( \mathcal{A} \) via a sequence of applications of (b) then \( A(a,i) \) is added to \( \mathcal{A} \) via a sequence of steps (iii).

It remains to argue that this algorithm can be implemented in NC\(^1\). Recall that Step (ii) does not depend on the size of \( \mathcal{A} \) and can thus be implemented in constant time (provided that step (i) has been performed). Step (i) can be implemented in constant time as well, as the number of equivalence classes depends only on \( \mathcal{T} \), and thus, constant: (polynomially many) processor units \( P_{abu} \), for \( (a,b,i) \in \text{ind}(\mathcal{A})^2 \times \text{tem}(\mathcal{A}) \).

For step (iii), it suffices to prove that deciding whether a finite word is a possible prefix for \( L(\mathcal{A}) \) (with a fixed Büchi automaton \( \mathcal{A} \)) is in NC\(^1\). As every regular language is in NC\(^1\) [Straubing, 1994], it suffices to show that the set of all possible prefixes for \( L(\mathcal{A}) \) is regular.

Let \( \mathfrak{A} = (Q, \Sigma, q_0, \Delta, F) \) be a Büchi automaton, where all states are terminating in the sense that, for each \( q \in Q \), there is a word \( \sigma_0 \sigma_1 \cdots \in L(\mathfrak{A}) \) and an accepting run \( q_0 q_1 \cdots \) involving \( q \) (this assumption can be made without loss of generality). Construct an NFA \( \mathfrak{B} = (Q, \Sigma, q_0, \Delta', \mathcal{Q}) \) by taking
\[
\Delta' = \{(q, \tau', q') \mid (q, \tau, q') \in \Delta, \tau' \subseteq \tau\}.
\]

Claim 4. \( L(\mathfrak{B}) \) is the set of all possible prefixes for \( L(\mathfrak{A}) \).

This finishes the description of the algorithm and the analysis of its complexity. \( \square \)

**Temporal Acyclicity**

**Theorem 10** \( \text{TAQ} \) answering over temporally acyclic \( \mathcal{T}E\mathcal{C}^0 \) (with rigid concepts) is \( \text{PTIME-complete} \) in data and combined complexity.

**Proof.** The lower bounds are inherited from \( \mathcal{EL} \). To prove the upper bounds, we establish a small quasimodel property. Let \( K = (\mathcal{T}, \mathcal{A}) \) be a temporal KB with a temporally acyclic TBox \( \mathcal{T} \) and \( \Omega = \{ \tau_d \mid d \in D \} \) its canonical quasimodel. By definition, there is a map \( \ell_0 : \mathcal{C} \to \mathbb{N} \) satisfying conditions (i)–(iii) for temporally acyclic TBoxes. Without loss of generality, we assume that \( \ell_0 \) is minimal in the sense that for all maps \( \ell' \) witnessing temporal acyclicity of \( \mathcal{T} \), we have \( \max_{A \in \mathcal{C}} \ell_0(A) \leq \max_{A \in \mathcal{C}} \ell'(A) \). It follows that \( \min_{A \in \mathcal{C}} \ell_0(A) = 0 \). Let \( n_T = \max_{A \in \mathcal{C}} \ell_0(A) \). By our assumption, \( n_T \leq |\mathcal{T}| \). Denote by \( b \) and \( u \) the numbers \( \min \mathcal{A} \) and \( \max \mathcal{A} \), respectively. We prove that the traces feature a small quasimodel property in the sense that, for every \( d \in D \), there is some \( \sigma_d \subseteq \mathcal{C} \) such that

1. \( \pi_a(j) = \sigma_a \) for all \( j > u + n_T \) or \( j < l - n_T \) and all \( a \in \text{ind}(\mathcal{A}) \);
2. \( \pi_{b,j}(j) = \sigma_B \) for all \( j > n_T \) or \( j < -n_T \) and \( B \in \mathcal{C} \).

Intuitively, \( \sigma_d \) is the set of concepts which contain \( d \) at all time instants, that is, the set of concepts which are ‘rigid’ for \( d \). This property means that the canonical quasimodel has a restricted temporal extension that can change only \( |\mathcal{T}| \) time points beyond the ABox.

We prove Properties (P1) and (P2) by giving an algorithm that constructs a small representation of the canonical quasimodel. The property will then follow from Lemma 17 below, which is part of the completeness proof of the algorithm.

The algorithm is an extension of the quasimodel construction procedure. Besides the finite parts \( \bar{\pi}_d \) of \( \pi_d \), it maintains additional maps \( \sigma_d \), which are initially empty, and for which we have the following completion rules for all \( d \in D \) and all \( j \) with \( l - n_T \leq j \leq u + n_T \):
1. If $A \in \pi_d(j)$ and $A$ is a rigid concept, then $A \in \sigma_d$;
2. $\sigma_d \subseteq \pi_d(j)$;
3. If $A, A' \in \sigma_d$ and $A \cap A' \subseteq B \in T$, then $B \in \sigma_d$;
4. If $A \in \sigma_d$ and $\bigcirc A \subseteq B \in T$, then $B \in \sigma_d$;
5. If $r(a, b, n) \in A$ and $r \in N_{R}^{ig}$, then $proj_{r}(\sigma_a) \subseteq \sigma_a$;\footnote{proj applied to a set of concepts (rather than a trace) is defined in the obvious way.}
6. If $A \in \sigma_d, A \subseteq \exists r.B, r \in N_{R}^{loc}$, then $proj_{r}(\pi_0(B(0))) \subseteq \sigma_d$;
7. If $A \in \sigma_d, A \subseteq \exists r.B, r \in N_{R}^{rig}$, then $proj_{r}(\pi_0(B(j))) \subseteq \sigma_d$;
8. If $A \in \pi_d(j)$, $A \subseteq \exists r.B, r \in N_{R}^{rig}$, then $proj_{r}(\sigma_B) \subseteq \sigma_d$.

Note further that also rules (T1), (T2), (Q1)–(Q4) are applied to $\pi_d$ in the range of $l - n_T \leq j \leq u + n_T$. The algorithm returns “yes” on query $A(a, i)$ if $A \in \pi_d(i)$.

It is routine to verify that the algorithm is sound. For completeness, assume that $A \notin \pi_d(i)$. We define a collection of traces $\pi_d(i)$, $d \in D$, by taking

$$\pi_d(i) = \begin{cases} \pi_d(i), & \text{if } d \in \text{ind}(A) \text{ and } l - n_T \leq i \leq u + n_T, \\ \hat{\pi_d}(i), & \text{if } d \notin \text{CN} \text{ and } 0 \leq \ell_0(d) + i \leq n_T, \\ \sigma_d(i), & \text{otherwise.} \end{cases}$$

and verify that $\Omega = \{\pi_d \mid d \in D\}$ is a quasimodel. This suffices since $A \notin \pi_d(i)$ and Theorem 4 yields $(T, \mathcal{A}) \notin \forall(a, i)$. Although Theorem 4 does not mention rigid concepts, they can be expressed in full $\mathcal{T} \mathcal{E} \mathcal{C}$\footnote{trivially preserves (a)—(c).}. It thus suffices to note that $\Omega$ respects rigid concepts: assume $A \in \pi_d(j)$ and $A$ is a rigid concept name; if $A \in \sigma_d$, then $A \in \pi(i)$ for all $i \in Z$ by rule 2; if $A \notin \sigma_d$, then $A \in \pi_d(i)$ for some $i$. But then, by rule 1, $A \notin \sigma_d$, a contradiction.

For showing that $\Omega$ is a quasimodel, we need to establish the following invariants of the maps $\pi_d$ and $\sigma_d$.

**Lemma 17** Let $CN_k = \{A \in CN \mid \ell_0(A) = k\}$. Then, for every $a \in \text{ind}(A)$, we have

(a) $\hat{\pi}_a(j) \subseteq \bigcup_{k \leq j - u} CN_k \cup \sigma_a$, for all $j$ with $u \leq j \leq u + n_T$;

(b) $\hat{\pi}_a(j) \subseteq \bigcup_{k \leq j - l} CN_k \cup \sigma_a$, for all $j$ with $l - n_T \leq j \leq l$.

For every $B \in CN$ with $\ell_0(B) = m$, we have

(c) $\hat{\pi}_B(j) \subseteq CN_{m+j} \cup \sigma_B$, for all $j$ with $0 \leq m + j \leq n_T$.

**Proof.** We prove these properties by induction on the number of rule applications in the construction of the canonical quasimodel. On empty maps, the three properties are trivially satisfied. Further note that rules 1–8 preserve the invariants since they only extend $\sigma_d$ (except for rule 3, where the claim is also clear). Moreover, it is immediate that any application of (Q1) or (Q2) trivially preserves (a)–(c).

For (Q3), we need to verify only (a) and (b). Assume that $r(a, b, n) \in A$ and some $B \in CN$ is added to $\pi_a(j)$ for some $j$ with $l - n_T \leq j \leq u + n_T$, that is, there is a $B' \in \pi_a(j)$ and $\exists r.B' \in B \in T$. If $r \in N_{R}^{ig}$ is a local role, then $j = n \in \text{tem}(A)$, and so (a) and (b) are trivially satisfied. Otherwise, $r$ is a rigid role. We show (a). By the induction hypothesis, either $B' \in \bigcup_{k \leq l-1} CN_k$ or $B' \in \sigma_a$. In the former case, $B \in A \in \sigma_B$, by the definition of $\ell_0$. In the latter case, we obtain $B \in \sigma_a$ by rule 5; thus, $B \in \pi_d(k)$ for all $k$ with $l - n_T \leq k \leq u + n_T$. Hence, (a) is preserved; (b) is considered similarly.

For (Q4), assume that $A \in \pi_d(n)$, $A \subseteq \exists r.B \in T$, $B' \in \pi_B(i)$ and $\exists r.B' \subseteq A' \in T$. If $r$ is a local role, then we need to consider $i = 0$ only, but then it suffices to note that $\ell_0(A) = \ell_0(B) = \ell_0(B') = \ell_0(A')$. If $r$ is rigid, we distinguish cases:

- If $A \in \sigma_d$, then, by rule 7, we have $A' \in \sigma_d$;
- If $B' \in \sigma_B$, then, by rule 8, we have $A' \in \sigma_d$;
- If neither of the first two cases applies, $d \in \text{ind}(A)$, and $j \geq u (j < l$ and $l \leq j \leq u$ are similar), then, by the induction hypothesis, we have $\ell_0(A) \geq j - u$ and $\ell_0(B') = \ell_0(B) + i = \ell_0(A) + i$. Thus, $\ell_0(A') = \ell_0(B') \geq j + i - u$.
- If either of the first two cases applies and $d \in \text{CN}$, then, by the induction hypothesis, $\ell_0(A) = \ell_0(d) + j$ and $\ell_0(B') = \ell_0(B) + i = \ell_0(A) + i$. Thus, we have $\ell_0(A') = \ell_0(B') = \ell_0(d) + i + j$.

It remains to verify that (T1) and (T2) preserve (a)–(c). We show (a); the other properties are analogous. Let $B \in \pi_d(j)$ and $\bigcirc_B \subseteq A \in T$. By the induction hypothesis, either $B \in \bigcup_{k \leq j-u} CN_k$ or $B \in \sigma_a$. In the former case, we have $\ell_0(B) \geq j - u$, whence, by the definition of $\ell_0$, we get $\ell_0(A) = \ell_0(B) - 1 \geq (j - 1) - u$, and thus $A \in \bigcup_{k \geq j-1-u} CN_k$. In the latter case, we obtain $A \in \sigma_a$ by rule 4. In either case, (a) is preserved. This finishes the proof of Lemma 17.

We are now ready to show that $\Omega$ is a quasimodel. By construction, the $\pi_d$ satisfy (Q1), (Q2), and (T1). For (T2), let $d \in \text{ind}(A)$; the case $d \in \text{CN}$ is analogous. Assume that $A \in \pi_d(j)$ and $\bigcirc_A \subseteq B \in T$. If $l - n_T < j < u + n_T$, then (T2) is satisfied because $\pi_d$ is closed under (T2) on this interval. If $j > u + n_T$ or $j < l - n_T$, then (T2) is due to rule 4. If $j = u + n_T$, then, by Lemma 17, $\ell_0(A) = n_T$, whence $\ell_0(B) < n_T$, and so (T2) holds because $\pi_d$ is closed under (T2).

For (Q3), observe first that it is trivially satisfied in case $r$ is a local role or $l - n_T \leq j \leq u + n_T$. So assume that $r$ is rigid, take $j > u + n_T$ (the case $j < l - n_T$ is symmetric), $B \in \pi_B(j)$, and $\exists r.B \subseteq B' \in T$. By the definition of $\pi_B$, we have $B \in \sigma_B$. By rule 5, $B' \in \sigma_a$, and so $B' \in \pi_a(j)$, by the definition of $\pi_a$.

For (Q4), assume that $A \in \pi_d(j)$, $A \subseteq \exists r.B \in T$, $B' \in \pi_B(i)$ for some $i \in Z$, and $\exists r.B' \subseteq A' \in T$. If $r$ is a local role, then we need to consider $i = 0$ only. If $l - n_T \leq j \leq u + n_T$, then $\pi_d$ is closed under (Q4) because $\pi_d$ is closed under (Q4) on that interval. If $j > u + n_T$
For any $r$, we obtain $A \in \sigma_d$ and, by rule 6, $A' \in \sigma_d$, and hence $A' \in \pi_d(j)$. Finally, consider the following cases where $r$ is rigid.

- If $A \in \sigma_d$, then $A' \in \sigma_d$ by rules 7 and 2.
- If $B' \in \sigma_B$, then, by rule 8, $A' \in \sigma_d$, and so, we obtain $A' \in \pi_d(j + i)$ by rule 2 and the definition of $\pi_d$.
- If neither of the two cases above applies, then, by Lemma 17, $0 \leq \ell_0(B') \leq n_T$ and $\ell_0(B') = \ell_0(B) + i$, and either $d \in \text{Ind}(A)$ and $l - n_T \leq j \leq u + n_T$ or $d \in \text{CN}$ and $0 \leq \ell_0(A) + i \leq n_T$. However, in both cases, (q4) is applied to $\pi_d$ and $\pi_{d'}$, so we are done.

Hence, $\Omega$ is indeed a quasimodel. It remains to note that the algorithm terminates after polynomially many steps: the bounds on the size of the $\pi_d$ and $\sigma_d$ imply that they can be represented in polynomial space, and the completion rules extend these data structures at each step. $\square$

Inflationary $\mathcal{TEL}^\infty$

**Theorem 11** T AQ answering over $\mathcal{TEL}^\infty_{inf}$ is PTIME-complete in both data and combined complexity.

**Proof.** The lower bounds are inherited from $\mathcal{EL}$. To prove the upper bounds, we establish some key properties of traces in the canonical quasimodel over $\mathcal{TEL}^\infty_{inf}$-TBoxes.

Let $(T, A)$ be a KB with a $\mathcal{TEL}^\infty_{inf}$-TBox $T$ and let $\Omega = \{\pi_d \mid d \in D\}$ be its canonical quasimodel. We remind the reader that, by definition, traces for $T$ are closed under (t1) for $\cap$ and (t2) for $\dot\cup$ and $\dot\cap$. We are going to show in Lemma 18 below that the traces $\pi_d$ in the canonical quasimodel are in fact all bow ties.

**Lemma 18** If $T$ is a $\mathcal{TEL}^\infty_{inf}$-TBox, then traces in the canonical quasimodel $\Omega = \{\pi_d \mid d \in D\}$ enjoy the following properties:

(a) $\pi_a$ is a $[\min A, \max A]$-bow tie, for each $a \in \text{ind}(A)$;

(b) $\pi_B$ is a $[0, 0]$-bow tie, for each $B \in \text{CN}$.

**Proof.** We show by induction on the number of rule applications that the traces in the quasimodel construction procedure satisfy (a) and (b). The basis of induction is immediate from (q1) and (q2). For the inductive step, we make a case distinction on which rule is applied. For (t1) and (t2), we assume that they are always applied exhaustively. Since $\dot\cap$ are the only temporal operators allowed in $T$, it can be readily verified, that

Claim. For any $[l, u]$-bow tie $\varrho$, the closure of $\varrho$ under (t1) and (t2) is an $[l, u]$-bow tie as well.

For (q3), let $r(a, b, n) \in A$. Property (b) is trivial and (a) is preserved because $\text{proj}_a(\pi_b)$ is an $[l, u]$-bow tie for any $[l, u]$-bow tie $\pi_b$ and the component-wise union of two $[l, u]$-bow ties is also an $[l, u]$-bow tie.

For (q4), let $A \in \pi_d(n)$ and $A \subseteq \exists r. B \in T$. By the induction hypothesis, $\pi_d$ is an $[l, u]$-bow tie for some $l, u$. Also observe that, by the induction hypothesis, $\pi_B$ is a $[0, 0]$-bow tie and therefore, for any $m \in \mathbb{Z}$, the map $\varrho_m : i \rightarrow \text{proj}_i(\pi_B)(i - m)$ is an $[m, m]$-bow tie. We consider the following three cases.

1. If $l \leq n \leq u$, then the application of (q4) extends $\pi_d$ with $\varrho_m$, which results in an $[l, u]$-bow tie, as required for (a) and (b).

2. If $n > u$, let $m_0 > u$ be minimal with $A \in \pi_d(m_0)$. Observe that (q4) is applicable to all $m \geq m_0$ and all such applications of (q4) can be captured by taking the union of $\pi_d$ and all $\varrho_m$, for $m \geq m_0$. Observe that $\bigcup_{m \geq m_0} \varrho_m$ is in fact $\text{proj}_i(\pi)$ for $\pi$ defined by taking

$$
\pi(n) = \begin{cases} X, & \text{if } n < m_0, \\
X \cup \pi_B(0) \cup \pi_B(n - m_0), & \text{if } n \geq m_0,
\end{cases}
$$

where $X$ be the maximum set at ‘the left end’ of $\pi_B$: $X = \bigcup_{n<0} \pi_B(n)$. Since $\pi_B$ is a $[0, 0]$-bow tie, the map $\pi$ is monotone, that is, $\pi(n) \subseteq \pi(n + 1)$ for all $n \in \mathbb{Z}$, and $\pi(n) = \pi(n - 1)$ for all $n < m_0$. It follows that also $\text{proj}_i(\pi)$ is monotone and $\text{proj}_i(\pi)(n) = \text{proj}_i(\pi)(n - 1)$ for all $n < m_0$. It is then easy to see that the component-wise union of $\pi_d$ and $\text{proj}_i(\pi)$ is an $[l, u]$-bow tie.

3. If $n < l$, then the construction is the mirror image of 2.

This completes the proof of Lemma 18. $\square$

We are now in a position to complete the proof of Theorem 11. First, observe that the traces maintained by the procedure for constructing the canonical quasimodel of a KBs with a $\mathcal{TEL}^\infty_{inf}$-TBox can be represented using a polynomial amount of space because, by Lemma 18, they are $[l, u]$-bow ties with both $l, u \in \text{tem}(A)$: one needs to store only the middle section and at most $\text{CN}$ steps at both ends. Second, the traces are extended with every rule application, and so, the procedure must terminate after polynomially many steps, which gives PTIME upper complexity bound for both data and combined complexity. $\square$

**References**
