Metric Temporal Logic for Ontology-Based Data Access over Log Data

D. Calvanese, E. Güzel Kalaycı, V. Ryzhikov, G. Xiao and **M. Zakharyaschev**

Faculty of Computer Science
Free University of Bozen-Bolzano, Italy
{calvanese,kalayci,ryzhikov,xiao}@inf.unibz.it

Department of Computer Science
Birkbeck, University of London, UK
michael@dcs.bbk.ac.uk

Abstract

We present a new metric temporal logic *HornMTL* over dense time and its datalog extension *datalogMTL*. The use of *datalogMTL* is demonstrated in the context of ontology-based data access over meteorological data. We show decidability of answering ontology-mediated queries for a practically relevant non-recursive fragment of *datalogMTL*. Finally, we discuss directions of the future work, including the potential use-cases in analyzing log data of engines and devices.

Introduction

The aim of ontology-based data access (OBDA) (Poggi et al. 2008) is, on one hand, to represent the information from various heterogeneous data sources in a unified and conceptually transparent way by means of *mappings*. On the other hand, the *ontology language* allows one to define concepts in terms of other concepts, and thereby represent frequently used query patterns as reusable concepts. The end-user, in that case, can obtain the required information by means of simple conceptual queries and is not required to know neither the structure of the source data nor the definitions of the concepts he is using.

Due to up-to-date requirements of industry (see, e.g., (Kharlamov et al. 2014)) the OBDA approach is being actively adopted in the context of the temporal data of streams and logs. Initially, only the classical non-temporal ontology languages were considered to mediate the access to temporal data (Gutiérrez-Basulto and Klarman 2012; Özcep et al. 2013; Baader, Borgwardt, and Lippmann 2013; Klarman and Meyer 2014). Later, the ontology languages with temporalized concepts were studied in this context (Artale et al. 2015; Kontchakov et al. 2016; Basulto, Jung, and Kontchakov 2016). Such concepts are defined by means of linear temporal logic (LTL); for example, the axiom

$$\mathsf{Hurricane} \leftarrow \mathsf{HurricaneForceWind} \land \mathbf{X}^{-}\mathsf{HurricaneForceWind}$$

defines a hurricane as hurricane force wind lasting for 1 hour ($\mathbf{X}^{-}$ is the *previous time* LTL operator). One easily notices that this definition works only if the temporal data arrives strictly in hourly periods, such as 13:21, 14:21, etc. If these periods are smaller and have a fixed length, the definition above can still be adjusted by using the conjunction of the form $\mathsf{HFW} \land \mathbf{X}^{-}\mathsf{HFW} \land \mathbf{X}^{-}\mathbf{X}^{-}\mathsf{HFW} \land \dots$. However, first, having the data with fixed-period timestamps is not always a realistic assumption, and, second, doing the adjustment above contradicts the OBDA philosophy, where the ontology user is not required to have knowledge of the structure of the data sources. Therefore, the following definition would be more natural

$$\mathsf{Hurricane} \leftarrow \boxminus_{\geqslant 0}^{\leqslant 1h}\mathsf{HurricaneForceWind},$$

where $\boxminus_{\geqslant 0}^{\leqslant 1h}$ is a metric temporal operator *during the previous hour*. The logic required to express such statements is a kind of *metric temporal logic* or *modal logic of metric spaces*; see (Koymans 1990; Kurucz, Wolter, and Zakharyaschev 2005) for surveys and further references.

In this paper, we introduce a metric temporal logic *HornMTL* with the operator $\boxminus_{\triangleright e}^{\triangleleft d}$, where $\triangleright$ is either $>$ or $\geqslant$ (and similarly for $\triangleleft$) and e, d are time *distances*, its future analogue $\boxplus_{\triangleright e}^{\triangleleft d}$, as well as their duals $\oplus_{\triangleright e}^{\triangleleft d}$ and $\ominus_{\triangleright e}^{\triangleleft d}$. We interpret this logic over a *dense* temporal domain. The reason for not considering a *discrete* domain is that we want to abstract from the *granularities* of time (periods of timestamps) in the data sources. In our logic, we allow the statements of the form $P@\iota$, where ι is an interval specified by a pair of time instants, to represent the conceptualized temporal data. The meaning of, say, $P@(t_1, t_2]$ is that P holds at all times t between t_1 (not including it) and t_2 (including it). We assume that we can convert data from any source with timestamped tuples to this format by means of mappings. For example, if a source contains the information of temperature measurements taken every hour, such as 13:21: -1°C, 14:21: 2°C, 15:21: -1°C, etc., we can conceptualize them as the statements PositiveTemp@(13:21, 14:21], etc. Note that whether to include the ends of intervals or not, as well as whether to consider 2°C to be the case in the hour preceding or following 14:21, is the choice of the mapping designer. We then extend *HornMTL* to *datalogMTL* that also allows for standard Datalog reasoning about objects of the application domain (weather stations, cities, sensors, etc.).

We present a few preliminary results on *datalogMTL*. First, we describe a use-case of OBDA over meteorological data with SQL mappings to a large real-world weather database and *datalogMTL* as an ontology language. Second, we develop an ontology-mediated query answering algorithm for a non-recursive fragment $\mathit{datalogMTL}_{nr}^{\Box}$ of *datalogMTL*. Finally, we report some preliminary evaluation results showing the feasibility of our approach.

HornMTL and *datalogMTL*

Syntax. We consider a propositional temporal logic *HornMTL* with the set of propositional variables $P_0, P_1, \dots$ over the temporal domain $\mathfrak{T}$ isomorphic to $(\mathbb{R}, \leqslant)$ with 0 and arithmetic operations $+, -$. That is, we assume *dense* time. Let $\mathsf{int}(\mathfrak{T})$ be the set of (non-empty) *intervals on* $\mathfrak{T}$, which are of the form $[t_1, t_2]$, $[t_1, t_2)$, $(t_1, t_2]$, and (t_1, t_2), where $t_i \in \mathfrak{T} \cup \{-\infty, \infty\}$, $\langle$ is either (or [, and $\rangle$ is either) or]. (We do not distinguish between the intervals $\langle t_1, \infty]$ and $\langle t_1, \infty)$, consider $\langle \infty, \infty \rangle$ to be empty, and analogously for $-\infty$. We also assume that $\leq$ is defined on $\mathfrak{T} \cup \{-\infty, \infty\}$ and $+, -$ are defined on pairs of elements from $\mathfrak{T}$ and $\{-\infty, \infty\}$, in a standard way.) Define a *data instance* $\mathcal{D}$ as a non-empty finite set of *data assertions* (or *facts*) of the form:

$$P_i@\iota,$$

where P_i is a propositional variable and $\iota \in \mathsf{int}(\mathfrak{T})$.

We use the temporal operators of the form:

- $\boxplus^{\lhd d}_{\rhd e}$ (always between e and d in the future),
- $\boxminus^{\lhd d}_{\rhd e}$ (always between e and d in the past),
- $\diamondplus^{\lhd d}_{\rhd e}$ (sometime between e and d in the future),
- $\diamondminus^{\lhd d}_{\rhd e}$ (sometime between e and d in the past),

where $\lhd$ is either $<$ or $\leqslant$, e, d are *distances*, that is, *positive* elements of $\mathfrak{T}$, and $\rhd$ is either $>$ or $\geqslant$. Thus, e.g., $\boxplus^{<d}_{\geqslant e}$ expresses 'always between e and d in the future including e and excluding d' and similarly for the other operators. We also impose the following consistency requirement on every operator $\mathbf{O}^{\lhd d}_{\rhd e}$ (henceforth we assume $\mathbf{O} \in \{\boxplus, \boxminus, \diamondplus, \diamondminus\}$, $\Box \in \{\boxplus, \boxminus\}$, and $\Diamond \in \{\diamondplus, \diamondminus\}$):

- there exists $t \in \mathfrak{T}$ such that $t \rhd e$ and $t \lhd d$.

Propositional *literals* are defined by the following grammar:

$$\lambda \ ::= \ P_i \ \mid \ \mathbf{O}^{\lhd d}_{\rhd e}\lambda.$$

An *ontology*, $\mathcal{O}$, is a finite set of *axioms* of the form:

$$\lambda \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k, \qquad \bot \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k. \tag{1}$$

A *knowledge base* (KB) is a pair $(\mathcal{O}, \mathcal{D})$.

Semantics. Consider an interpretation $\mathfrak{M} = (\mathfrak{T}, \cdot^{\mathfrak{M}})$ such that $P_i^{\mathfrak{M}} \subseteq \mathfrak{T}$ for each propositional variable P_i and write $\mathfrak{M}, t \models P_i$ when $t \in P_i^{\mathfrak{M}}$ for $t \in \mathfrak{T}$. As usual, it is assumed that $\mathfrak{M}, t \not\models \bot$ for all $t \in \mathfrak{T}$. We extend the definition of $\models$ to λ as follows:

$$\mathfrak{M}, t \models \boxplus^{\lhd d}_{\rhd e}\lambda \quad \text{iff} \quad \mathfrak{M}, t' \models \lambda \text{ for all } t' \text{ such that } t' - t \rhd e \text{ and } t' - t \lhd d, \tag{2}$$

$$\mathfrak{M}, t \models \boxminus^{\lhd d}_{\rhd e}\lambda \quad \text{iff} \quad \mathfrak{M}, t' \models \lambda \text{ for all } t' \text{ such that } t - t' \rhd e \text{ and } t - t' \lhd d, \tag{3}$$

$$\mathfrak{M}, t \models \diamondplus^{\lhd d}_{\rhd e}\lambda \quad \text{iff} \quad \mathfrak{M}, t' \models \lambda \text{ for some } t' \text{ such that } t' - t \rhd e \text{ and } t' - t \lhd d, \tag{4}$$

$$\mathfrak{M}, t \models \diamondminus^{\lhd d}_{\rhd e}\lambda \quad \text{iff} \quad \mathfrak{M}, t' \models \lambda \text{ for some } t' \text{ such that } t - t' \rhd e \text{ and } t - t' \lhd d. \tag{5}$$

We say that $\mathfrak{M}$ satisfies a data assertion $P@\iota$ if $\mathfrak{M}, t \models P$ for *all* $t \in \iota$. We say that $\mathfrak{M}$ satisfies an ontology axiom $\lambda \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k$ (respectively, $\bot \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k$), if $\mathfrak{M}, t \models \lambda_i$, for all $i = 1, \dots, k$, imply $\mathfrak{M}, t \models \lambda$ (resp., $\mathfrak{M}, t \models \bot$), for *every* $t \in \mathfrak{T}$. Thus, the ontology axioms are *global*. We say that $\mathfrak{M}$ satisfies a data instance $\mathcal{D}$ (resp., ontology $\mathcal{O}$) if it satisfies each statement in it. Finally, we say that $\mathfrak{M}$ satisfies a knowledge base $(\mathcal{O}, \mathcal{D})$ and write $\mathfrak{M} \models (\mathcal{O}, \mathcal{D})$ if $\mathfrak{M}$ satisfies both $\mathcal{O}$ and $\mathcal{D}$.

Our main reasoning problem is *query answering*. Define an *atomic query* (AQ) as an expression $P@\delta$, where P is a proposition and δ is an *interval variable*. An ontology $\mathcal{O}$ and an AQ $P@\delta$ constitute an *ontology-mediated query* (OMQ) $\boldsymbol{Q}(\delta) = (\mathcal{O}, P@\delta)$. A *certain answer* to $\boldsymbol{Q}(\delta)$ over $\mathcal{D}$ is any interval $\iota \in \mathsf{int}(\mathfrak{T})$ such that $\mathfrak{M} \models (\mathcal{O}, \mathcal{D})$ implies $\mathfrak{M}, t \models P$ for all $t \in \iota$.

***HornMTL*$^\Box$ fragment.** We consider one important fragment *HornMTL*$^\Box$ of *HornMTL*, where the operators $\diamondplus^{\lhd d}_{\rhd e}$ and $\diamondminus^{\lhd d}_{\rhd e}$ are disallowed in the *heads* of the rules. Note that each *HornMTL*$^\Box$ KB can be converted to KB that has $\boxplus^{\lhd d}_{\rhd e}$ and $\boxminus^{\lhd d}_{\rhd e}$ operators only, and the original KB is a conservative extension of it. For example, an axiom $R \leftarrow P \wedge \diamondminus^{\lhd d}_{\rhd e}Q$ can be replaced by the pair of axioms $R \leftarrow P \wedge Q'$ and $\boxplus^{\lhd d}_{\rhd e}Q' \leftarrow Q$. Finally, we consider a *non-recursive* fragment *HornMTL*$^\Box_{nr}$ of *HornMTL*$^\Box$ by adopting the simplest definition of non-recursivivity: consider the relation $\prec$ on the symbols of $\mathcal{O}$ defined as $P \prec Q$ iff there is an axiom in $\mathcal{O}$, where P occurs in the head and Q in the body (P *depends on* Q). We require that $P \prec^* P$ for no symbol P in $\mathcal{O}$, where $\prec^*$ is a transitive closure of $\prec$.

***datalogMTL*.** Consider the predicate symbols $P_0, P_1, \dots$, each of some arity $m \geq 0$, and a set of object variables $x_0, x_1, \dots$. Data instances $\mathcal{D}$ here contain assertions $P(\boldsymbol{c})@\iota$, where P is an m-ary predicate symbol, $\boldsymbol{c}$ an m-tuple of individual constants, and $\iota \in \mathsf{int}(\mathfrak{T})$. This assertion says that $P(\boldsymbol{c})$ is true at ι. We denote by $\mathsf{ind}(\mathcal{D})$ the set of all individual constants in $\mathcal{D}$. An ontology $\mathcal{O}$ is a finite set of axioms of the form (1) with the literals λ defined by the grammar:

$$\lambda \ ::= (\tau \neq \tau') \ \mid \ (\tau = \tau') \ \mid \ P(\boldsymbol{x}) \ \mid \ \mathbf{O}^{\lhd d}_{\rhd e}\lambda,$$

where P is a predicate symbol of arity m, $\boldsymbol{x}$ is a vector of m variables, and τ, τ' are individual terms, i.e., variables or constants. We also impose other standard datalog restrictions on our programs, and forbid (in)equality predicates in the heads. We call the predicates occurring in $\mathcal{D}$ *extensional* and those occurring in the head of the axioms of $\mathcal{O}$ *intentional*. An *interpretation*, $\mathfrak{M}$, is based on the domain $\Delta = \mathsf{ind}(\mathcal{D})$ (for the individual variables and constants) and $\mathfrak{T}$. For any m-ary predicate P, m-tuple $\boldsymbol{c}$ from Δ and $t \in \mathfrak{T}$, $\mathfrak{M}$ specifies whether P is *true on* $\boldsymbol{c}$ *at* t, in which case we write $\mathfrak{M}, t \models P(\boldsymbol{c})$. Let ν be an *assignment* of elements of Δ to individual terms (we adopt the standard name assumption: $\nu(c) = c$, for every individual constant c). We set:

$$\mathfrak{M}, t \models^{\nu} \tau \neq \tau' \text{ iff } \nu(x_0) \neq \nu(x_1),$$
$$\mathfrak{M}, t \models^{\nu} \tau = \tau' \text{ iff } \nu(x_0) = \nu(x_1),$$
$$\mathfrak{M}, t \models^{\nu} P(\boldsymbol{x}) \text{ iff } \mathfrak{M}, t \models^{\nu} P(\nu(\boldsymbol{x})),$$

and use inductively the formulas (2)–(5) with $\models^{\nu}$ instead of $\models$ for the cases $\mathbf{O}^{\lhd d}_{\rhd e}\lambda$. We say $\mathfrak{M}$ satisfies an ontology axiom $\lambda \leftarrow \lambda_1 \wedge \cdots \wedge \lambda_k$ (respectively, $\perp \leftarrow \lambda_1 \wedge \cdots \wedge \lambda_k$), if $\mathfrak{M}, t \models^{\nu} \lambda_i$ for each i implies $\mathfrak{M}, t \models^{\nu} \lambda$ (resp., $\mathfrak{M}, t \models^{\nu} \perp$), for *every* $t \in \mathfrak{T}$ and assignment ν. Finally, $\mathfrak{M}$ satisfies a data assertion $P(\boldsymbol{c})@\iota$ if $\mathfrak{M}, t \models P(\boldsymbol{c})$ for each $t \in \iota$, and $\mathfrak{M} \models (\mathcal{O}, \mathcal{D})$ is defined straightforwardly.

AQs are defined as $P(\boldsymbol{x})@\delta$, where P is a predicate symbol of arity m, and δ is an interval variable. An ontology-mediated query is defined $\boldsymbol{Q}(\boldsymbol{x}, \delta) = (\mathcal{O}, P(\boldsymbol{x})@\delta)$. A *certain answer* to $\boldsymbol{Q}(\boldsymbol{x}, \delta)$ over $\mathcal{D}$ is any pair $(\boldsymbol{c}, \iota)$, such that $\boldsymbol{c} = \nu(\boldsymbol{x})$ for some ν, and $\mathfrak{M} \models (\mathcal{O}, \mathcal{D})$ implies $\mathfrak{M}, t \models P(\boldsymbol{c})$ for all $t \in \iota$.

Note that *HornMTL* is a fragment of *datalogMTL* (where all predicates have arity 0). We also consider the fragments *datalogMTL*$^{\Box}$ and *datalogMTL*$^{\Box}_{nr}$ defined with the same syntactic restrictions as *HornMTL*$^{\Box}$ and *HornMTL*$^{\Box}_{nr}$.

Weather Use Case

Our OBDA approach can be used to analyze meteorological data through ontology-mediated queries. The MesoWest[1] project makes publicly available historical records of the weather stations across the US. This data is available in the relational tables $\mathsf{Weather}$ containing the following fields:

ID. Station ID. Example: KHYS.

TIME. Timestamp. Example: 11-11-2015 8:55 CST.

TMP. Temperature. Example: 15.6° C.

SKNT. Wind Speed. Example: 9.2 km/h.

P01I. Precipitation in 1 hour. Example: 0.09 cm.

Moreover, there are metadata tables $\mathsf{Metadata}$ containing, in particular, location information of stations in the fields:

ID. Station ID. Example: KHYS.

COUNTY. Example: Ellis.

STATE. Example: Kansas.

We can conceptualize this raw data by means of the SQL mappings. For example, to extract the data for the extensional predicate $\mathsf{Precipitation}(x)@\langle t_1, t_2\rangle$ (with the meaning precipitation occurs at x during $\langle t_1, t_2\rangle$), we can use the following SQL query:

```
SELECT ID AS x,
  lag(TIME) over (partition
       by ID order by TIME) AS t1,
  TIME AS t2, "(" AS ⟨, "]" AS ⟩
FROM Weather
WHERE P01I > lag(P01l)
  over(partition by ID order by TIME)
```

[1] http://mesowest.utah.edu/

That is, we extract the intervals of the shape $(t_1, t_2]$, where t_1 and t_2 are the two *next* timestamps for a given station. The ends of the interval are chosen to reflect the fact that, e.g., the precipitation is measured *accumulatively* and the device produces the output in the end of the measurement interval. Analogously to $\mathsf{Precipitation}$, we populate by the data the other extensional predicates, such as $\mathsf{PositiveTemp}$ (temperature well above 0° C), $\mathsf{HurricaneForceWind}$ (wind with the speed above 118 km/h), $\mathsf{TempAbove24}$ and $\mathsf{TempAbove41}$ (temperature above 24 and 41° C).

Consider the ontology containing the axioms:

$$\mathsf{Rain}(x) \leftarrow \mathsf{PositiveTemp}(x) \wedge \mathsf{Precipitation}(x),$$
$$\boxminus^{\leqslant 1h}_{\geqslant 0}\mathsf{Hurricane}(x) \leftarrow \boxminus^{\leqslant 1h}_{\geqslant 0}\mathsf{HurricaneForceWind}(x),$$
$$\boxminus^{\leqslant 24h}_{\geqslant 0}\mathsf{ExcessiveHeat}(x) \leftarrow \boxminus^{\leqslant 24h}_{\geqslant 0}\mathsf{TempAbove24}(x) \wedge \diamondminus^{\leqslant 24h}_{\geqslant 0}\mathsf{TempAbove41}(x),$$

The second axiom is already discussed in the introduction (here we use a slightly modified version to say that hurricane holds also at the time point, when the hurricane force wind begins), whereas the last axiom formalizes the definition of the situation when an excessive heat warning should be issued according to the US Weather Forecast Offices (24 hours with the minimal temperature above 24° C and the maximal above 41° C).

We can also populate the binary predicate $\mathsf{LocationOf}(x, y)@\langle t_1, t_2\rangle$ by using:

```
SELECT COUNTY AS x, ID AS y,
  −∞ AS t1, ∞ AS t2, "(" AS ⟨, ")" AS ⟩
FROM Metadata
```

Note that we assume that $\mathsf{LocationOf}$ holds between a county and a station *globally*. It is now possible to define:

$$\mathsf{HurricaneAffectedCounty}(x) \leftarrow \mathsf{LocationOf}(x, y) \wedge \mathsf{Hurricane}(y),$$
$$\mathsf{SpreadRainCounty}(x) \leftarrow \mathsf{LocationOf}(x, y) \wedge \mathsf{LocationOf}(x, z) \wedge (y \neq z) \wedge \mathsf{Rain}(y) \wedge \mathsf{Rain}(z).$$

Query Answering in *datalogMTL*$^{\Box}_{nr}$

In this section we first present an algorithm for computing certain answers to an *HornMTL*$^{\Box}_{nr}$ OMQ $\boldsymbol{Q}(\delta) = (\mathcal{O}, P@\delta)$ over $\mathcal{D}$.

Normal form for *HornMTL*$^{\Box}_{nr}$. Our procedure works on the ontology $\mathcal{O}$ containing only the clauses of the shape:

$$P \leftarrow Q \wedge R, \quad \perp \leftarrow Q \wedge R,$$
$$\boxplus^{\lhd d}_{\rhd e} P \leftarrow Q, \quad \boxminus^{\lhd d}_{\rhd e} P \leftarrow Q,$$
$$P \leftarrow \boxplus^{\lhd d}_{\rhd e} Q, \quad P \leftarrow \boxminus^{\lhd d}_{\rhd e} Q$$

It is an easy exercise to verify that every *HornMTL*$^{\Box}_{nr}$ can be brought to the normal form by performing the following operations:

– Substitute the axioms of the shape $\lambda \leftarrow \lambda_1 \wedge \cdots \wedge \lambda_k$ for $k \geq 3$ by $k - 1$ axioms with binary conjunctions using fresh symbols. Analogously for the axioms with $\perp$ in the head.

– Remove $\Diamond_{\rhd e}^{\lhd d}\lambda$ literals in the body of the axioms as sketched in Preliminaries.

– Remove the nested modalities $\boxminus_{\rhd e}^{\lhd d}\lambda$ by substituting them for $\boxminus_{\rhd e}^{\lhd d}P_\lambda$, for a fresh symbols P_λ, and adding:

  – $P_\lambda \leftarrow \lambda$, if $\boxminus_{\rhd e}^{\lhd d}\lambda$ occurred in the body of the axiom,

  – $\lambda \leftarrow P_\lambda$, if $\boxminus_{\rhd e}^{\lhd d}\lambda$ occurred in the head of the axiom.

– Remove the axioms of the shape $\lambda_0 \leftarrow \lambda_1 \wedge \lambda_2$, if $\lambda_i = \boxminus_{\rhd e}^{\lhd d}P$ for some $0 \leq i \leq 2$, as described in the previous step. Analogously for the axioms with $\bot$ in the head.

It can be readily verified that the resulting ontology in the normal form is in *HornMTL*$^{\boxminus}_{nr}$.

Algorithm. We first assume that the facts of $\mathcal{D}$ are stored in the tables of the shape $P_i^*(t_1, t_2, \langle, \rangle)$, where $t_1, t_2 \in \mathfrak{T}$, $\langle$ is either (or [, and $\rangle$ is either) or]. E.g., for $\mathcal{D} = \{P_i@(t_1, t_2], P_i@[t_1', t_2']\}$ we produce the table P_i^* with two tuples $\{(t_1, t_2, (,]), (t_1', t_2', [,])\}$. Consider an intentional symbol P and assume that for all Q such that $P \prec Q$ the tables Q^* are computed. Consider now the cases:

$\boldsymbol{P \leftarrow Q \wedge R}$**.** Then P^* is computed as the *minimal* table satisfying the condition:

$$Q^*\big(t_1,t_2,\langle,\rangle\big) \wedge R^*\big(t_1', t_2', \langle', \rangle'\big) \wedge \\ \mathsf{ints}\big(t_1, t_2, \langle, \rangle, t_1', t_2', \langle', \rangle'\big) \to P^*\big(t_1'', t_2'', \langle'', \rangle''\big),$$

where $\mathsf{ints}(t_1, t_2, \langle, \rangle, t_1', t_2', \langle', \rangle')$ is $\top$ if $\langle t_1, t_2\rangle \cap \langle' t_1', t_2'\rangle' \neq \emptyset$ (the intervals intersect), otherwise it is $\bot$, and $\langle'' t_1'', t_2''\rangle'' = \langle t_1, t_2\rangle \cap \langle' t_1', t_2'\rangle'$ (the result of the intersection). Note that P^* is computed as a temporal join (Gao et al. 2005) of Q^* and R^*. We also create a table $\bot^*$ for the axioms $\bot \leftarrow Q \wedge R$.

$\boldsymbol{\boxplus_{\rhd e}^{\lhd d} P \leftarrow Q}$**.** Then P^* is computed as a minimal table satisfying:

$$Q^*\big(t_1, t_2, \langle, \rangle\big) \to P^*\big(t_1 + e, t_2 + d, \mathsf{ed}\big(\langle, \rhd\big), \mathsf{ed}\big(\rangle, \lhd\big)\big)$$

where the *edge function* $\mathsf{ed}(\langle, \rhd)$ returns [, if $\langle$ is [and $\rhd$ is $\geqslant$, and (, otherwise. Then $\mathsf{ed}(\rangle, \lhd)$ is defined symmetrically. For example, if $Q^* = \{(t_1, t_2, (,])\}$ and the axiom is $\boxplus_{\geqslant e}^{\leq d} P \leftarrow Q$, then $P^* = \{(t_1 + e, t_2 + d, (,))\}$. The axiom $\boxminus_{\rhd e}^{\lhd d} P \leftarrow Q$ is handled analogously.

$\boldsymbol{P \leftarrow \boxplus_{\rhd e}^{\lhd d} Q}$**.** Consider the following example: let $Q^* = \{(t_1, t_2, (,]), (t_2, t_3, (,))\}$ and the axiom $P \leftarrow \boxminus_{\geqslant e}^{\leq d} Q$ such that $d - e < t_3 - t_1$. Then, according to the semantics, $P^* = \{(t_1 - e, t_3 - d, (,])\}$. In order to compute P^* correctly we need to consider the *concatenation* of the intervals $(t_1, t_2]$ and (t_2, t_3). To compute P^* in general we first produce a *closure* Q' of Q^* as the minimal table satisfying:

$$Q^*\big(t_1, t_2, \langle, \rangle\big) \to Q'\big(t_1, t_2, \langle, \rangle\big),$$

$$Q^*\big(t_1, t_2, \langle, \rangle\big) \wedge Q'\big(t_1', t_2', \langle', \rangle'\big) \wedge (t_2' \leq t_2) \wedge \\ \mathsf{ints}\big(t_1, t_2, \langle, \rangle, t_1', t_2', \langle', \rangle'\big) \to Q'\big(t_1', t_2, \langle', \rangle\big).$$

After that P^* can be obtained by:

$$Q'\big(t_1, t_2, \langle, \rangle\big) \wedge \mathsf{fit}\big(t_1, t_2, \langle, \rangle, e, d, \rhd, \lhd\big) \to \\ P^*\big(t_1 - e, t_2 - d, \mathsf{de}\big(\langle, \rhd\big), \mathsf{de}\big(\rangle, \lhd\big)\big),$$

where $\mathsf{fit}\big(t_1, t_2, \langle, \rangle, e, d, \rhd, \lhd\big)$ is $\top$, if there exists $t \in \mathfrak{T}$ such that $\{t + t' \mid t' \rhd e \text{ and } t' \lhd d\} \subseteq \langle t_1, t_2\rangle$, and $\bot$ otherwise. Essentially, fit holds if the segment $\{t' \mid t' \rhd e \text{ and } t' \lhd d\}$ can be shifted so that it *fits* inside $\langle t_1, t_2\rangle$. Finally, another edge function de is needed to compute the ends of the resulting interval. Here $\mathsf{de}\big(\langle, \rhd\big)$ is [, if either $\langle$ is (and $\rhd$ is $>$, or $\langle$ is [and $\rhd$ is $\geqslant$; otherwise $\mathsf{de}\big(\langle, \rhd\big)$ is (. The definition of $\mathsf{de}\big(\rangle, \lhd\big)$ is symmetric. The axiom $\boxminus_{\rhd e}^{\lhd d} P \leftarrow Q$ is handled analogously. Observe that the computation of Q' requires *recursion*.

Clearly, when P occurs in the head of several axioms, the table P^* is taken equal to the union of the tables computed above. In fact, for every symbol P in $\mathcal{O}$ the algorithm computes P^* that, for a consistent KB $(\mathcal{O}, \mathcal{D})$, satisfies:

- for every $t \in \mathfrak{T}$, there exists a certain answer ι to OMQ $\boldsymbol{Q}(\delta) = (\mathcal{O}, P@\delta)$ over $\mathcal{D}$ such that $t \in \iota$ iff there exists a tuple $\big(t_1, t_2, \langle, \rangle\big)$ in P^* such that $t \in \langle t_1, t_2\rangle$.

This correctness follows directly from the semantics of *HornMTL*$^{\boxminus}_{nr}$. Then, if the table $\bot^*$ is empty, as an output of the OMQ $\boldsymbol{Q}(\delta) = (\mathcal{O}, G@\delta)$ over $\mathcal{D}$ we produce the table G^* (otherwise, we return G^* with one special tuple $\big(-\infty, \infty, (,)\big)$ as $(\mathcal{O}, \mathcal{D})$ is inconsistent). Clearly, the correctness above guarantees that G^* represents the set of all certain answers.

One can extend the approach presented above to OMQ answering in *datalogMTL*$^{\boxminus}_{nr}$. Indeed, it is possible to convert an arbitrary *datalogMTL*$^{\boxminus}_{nr}$ ontology to the one in the normal form *similar* to that used above. The tables P^* need to contain the tuples of the shape $\big(c_1, \dots, c_m, t_1, t_2, \langle, \rangle\big)$, where m is the arity of P. The rules for processing the temporal axioms essentially remain the same. The rules for computing the conjunctions (joins) need to be adjusted to correctly handle the individual arguments of the predicates.

Discussion and Future Work

Initial Experiments. We made experiments to evaluate the performance of the proposed algorithm on the $\mathsf{Hurricane}(x)@\delta$ and $\mathsf{ExcessiveHeat}(x)@\delta$ OMQs with the ontology from the weather use case. We implemented the algorithm of the previous section, for a given OMQ, as an SQL query using WITH clause and the RECURSIVE operator. That is, the intermediate tables of the algorithm are defined as a sequence of virtual SQL tables. The configuration of the computer that was used for the experiments is Intel Core i5 @ 2.7 GHz, 8 GB RAM with 1867 MHz DDR3 and OS X El Capitan operating system in version 10.11.4. The weather data is stored in 64 bit PostgreSQL version 9.4.5. We ran the queries over a table including 140 881 rows. It took 3 199 ms for Hurricane and 481 876 ms for ExcessiveHeat to retrieve the results. We interpret this outcome as a positive indication of the feasibility of our approach: even a straightforward implementation appears to work. We foresee the following three directions of the future work:

New Use Cases. Our language is capable of expressing complex patterns of events that are of interest for such purposes as diagnostics of engines or devices. The axiom

$$\mathsf{SmoothShutDown} \leftarrow \mathsf{IdleRPM} \wedge \boxminus_{\geqslant 0}^{\leqslant 15min} \mathsf{IntermRPM} \wedge \diamondminus_{\geqslant 15min}^{\leqslant 25min} \mathsf{RunningRPM},$$

for instance, describes the event of smooth shutdown of an engine as being in an idle state after having intermediate speed (RpM) for 15 minutes and having a running speed before that (not further than 25 minutes). The axiom:

$$\mathsf{ConsHighVibration} \leftarrow \boxminus_{\geqslant 0}^{\leqslant 50sec} \diamondminus_{>0}^{\leqslant 10sec} \mathsf{HighVibration}$$

describes consistent high vibration as high vibration occurring every 10 seconds during a minute. Our OBDA approach seems to be able to capture many industrial use-cases. In the future, we plan to investigate such potential applications.

Open Theoretical Problems. At the moment, we do not know whether OMQ answering in *HornMTL* is decidable. In fact, this question is open even for the fragment *HornMTL*$^{\Box}$. We plan to obtain complexity results for those languages, and we are particularly interested in *data complexity* (that is, the complexity in the size of $\mathcal{D}$ when $\boldsymbol{Q}(\delta)$ is assumed to be fixed). It is also important to understand how the complexity results for *HornMTL* carry over to *datalogMTL*. To achieve our goal, we plan to study various techniques developed in the area of metric temporal logics (Ouaknine and Worrell 2005; 2008; Hirshfeld and Rabinovich 2005) and modal logics over metric spaces (Kutz et al. 2003; Sheremet, Wolter, and Zakharyaschev 2010; Wolter and Zakharyaschev 2005).

Implementation and Optimizations. The proposed query answering algorithm for *datalogMTL*$^{\Box}_{nr}$ clearly allows for optimizations. For example, computing the transitive closure of the table Q^* when processing the axiom $P \leftarrow \boxplus_{\rhd e}^{\lhd d} Q$ seems to be avoidable. Moreover, our algorithm does not make any assumption regarding the temporal ordering of the tuples. If such a realistic assumption is made, we may be able to develop more efficient algorithms, in particular, by using indexes on timestamps.

References

Artale, A.; Kontchakov, R.; Kovtunova, A.; Ryzhikov, V.; Wolter, F.; and Zakharyaschev, M. 2015. First-order rewritability of temporal ontology-mediated queries. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, 2706–2712.

Baader, F.; Borgwardt, S.; and Lippmann, M. 2013. Temporalizing ontology-based data access. In *Proc. of the 24th Int. Conf. on Automated Deduction (CADE-24)*, volume 7898 of *LNCS*, 330–344. Springer.

Basulto, V. G.; Jung, J.; and Kontchakov, R. 2016. Temporalized EL ontologies for accessing temporal data: Complexity of atomic queries. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*. AAAI Press.

Gao, D.; Jensen, C. S.; Snodgrass, R. T.; and Soo, M. D. 2005. Join operations in temporal databases. *The VLDB Journal* 14(1):2–29.

Gutiérrez-Basulto, V., and Klarman, S. 2012. Towards a unifying approach to representing and querying temporal data in description logics. In *Proc. of the 6th Int. Conf. on Web Reasoning and Rule Systems (RR 2012)*, volume 7497 of *LNCS*, 90–105. Springer.

Hirshfeld, Y., and Rabinovich, A. 2005. Timer formulas and decidable metric temporal logic. *Information and Computation* 198(2):148 – 178.

Kharlamov, E.; Solomakhina, N.; Özçep, Ö.; Zheleznyakov, D.; Hubauer, T.; Lamparter, S.; Roshchin, M.; Soylu, A.; and Watson, S. 2014. How semantic technologies can enhance data access at Siemens Energy. In *Proc. of the 13th Int. Semantic Web Conf. (ISWC 2014), Part I*, volume 8796 of *LNCS*, 601–619. Springer.

Klarman, S., and Meyer, T. 2014. Querying temporal databases via OWL 2 QL. In *Proc. of the 8th Int. Conf. on Web Reasoning and Rule Systems (RR 2014)*, volume 8741 of *LNCS*, 92–107. Springer.

Kontchakov, R.; Pandolfo, L.; Pulina, L.; Ryzhikov, V.; and Zakharyaschev, M. 2016. Temporal and spatial obda with many-dimensional halpern-shoham logic. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*. AAAI Press.

Koymans, R. 1990. Specifying real-time properties with metric temporal logic. *Real-Time Systems* 2(4):255–299.

Kurucz, A.; Wolter, F.; and Zakharyaschev, M. 2005. Modal logics for metric spaces: Open problems. In *We Will Show Them! Essays in Honour of Dov Gabbay, Volume Two*, 193–108.

Kutz, O.; Wolter, F.; Sturm, H.; Suzuki, N.; and Zakharyaschev, M. 2003. Logics of metric spaces. *ACM Trans. Comput. Log.* 4(2):260–294.

Ouaknine, J., and Worrell, J. 2005. On the decidability of metric temporal logic. In *Proceedings of the 20th Annual IEEE Symposium on Logic in Computer Science*, LICS '05, 188–197. Washington, DC, USA: IEEE Computer Society.

Ouaknine, J., and Worrell, J. 2008. Some recent results in metric temporal logic. In *Proceedings of the 6th International Conference on Formal Modeling and Analysis of Timed Systems*, FORMATS '08, 1–13. Berlin, Heidelberg: Springer-Verlag.

Özcep, O.; Möller, R.; Neuenstadt, C.; Zheleznyakov, D.; and Kharlamov, E. 2013. A semantics for temporal and stream-based query answering in an OBDA context. Technical report, Deliverable D5.1, FP7-318338, EU.

Poggi, A.; Lembo, D.; Calvanese, D.; De Giacomo, G.; Lenzerini, M.; and Rosati, R. 2008. Linking data to ontologies. *J. on Data Semantics* X:133–173.

Sheremet, M.; Wolter, F.; and Zakharyaschev, M. 2010. A modal logic framework for reasoning about comparative distances and topology. *Ann. Pure Appl. Logic* 161(4):534–559.

Wolter, F., and Zakharyaschev, M. 2005. A logic for metric and topology. *J. Symb. Log.* 70(3):795–828.