

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

D. Calvanese¹, E. Güzel Kalaycı¹, V. Ryzhikov¹, G. Xiao¹ and
M. Zakharyashev²

¹ Free University of Bozen-Bolzano, Italy

² Birkbeck, University of London, UK

ONTOLP Workshop 2016, New York, USA

July 11, 2016

Temporal OBDA Setting

Raw Log Data mappings Conceptualized Temporal Data ontology End-User Queries

time	field1	...
...

⇒

$A(a)@l$
 $P(a, b)@l$

⇒

$q(x_1, \dots, x_n,$
 $\delta_1, \dots, \delta_m)$

Mappings: Example

Raw Log Data:

station id	time	wind speed
...
KYHS	13:25	105
KYHS	13:45	120
KYHS	14:30	125
KYHS	14:40	95
...

HurricaneForceWind(*x*)@(*t*₁, *t*₂)

"wind with the speed greater than 118 km/h is hurricane force"

Conceptualized

Temporal Data:

HurricaneForceWind(*KYHS*)@(*13:25*, *13:45*)

HurricaneForceWind(*KYHS*)@(*13:45*, *14:30*)

mappings:

SELECT

STATION_ID AS *x*,

lag(*TIME*) over (partition

by *STATION_ID* order by *TIME*) AS *t*₁,

TIME AS *t*₂,

"(" AS *<*,

"]" AS *>*

FROM *Weather*

WHERE *WIND_SPEED* > 118

Concept HurricaneForceWind(*x*) is temporal

Temporal Ontologies

- Declarative language to define **temporal concepts** in terms of other temporal concepts
- Capture temporal events
- “hurricane is a hurricane force wind lasting one hour or longer”

$$\text{Hurricane}(x) \leftarrow \boxminus_{\geq 0}^{\leq 1h} \text{HurricaneForceWind}(x)$$

- $\boxminus_{\geq 0}^{\leq 1h}$ is a metric temporal operator *‘during the previous hour’*

HurricaneForceWind(KYHS)@(13:25, 13:45] \Rightarrow Hurricane(KYHS)@(14:25, 14:30]
HurricaneForceWind(KYHS)@(13:45, 14:30]

We present temporal knowledge base language **datalogMTL**

datalogMTL: Syntax

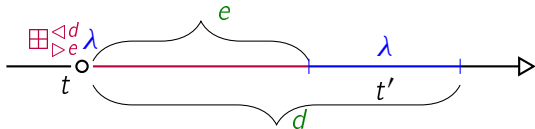
- Predicate symbols P_0, P_1, \dots , each of some arity $m \geq 0$
- Data instances \mathcal{D} with $P(\mathbf{c})@I$, where \mathbf{c} is an m -tuple of individual constants, and I an interval of \mathbb{R} : $P(\mathbf{c})$ is true at I
- An ontology \mathcal{O} is a finite set of axioms:

$$\lambda \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k, \quad \perp \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k$$

- Literals: $\lambda ::= P(\mathbf{x}) \mid \mathbf{O}_{\triangleright_e}^{\triangleleft_d} \lambda$,
 also $(x \neq x')$ and $(x = x')$ not in the head
- Temporal operators $\mathbf{O}_{\triangleright_e}^{\triangleleft_d}$:
 - $\boxplus_{\triangleright_e}^{\triangleleft_d}$ (always between e and d in the future),
 - $\boxminus_{\triangleright_e}^{\triangleleft_d}$ (always between e and d in the past),
 - $\diamond_{\triangleright_e}^{\triangleleft_d}$ (sometime between e and d in the future),
 - $\blacklozenge_{\triangleright_e}^{\triangleleft_d}$ (sometime between e and d in the past),
- \triangleleft is either $<$ or \leq , \triangleright is either $>$ or \geq .

datalogMTL: Semantics

- Interpretation, \mathfrak{M} , is based on the domain Δ of individuals in \mathcal{D} and the temporal domain \mathbb{R}
- For any m -ary predicate P , m -tuple \mathbf{c} from Δ and $t \in \mathbb{R}$, \mathfrak{M} specifies whether P is *true on \mathbf{c} at t* , we write $\mathfrak{M}, t \models P(\mathbf{c})$
- λ of shape $P(\mathbf{x})$ are interpreted as usual: $\mathfrak{M}, t \models^v \lambda$ using **assignment v**
- $\mathfrak{M}, t \models^v \boxplus_{\triangleright e}^{\triangleleft d} \lambda$ iff $\mathfrak{M}, t' \models^v \lambda$
for **all** t' such that $t' - t \triangleright e$ and $t' - t \triangleleft d$



- $\mathfrak{M}, t \models^v \boxminus_{\triangleright e}^{\triangleleft d} \lambda$ iff $\mathfrak{M}, t' \models^v \lambda$
for **some** t' such that $t' - t \triangleright e$ and $t' - t \triangleleft d$
- Axioms $\lambda \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k$, $\perp \leftarrow \lambda_1 \wedge \dots \wedge \lambda_k$ are interpreted **globally** (hold at **all** times)

Queries

- We consider **atomic queries** $P(\mathbf{x})@δ$
 - ▶ P is a predicate symbol of arity m
 - ▶ $δ$ is an interval variable
- Ontology-mediated query $\mathbf{q}(\mathbf{x}, δ) = (\mathcal{O}, P(\mathbf{x})@δ)$
- Certain answer to $\mathbf{q}(\mathbf{x}, δ)$ over \mathcal{D} is (\mathbf{c}, ι) such that
 - ▶ $\mathbf{c} = \nu(\mathbf{x})$ for some ν ,
 - ▶ $\mathfrak{M} \models (\mathcal{O}, \mathcal{D})$ implies $\mathfrak{M}, t \models P(\mathbf{c})$ for all $t \in \iota$

Example: Weather Ontology

$\text{Rain}(x) \leftarrow \text{PositiveTemp}(x) \wedge \text{Precipitation}(x),$

$\boxplus_{\geq 0}^{\leq 1h} \text{Hurricane}(x) \leftarrow \boxplus_{\geq 0}^{\leq 1h} \text{HurricaneForceWind}(x),$

$\boxplus_{\geq 0}^{\leq 24h} \text{ExcessiveHeat}(x) \leftarrow \boxplus_{\geq 0}^{\leq 24h} \text{TempAbove24}(x) \wedge$

$\boxplus_{\geq 0}^{\leq 24h} \text{TempAbove41}(x),$

$\text{HurricaneAffectedCounty}(x) \leftarrow \text{LocationOf}(x, y) \wedge \text{Hurricane}(y),$

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

Queries: $\text{Hurricane}(x)@{\delta}, \text{ExcessiveHeat}(x)@{\delta}$

Related Work

Consider **HornMTL**, a propositional fragment of **datalogMTL**:

$$\lambda \leftarrow \lambda_1 \wedge \cdots \wedge \lambda_k, \quad \perp \leftarrow \lambda_1 \wedge \cdots \wedge \lambda_k,$$
$$\lambda ::= P \mid \boxplus_{\triangleright e}^{\triangleleft d} \lambda \mid \boxminus_{\triangleright e}^{\triangleleft d} \lambda \mid \blacklozenge_{\triangleright e}^{\triangleleft d} \lambda \mid \blacktriangleleft_{\triangleright e}^{\triangleleft d} \lambda$$

Extending to **Boolean** clauses:

- Metric Temporal Logic
- Modal Logic of Metric Spaces
- Many known decidability+complexity/undecidability results

Our Contribution

- A fragment $\text{datalogMTL}_{nr}^{\square}$ of datalogMTL that is decidable
- $\diamond_{\triangleright e}^{\triangleleft d} \lambda$ and $\diamond_{\triangleright e}^{\triangleleft d} \lambda$ do not appear in the head of rules
- Non-recursiveness: $P(\mathbf{x})$ cannot be defined in terms of $P(\mathbf{x})$
- Algorithm for query answering based on **temporal joins** and **coalescing** for intervals
- We constructed an **SQL** query (using **WITH** clause and **RECURSIVE** operator) for the weather ontology above
- We ran the query on **real-world weather data**: termination on reasonable size data in reasonable time

Future Work

- Study complexity of full **datalogMTL** and its fragments other than **datalogMTL_{nr}**
- Establish connection with works on Metric Temporal Logic and Modal Logic of Metric Spaces
- Consider other practical use-cases: **monitoring of engines and devices**

SmoothShutDown \leftarrow IdleRPM \wedge $\exists_{>0}^{\leq 15min}$ IntermRPM \wedge

$\diamond_{\geq 15min}^{\leq 25min}$ RunningRPM,

ConsHighVibration \leftarrow $\exists_{>0}^{\leq 50sec}$ $\diamond_{>0}^{\leq 10sec}$ HighVibration

Algorithm Idea

- Store data $P(c_1, \dots, c_n) @ \langle t_1, t_2 \rangle$ in tables P^* with attributes $c_1, \dots, c_n, t_1, t_2, \langle, \rangle$
- For, e.g., $P(x) \leftarrow Q(x) \wedge R(x)$ add the tuple $c, t_1'', t_2'', [,]$ to P^* if there are $c, t_1, t_2, [,]$ in Q^* and $c, t_1', t_2', [,]$ in R^* , such that $[t_1, t_2] \cap [t_1', t_2'] \neq \emptyset$, where $[t_1'', t_2''] = [t_1, t_2] \cap [t_1', t_2']$
- For, e.g., $\boxplus_{\geq e}^d P(x) \leftarrow Q(x)$, add the tuple $c, t_1 + e, t_2 + d, [,]$ to P^* if there is a tuple $c, t_1, t_2, [,]$ in Q^*
- For, e.g., $P(x) \leftarrow \boxplus_{\geq e}^d Q(x)$, add the tuple $c, t_1 - e, t_2 - d, [,]$ to P^* , such that $c, t_1, t_2', [,]$ is in Q^* , $c, t_1', t_2, [,]$ is in Q^* and $[t_1, t_2]$ is covered by some set of intervals of shape $[t_1'', t_2'']$ for which $c, t_1'', t_2'', [,]$ is in Q^*