Compact Answers to Temporal Path Queries (Extended Version)

Anonymous author
Anonymous affiliation

Abstract

Several proposals have been made recently to extend graph databases with temporal properties, in order to store and access information about the evolution of data over time. In particular, Arenas et al. have proposed a model where facts are annotated with validity intervals, and queried via so-called Temporal Regular Path Queries (TRPQs), which extend regular path queries with temporal navigation. An important question that is left open is how to represent answers to such queries in a form that is not only finite, but also compact, and how to maintain such properties during query evaluation. We investigate four compact representations of answers to a TRPQ that rely on alternative ways of encoding sets of intervals. We discuss their respective advantages and drawbacks, in terms of finiteness, conciseness, uniqueness, and computational cost. Notably, the most refined representation can handle dense time. We also carry out a small evaluation that validates some our hypotheses about the conciseness of some representations.

2012 ACM Subject Classification Information systems → Query languages for non-relational engines

Keywords and phrases temporal databases, graph databases, regular path queries

Digital Object Identifier 10.4230/LIPIcs.ISWC.2025.101

1 Introduction

With the growing popularity of graph database (DB) engines, several proposals have been made recently to extend graphs with temporal properties, in order to store and access information about the evolution of data over time [2, 11, 27, 8, 22, 26]. We focus here on *Temporal Graphs* (TGs), where each fact is labeled with a set of time intervals that specify its validity. Equivalently, a TG can be viewed as a sequence of "snapshot" graphs, one for each time point, which consists of all facts that hold at that time point. Figure 1 represents a TG with time unit one hour. For conciseness, we represent it as a so-called *Property Graph*, one of the most popular graph data models [18]. In such a graph, both vertices (like `n1`) and edges (like `e1`) can carry attributes and a type. However, without loss of expressivity, the same data could be represented as a (less concise) edge-labelled graph, with time intervals associated to each edge.

In order to query such a graph, a sensible approach consists in extending a graph query language with temporal operators. Graph query languages, such as Cypher [18] or SPARQL [20], are based on navigational queries, whose basic form are so-called *Regular Path Queries* (RPQs). An RPQ q is a regular expression, and a pair $\langle o_1, o_2\rangle$ of objects in a graph is in the answer to q if there exists a path from o_1 to o_2 whose concatenated labels match this regular expression.

A natural extension of such queries consists in allowing navigation not only through the graph, but also through time. To this end, we consider *Temporal RPQs* (TRPQs), originally proposed by [2], which extend RPQs with a temporal navigation operator, allowing navigation from one object in a snapshot graph to the *same* object in a past or future snapshot graph. Hence, the answer to a TRPQ is a set of pairs $\langle\langle o_1, t_1\rangle, \langle o_2, t_2\rangle\rangle$, where o_1 and o_2 are associated with a time point each, respectively t_1 and t_2.

Surprisingly, this simple idea is an important departure from the way query answers are traditionally represented in temporal databases, where each tuple is instead associated with

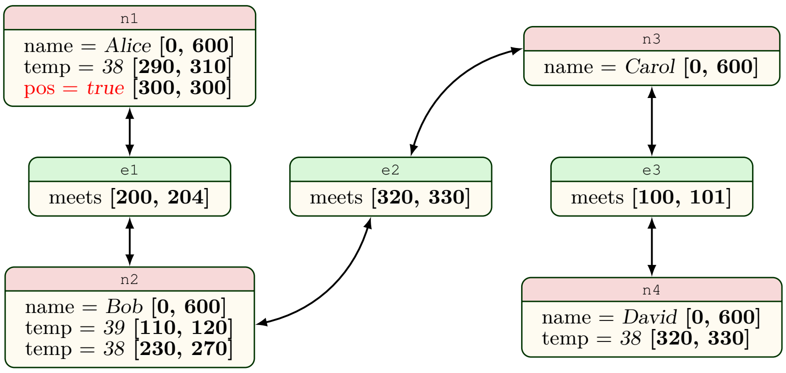


Figure 1 A Temporal Property Graph (TPG)

a *single* time point or interval for validity.[1] In particular, a central problem for traditional temporal query answering is producing answers in a *compact* form, using time intervals (this is even a necessity over dense time, to ensure that sets of answers are finite). A natural solution to this problem consists in computing answers in so-called *coalesced* form, using time intervals. This solution has long been adopted by temporal DB engines (e.g., [30, 14]) and also adapted for graph query languages such as a T-GQL [11]. However, in those approaches the time points assigned to a tuple are coalesced into intervals, and thus temporal joins only require computing interval intersection. But to our knowledge, this has not been investigated in a setting where the validity of each answer is associated with a *pair* of time points. Indeed, this question was left open by [2].

As an illustration, consider the TRPQ q_1 below that retrieves all pairs $\langle\langle p_1, t_1\rangle, \langle p_2, t_2\rangle\rangle$ such that person p_1 tested positive at time t_1, p_1 met p_2 within a week prior to t_1, and p_2 had high temperature at time t_2, less than two days after the meeting:

$$q_1 := \ (\texttt{pos} = \textit{true})/\mathrm{T}_{[-168,0]}/\mathrm{F}/\texttt{meets}/\mathrm{F}/\mathrm{T}_{[0,48]}/(\texttt{temp} \geq \textit{38})$$

The expressions $\texttt{pos} = \textit{true}$, $\texttt{meets}$ and $\texttt{temp} \geq \textit{38}$ 'locally' check whether a node or edge satisfies a certain property. The operator F stands for (atemporal) forward navigation, either from a node to an edge or conversely. The operator $\mathrm{T}_{[-168,0]}$ stands for temporal navigation in the past by at most a week (168 hours), and $\mathrm{T}_{[0,48]}$ for temporal navigation in the future by at most two days. There are 23 answers to q_1 over the TPG of Figure 1, precisely one tuple $\langle\langle n_1, 300\rangle, \langle n_2, t_2\rangle\rangle$ for each integer t_2 in the interval $[230, 252]$.

Another interesting feature of TRPQs is the transitive closure operator (written $[m, \_]$), inherited from RPQs. When applied to a temporal domain, this operator offers a natural way to express reachability under certain temporal constraints. For instance, let us assume that our virus may be carried at most one week by the same person. Then the query

$$q_2 := \ (\mathrm{T}_{[0,168]}/\mathrm{F}/\texttt{meets}/\mathrm{F})[1, \_]$$

returns all pairs $\langle\langle p_1, t_1\rangle, \langle p_2, t_2\rangle\rangle$ such that if p_1 was carrying the virus at time t_1, then it may have transitively transmitted it to person p_2 at time t_2. So the query

$$q_3 := (\texttt{pos} = \textit{true})/\mathrm{T}_{[-168,0]}/\mathrm{F}/\texttt{meets}/\mathrm{F}/q_2$$

[1] Bitemporal databases [21, 25] do associate two timepoints (or intervals) to each tuple, but only one of these stands for validity, while the other one represents the (orthogonal) notion of transaction time.

identifies people at risk (namely *Alice*, *Bob*, and *Carol*).

As we showed above, the 23 answers to the TRPQ q_1 can be coalesced with a single time interval for t_2. And similarly, for q_3, the 16 answers can be coalesced with only two time intervals. It is easy to see that computing all answers before summarizing them may be inefficient. For instance, a naive evaluation of query q_1 over the TPG of Figure 1 may join the 169 answers to the subquery $(\texttt{pos} = \mathit{true})/\mathrm{T}_{[0,-168]}$ with the 20 answers to the subquery $\mathrm{F}/\texttt{meets}$. Worse, a change of time granularity may have a dramatic impact on performance. E.g., the TPG of Figure 1 does not allow representing meetings shorter than an hour. But adopting minutes as a time unit instead of hours would multiply by 60 the cardinality of the operands of each join. So it is essential to not only represent answers in a compact way, but also to maintain compactness during query evaluation. To our knowledge, both problems are still open for the case where each answer tuple carries two time points. To address this, we provide the following contributions:

- We define four alternative compact representation of TRPQ answers, from less concise but conceptually simpler ones to more concise but also more complex ones.
- We analyze the respective advantages and drawbacks of these four formats, in terms of finiteness (over dense time), compactness (when finite), uniqueness, as well as computational cost of query answering and minimizing a set of tuples.

The rest of the paper is organized as follows. Before formalizing TGs and TRPQs in Section 3, we provide in Section 2 an informal overview of the four representations, which we then study in detail in Section 4. In Section **??**, we report an empirical evaluation of the compactness of answers under the first two representations, using the same dataset as [2]. Finally, in Section 5, we present the related work.

2 Summary of results

This section provides an intuitive presentation of the 4 compact representations of answers to TRPQs defined and studied in this article. Figure 3 summarizes our main findings for each of them. The precise meaning of columns in this table is explained in Section 4.

A key insight to understand these representations is the trade-off between folding either (pairs of) *time points*, or *distances* between time points. Let us consider the query

$$q_4 = (\texttt{name} = \mathit{Alice})/\mathrm{F}/\texttt{meets}/\mathrm{T}_{[2,3]}/\mathrm{F}$$

The answers to this query over the graph of Figure 1 (assuming discrete time) are listed in Figure 2 (upper left). The first compact representation is obtained by folding start time points into intervals, while grouping answers by objects and distance between start and end point. We use $\mathcal{U}^t$ to denote this format, which yields in our example the tuples in Figure 2 (upper right). This solution may be better-suited to inputs where time intervals in the graph are larger than the ones present in the query (such as the interval $[2,3]$ in Query q_4). For instance, even if one extends the duration of the meeting between Alice and Bob to 10 hours, from time 200 to 210, there are still only two compact answers under $\mathcal{U}^t$, one for each distance in the interval $[2,3]$, namely $\langle n_1, [200,208], 2, n_2\rangle$ and $\langle n_1, [200,207], 3, n_2\rangle$. But, the number of tuples may grow linearly in the length of the distance interval in the query.

A second, symmetric solution consists in folding distances, while grouping tuples by objects and starting time (or alternatively, end time). We call this format $\mathcal{U}^d$, which yields in our example the tuples of Figure 2 (middle left). In contrast to $\mathcal{U}^t$, this format may be better-suited when time intervals in the query are larger than those in the input graph.

Start point		End point	
Object	Time	Object	Time
n_1	200	n_2	202
n_1	201	n_2	203
n_1	202	n_2	204
n_1	200	n_2	203
n_1	201	n_2	204

Repr.	Start point			End point
	Object	Time	Distance	Object
$\mathcal{U}^t$	n_1	$[200, 202]$	2	n_2
	n_1	$[200, 201]$	3	n_2

Repr.	Start point		End point	
	Object	Time	Object	Dist.
$\mathcal{U}^d$	n_1	200	n_2	$[2, 3]$
	n_1	201	n_2	$[2, 3]$
	n_1	202	n_2	$[2, 2]$

Repr.	Start point			End point
	Object	Time	Dist.	Object
$\mathcal{U}^{td}$	n_1	$[200, 201]$	$[2, 3]$	n_2
	n_1	$[202, 202]$	$[2, 2]$	n_2

Repr.	Start point			End point	b	e
	Object	Time	Dist.	Object		
$\mathcal{U}^s$	n_1	$[200, 202]$	$[2, 3]$	n_2	200	201

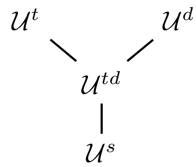


Figure 2 Answers to Query q_4 in non-compact form (upper left) and in the four compact representations. The diagram (bottom right) shows the comparative sizes of the compact answers in terms of number of tuples, where smaller sizes are at the bottom.

In our example, increasing the distance interval in Query q_4 to $[0, 3]$ would not affect the number of tuples. However, this number may grow linearly in the duration of the meeting.

Besides, a limitation of these two solutions is that they cannot handle intervals that represent dense time. Indeed, under both formats, the set of compact answers to a query may be infinite, if the query contains non-singleton intervals in the former case, and if the graph contains such intervals in the latter case. So a natural question is whether one can combine these two solutions, i.e., fold both time points and distances. We call this format $\mathcal{U}^{td}$. In our example, this yields the tuples of Figure 2 (middle right). We show that (maybe surprisingly) this third solution still cannot accommodate for dense time, if the query contains joins. Moreover, even though $\mathcal{U}^{td}$ is more compact than the two previous formats, the number of tuples may still be linear in the length of the input intervals. Also, uniqueness of representation is lost, in the sense that there may exist several (cardinality) minimal sets of tuples under this view that represent the set of answers to a query. In addition, minimizing a set of tuples (over discrete or dense time) becomes intractable, whereas it is in $O(n \log n)$ for the two previous representations. We also consider a variant of this representation where the answers represented by each tuple must be disjoint, which yields a potentially larger number of tuples. In this case, we regain tractability of minimization, but not uniqueness.

Finally, we define a more complex representation that not only can handle dense time, but also ensures that the number of tuples is independent of the length of the input time intervals. This format, called $\mathcal{U}^s$, extends the previous one with two values b and e. Answers in our example can now be represented with a single tuple, shown in Figure 2 (bottom left). In this tuple, the distance interval $[2, 3]$, together with b and e, specifies a range of distances δ_t for every time point t in the interval $[200, 202]$, precisely $[2 + \max(0, b - t), 3 - \max(0, t - e)]$,

	Finite (dense time)	Unique	Size (star-free q) data int.	Size (star-free q) query int.	Minimization	Query answering
$\mathcal{U}^t$	no	yes	$O(1)$	$\Omega(n)$	$O(n \log n)$	PSPACE-c
$\mathcal{U}^d$	no	yes	$\Omega(n)$	$O(1)$	$O(n \log n)$	PSPACE-c
$\mathcal{U}^{td}$	no	no	$O(1)$	$\Omega(n)$	NP-h / $O(n^{2.5})$	PSPACE-c
$\mathcal{U}^s$	yes	no	$O(1)$	$O(1)$	NP-h	PSPACE-h

Figure 3 Summary of results, where $\mathcal{U}^t$ stands for folding time points, $\mathcal{U}^d$ for folding distances, $\mathcal{U}^{td}$ for folding both, and $\mathcal{U}^s$ for the more complex representation (with extra "shifting" parameters). Minimization for $\mathcal{U}^{td}$ is tractable if redundant tuples are disallowed.

where 2 and 3 are the boundaries of the distance interval. Under this interpretation, it can be verified that this tuple represents all 5 original answers. For instance, if t is 200, then $\delta_t = [2 + \max(0, 200 - 200), 3 - \max(0, 200 - 201)] = [2, 3]$, and the two end points associated to 200 in the original answers are indeed $200 + 2 = 202$ and $200 + 3 = 203$. Similarly, if t is 202, then $\delta_t = [2+\max(0, 200-202), 3-\max(0, 202-201)] = [2, 2]$. More, this representation remains correct if the intervals of the graph and query are understood over dense time.

This last solution overcomes the limitations of the previous ones, in the sense that answers to a TRPQ can always be represented in a finite way, and their representation is guaranteed to be more compact. The price to pay however is an arguably less readable format. Besides, minimizing a set of tuples under this view is intractable.

3 Preliminaries

Sets, relations, order. For a binary relation R, we denote by $\mathsf{dom}(R)$ its domain and by $\mathsf{range}(R)$ its range. For a set S and a (possibly partial) order $\preceq$ over S, we denote by $\max_{\preceq} S$ the set of maximal elements in S w.r.t. $\preceq$, i.e., $\{s \in S \mid s \preceq s' \text{ implies } s = s' \text{ for all } s' \in S\}$.

Intervals, complement. We use $\mathsf{intv}(\mathbb{Z})$ (resp., $\mathsf{intv}(\mathbb{Q})$) for the set of nonempty intervals over $\mathbb{Z}$ (resp., $\mathbb{Q}$). If α is an interval, we use b_α for its beginning, e_α for its end, ${}_\alpha\lfloor$ for its left delimiter (either "(" or "["), and $\rfloor_\alpha$ for its right delimiter (either ")" or "]"). For instance, if $\alpha = [4, 6)$, then $b_\alpha = 4$, $e_\alpha = 6$, ${}_\alpha\lfloor$ is "[" and $\rfloor_\alpha$ is ")".

If $\alpha \in \mathsf{intv}(\mathbb{Z})$ (resp. $\mathsf{intv}(\mathbb{Q})$) and $t \in \mathbb{Z}$ (resp. $\mathbb{Q}$), we use $\alpha + t$ (resp. $\alpha - t$) for the interval identical to α, but where boundaries are shifted by t (resp $-t$), i.e. $\alpha + t$ is the interval ${}_\alpha\lfloor\ b_\alpha + t, e_\alpha + t\ \rfloor_\alpha$. If α and β are intervals, we use $\alpha + \beta$ for the interval with beginning $b_\alpha + b_\beta$, with end $e_\alpha + e_\beta$, and such that ${}_{\alpha+\beta}\lfloor$ is "(" iff both ${}_\alpha\lfloor$ and ${}_\beta\lfloor$ are "(", and similarly for $\rfloor_{\alpha+\beta}$. We also use $\alpha \ominus \beta$ for the interval with beginning $b_\alpha - e_\beta$, with end $e_\alpha - b_\beta$, and such that ${}_{\alpha\ominus\beta}\lfloor$ is "[" iff both ${}_\alpha\lfloor$ and $\rfloor_\beta$ are "[", and similarly for $\rfloor_{\alpha\ominus\beta}$.

Let $\alpha \in \mathsf{intv}(\mathbb{Z})$, and let S be a finite set of intervals s.t. $\bigcup S \subseteq \alpha$. We us $\mathsf{compl}\,(S, \alpha)$ to denote the complement of $\bigcup S$ in α represented as maximal intervals, i.e. :

$$\mathsf{compl}\,(S, \alpha) = \max_{\subseteq}\{I \subseteq \bigcup S \setminus \alpha \mid I \in \mathsf{intv}(\mathbb{Z})\}$$

And similarly for intervals in $\mathbb{Q}$.

Temporal Graphs. We adopt the same data model as in [2], with slight modifications in order to accommodate for either discrete or dense time, and to generalize the approach beyond Property Graphs (PGs). As a first modification, we assume that the underlying temporal domain $\mathcal{T}$ can be either discrete or dense. For simplicity, we assume that $\mathcal{T}$ is $\mathbb{Z}$ in the former case, and $\mathbb{Q}$ in the latter case. As a second modification, we abstract away from the specific representation of classes, labels, and attributes in PGs. Instead, we use a generic set *Pred* of boolean predicates whose validity for a given node (or edge) and time point can be checked locally, meaning that this verification is independent of the topology of the graph. For instance, over the graph of Figure 1, such predicates may be $\{\texttt{name} = \texttt{Alice}\}$, $\{\texttt{temp} = 38\}$, or `meets` (i.e., whether an edge has label `meets`). Formally, a *Temporal Graph* (TG) is a tuple $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$, where:

- N and E are finite sets of nodes and edges respectively, with $N \cap E = \emptyset$,
- $\mathsf{conn}\colon E \to N \times N$ maps an edge to its source and target,
- $\mathcal{T}_G$ is a closed-closed interval over $\mathcal{T}$, called the *active temporal domain*, and
- $\mathsf{val}\colon (N \cup E) \times \mathit{Pred} \to 2^{\mathsf{intv}(\mathcal{T}_G)}$ assigns a finite set of disjoint and pairwise non-adjacent intervals to each object o and predicate p, indicating when p holds for o.

If $\mathsf{conn}(e) = (n_1, n_2)$, we use $\mathsf{src}(e)$ for n_1 and $\mathsf{tgt}(e)$ for n_2.

Temporal Regular Path Queries. We study the query language introduced in [2], with minor modifications that allow us to handle dense time and to abstract away from Cypher and Property Graphs, so that our approach may be applied to other graph data model (with time intervals) and other (RPQ-based) graph query languages. A *Temporal Regular Path Query* (TRPQ) is an expression for the symbol "path" in the following grammar:

$$\mathsf{path} ::= \mathsf{test} \mid \mathsf{axis} \mid (\mathsf{path}/\mathsf{path}) \mid (\mathsf{path} + \mathsf{path}) \mid \mathsf{path}[m, n] \mid \mathsf{path}[m, \_]$$

$$\mathsf{test} ::= \mathit{pred} \mid (?\mathsf{path}) \mid < k \mid \mathsf{test} \vee \mathsf{test} \mid \mathsf{test} \wedge \mathsf{test} \mid \neg\mathsf{test}$$

$$\mathsf{axis} ::= \mathrm{F} \mid \mathrm{B} \mid \mathrm{T}_\delta$$

with $k \in \mathcal{T}, \delta \in \mathsf{intv}(\mathcal{T})$, $m, n \in \mathbb{N}^+$, and $m \leq n$.

The operator F (resp. B) stands for forward (resp., backward) atemporal navigation within a graph, either from a node to an edge or from an edge to a node. The temporal navigation operator T_δ stands for navigation in time by any distance in the interval δ. The terminal symbol *pred* stands for any element of *Pred*, i.e., a Boolean predicate that can be evaluated locally for one object and time point, as explained above. Similarly, the Boolean predicate $< k$ evaluates whether a time point is strictly inferior to k. The other operators are either Boolean combinations of these, or standard RPQ (a.k.a. regular epxression) operators. In particular, $\mathsf{path}[m, \_]$ stands for Kleene closure. The formal semantics of TRPQs is provided in Figure 4, where $[\![q]\!]_G$ is the evaluation of a TRPQ q over a TG G. In this defnition, we use q^i for the TRPQ defined inductively by $q^1 = q$ and $q^{i+1} = q^i/q$. For convenience, we represent (w.l.o.g.) an answer as two objects, one time point and a distance, rather than two objects and two time points, i.e. we use tuples of the form $\langle o_1, o_2, t, d\rangle$ rather than $\big\langle\langle o_1, t\rangle, \langle o_2, t{+}d\rangle\big\rangle$. We also use $\mathcal{U}_G$ to denote the set of tuples of this form with objects in G i.e., if $G = \langle N, E, \mathcal{T}_G, \mathsf{conn}, \mathsf{val}\rangle$, then $\mathcal{U}_G = \{\langle o_1, o_2, t, d\rangle \mid o_1, o_2 \in (N \cup E) \text{ and } t, d \in \mathcal{T}\}$. And we simply write $\mathcal{U}$ when G is clear from the context.

4 Compact answers

In this section, we define and study the four compact representation formats of answers to a TRPQ sketched in Section 2. Each of these representations can be viewed as an output format for the set of answers in $\mathcal{U}$ to a query. We specify each format as a set of admissible tuples, noted $\mathcal{U}^t$, $\mathcal{U}^d$, $\mathcal{U}^{td}$ and $\mathcal{U}^s$ respectively. Let $\mathcal{U}^x$ be any of these four sets. A tuple $\mathbf{u}$ in $\mathcal{U}^x$ represents a subset of $\mathcal{U}$, which we call the *unfolding* of $\mathbf{u}$. And the unfolding of a *set* $U \subseteq \mathcal{U}^x$ of such tuples is the union of their unfoldings. We say that U is *compact* if it is

$$
\begin{array}{rcl}
[\![pred]\!]_G & = & \{\langle o,o,t,0\rangle \mid t \in \tau \wedge \tau \in \mathsf{val}(o, pred)\} \\
[\![\mathrm{T}_\delta]\!]_G & = & \{\langle o,o,t,d\rangle \mid o \in (N \cup E), t \in \mathcal{T}, d \in \delta\} \\
[\![\mathrm{F}]\!]_G & = & \{\langle v,e,t,0\rangle \mid \mathsf{src}(e) = v, t \in \mathcal{T}\} \cup \{\langle e,v,t,0\rangle \mid \mathsf{tgt}(e) = v, t \in \mathcal{T}\} \\
[\![\mathrm{B}]\!]_G & = & \{\langle v,e,t,0\rangle \mid \mathsf{tgt}(e) = v, t \in \mathcal{T}\} \cup \{\langle e,v,t,0\rangle \mid \mathsf{src}(e) = v, t \in \mathcal{T}\} \\
[\![?\mathsf{path}]\!]_G & = & \{\langle o,o,t,0\rangle \mid \langle o,o',t,d\rangle \in [\![\mathsf{path}]\!]_G \text{ for some } o' \in N \cup E \text{ and } d \in \mathcal{T}\} \\
[\![\mathsf{test}_1 \vee \mathsf{test}_2]\!]_G & = & [\![\mathsf{test}_1]\!]_G \cup [\![\mathsf{test}_2]\!]_G \\
[\![\mathsf{test}_1 \wedge \mathsf{test}_2]\!]_G & = & [\![\mathsf{test}_1]\!]_G \cap [\![\mathsf{test}_2]\!]_G \\
[\![\neg\mathsf{test}]\!]_G & = & \big(\{\langle o,o\rangle \mid o \in N \cup E\} \times \mathcal{T} \times \{0\}\big) \setminus [\![\mathsf{test}]\!]_G \\
[\![\mathsf{path}_1/\mathsf{path}_2]\!]_G & = & \{\langle o_1,o_3,t,d_1+d_2\rangle \mid \\
 & & \exists o_2 \colon \langle o_1,o_2,t,d_1\rangle \in [\![\mathsf{path}_1]\!]_G \wedge \langle o_2,o_3,t+d_1,d_2\rangle \in [\![\mathsf{path}_2]\!]_G\} \\
[\![\mathsf{path}_1 + \mathsf{path}_2]\!]_G & = & [\![\mathsf{path}_1]\!]_G \cup [\![\mathsf{path}_2]\!]_G \\
[\![\mathsf{path}[m,n]]\!]_G & = & \bigcup_{k=m}^{n} [\![\mathsf{path}^k]\!]_G \\
[\![\mathsf{path}[m,\_]]\!]_G & = & \bigcup_{k \geq m} [\![\mathsf{path}^k]\!]_G
\end{array}
$$

Figure 4 Semantics of TRPQs

finite and if no strictly smaller (w.r.t. cardinality) subset of $\mathcal{U}^x$ has the same unfolding. A set $V \subseteq \mathcal{U}$ can be *finitely represented* (in $\mathcal{U}^x$) if there is a finite $U \subseteq \mathcal{U}^x$ with unfolding V.

4.1 Folding time points ($\mathcal{U}^t$)

Tuples under this view are identical to elements of $\mathcal{U}$, but where the time points associated to the first object are represented as intervals. The universe $\mathcal{U}^t$ of tuples is defined as

$$\mathcal{U}^t = \{\langle o_1, o_2, \tau, d\rangle \mid o_1, o_2 \in N \cup E, \tau \in \mathsf{intv}(\mathcal{T}), d \in \mathcal{T}\},$$

and the unfolding of $\langle o_1, o_2, \tau, d\rangle$ is $\{\langle o_1, o_2, t, d\rangle \mid t \in \tau\}$.

Inductive representation. In order to study when the answers $[\![q]\!]_G$ to a TRPQ q over a TG G can be finitely represented in $\mathcal{U}^t$, and what the size of such a representation may be, we define by induction on q a (not necessarily compact) representation of $[\![q]\!]_G$ in $\mathcal{U}^t$, noted $(\!|q|\!)^t_G$, wich also paves the way for an implementation. For instance, in the case where q is of the form *pred*, we define $(\!|pred|\!)^t_G$ as $\{\langle o, o, \tau, 0\rangle \mid o \in (N \cup E), \tau \in \mathsf{val}(o, pred)\}$.

The full definition of $(\!|q|\!)^t_G$ and a proof of correctness are provided in appendix. We only highlight here the least obvious operator, namely the temporal join q_1/q_2, illustrated with Figure 5, and defined as follows:

$$
\begin{aligned}
(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^t_G = \Big\{ & \langle o_1, o_3, ((\tau_1 + d_1) \cap \tau_2) - d_1, d_1 + d_2\rangle \mid \exists o_2 \colon \langle o_1, o_2, \tau_1, d_1\rangle \in (\!|\mathsf{path}_1|\!)^t_G \\
& \wedge \langle o_2, o_3, \tau_2, d_2\rangle \in (\!|\mathsf{path}_2|\!)^t_G \ \wedge\ (\tau_1 + d_1) \cap \tau_2 \neq \emptyset \Big\}
\end{aligned}
$$

Finiteness over dense time. Over discrete time, trivially, $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^t$ (and in any of the three other representations that we will consider below).[2] But over dense time, this is not always possible. For instance, let G be the graph of Figure 1 over dense time, and consider again query q_4. Then $\langle n_1, n_2, 200, d\rangle \in [\![q_4]\!]_G$ for every rational

[2] Recall that we assume the active temporal domain $\mathcal{T}_G$ of G to be bounded.

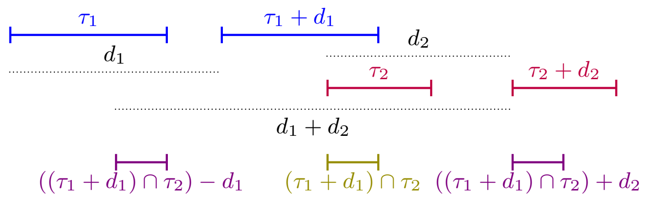


Figure 5 Join of two tuples in $\mathcal{U}^t$, whose intervals are depicted in blue and red respectively. The pair of intervals represented by the output tuple is depicted in violet.

number d in $[2, 3]$. And no tuple in $\mathcal{U}^t$ can represent more than one of these tuples. From the definition of $(\!|q|\!)^t_G$, the only possible source of non-finiteness is the temporal navigation operator T_δ, and only if δ specifies a certain *range* rather than a fixed distance:

▶ **Proposition 1.** *Let $G = \langle N, E, \mathcal{T}_G, \mathsf{conn}, \mathsf{val}\rangle$ be a TPG over dense time and q be a TRPQ such that δ is a singleton interval for every operator of the form T_δ in q. Then $(\!|q|\!)^t_G$ is finite.*

Compactness. If $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^t$, then a natural requirement on this representation is conciseness. It is easy to see that a finite set $U \subseteq \mathcal{U}^t$ is compact iff all time intervals for the same o_1, o_2 and d within U are coalesced. Formally, let $\sim$ denote the binary relation over $\mathcal{U}^t$ defined as $\langle o_1, o_2, \tau_1, d_1\rangle \sim \langle o_3, o_4, \tau_2, d_2\rangle$ iff $\langle o_1, o_2, d_1\rangle = \langle o_3, o_4, d_2\rangle$ and $\tau_1 \cup \tau_2 \in \mathsf{intv}(\mathcal{T})$. Then U is compact iff $\mathbf{u}_1 \not\sim \mathbf{u}_2$ for all $\mathbf{u}_1, \mathbf{u}_2 \in U$ s.t. $\mathbf{u}_1 \neq \mathbf{u}_2$. More, there is a unique way to coalesce a finite set of intervals. Therefore if $V \subseteq \mathcal{U}$ can be finitely represented in $\mathcal{U}^t$, then V also has a unique compact representation in $\mathcal{U}^t$.

However, the definition of $(\!|q|\!)^t_G$ does not preserve compactness. For instance, $(\!|q_1 + q_2|\!)^t_G$ may not be compact, even if $(\!|q_1|\!)^t_G$ and $(\!|q_2|\!)^t_G$ are. This observation generalizes to each of the (unary or binary) operators of the language, except for $\wedge$ and $\neg$. Besides, $(\!|\mathrm{T}_\delta|\!)^t_G$, when finite, is not guaranteed to be compact either. Coalescing a set of intervals is known to be in $O(n \log n)$, and efficient implementations have been devised (see Section 5). For this reason, coalescing intermediate results in $(\!|q|\!)^t_G$ may be a viable query evaluation strategy, in order for instance to reduce the size of the operands of a (worst-case quadratic) temporal join.

Size of compact answers. The size of the compact representation of $[\![q]\!]_G$ may be affected by the size of the active temporal domain $\mathcal{T}_G$, even if all time intervals in the query and graph are singletons, as soon as q contains an occurrence of the closure operator $[m, \_]$, as illustrated with the following example.

▶ **Example 2.** Consider a TG G with a single node o and no edge, let p be a boolean predicate such that $\mathsf{val}(o, p) = \{[0, 0]\}$, and let q be the query $(p/\mathrm{T}_{[2,2]})[1, \_]$. Then the compact representation of $[\![q]\!]_G$ in $\mathcal{U}^t$ is $\{\langle o, o, [0, 0], d\rangle \mid d \in D\}$, where $D = \{d \in \mathcal{T}_G \cap \mathbb{N}^+ \mid d \bmod 2 = 0\}$. ◀

The same observation holds (with the same example) for the three representations below.

For queries without closure operator, which we call *star-free*, this property does not hold, but the size of the compact representation of $[\![q]\!]_G$ in $\mathcal{U}^{td}$ may be affected by the size of the intervals present in the query. To formalize this, we introduce a notation that we will reuse for the other three representations. Let $\mathcal{U}^x$ be one of $\mathcal{U}^t$, $\mathcal{U}^d$, $\mathcal{U}^{td}$ or $\mathcal{U}^s$. Consider a TG G and a star-free query q such that $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^x$. Fix G and q, with the exception of $\mathcal{T}_G$ and time intervals in q, so that the cumulated length n of these intervals may grow arbitrarily, with the only requirement that $[\![q]\!]_G$ can still be finitely represented in $\mathcal{U}^x$. We use $\mathsf{size}^\delta([\![q]\!]_G, \mathcal{U}^x)$ for the cardinality of a compact representation of $[\![q]\!]_G$ in $\mathcal{U}^x$, expressed as a function of n. We also use $\mathsf{size}^{\tau}(\llbracket q \rrbracket_G, \mathcal{U}^x)$ with the same meaning, but where $\mathcal{T}_G$ and intervals in G may grow arbitrarily. The following result says that the size of the compact representation of $\llbracket q \rrbracket_G$ in $\mathcal{U}^t$ (when it exists) may be affected by the size of the intervals present in q, but not the ones used to label G.

▶ **Proposition 3.** *Let q be a star-free TRPQ and G a TG such that $\llbracket q \rrbracket_G$ can be finitely represented in $\mathcal{U}^t$. Then $\mathsf{size}^{\tau}(\llbracket q \rrbracket_G, \mathcal{U}^t) = O(1)$ and $\mathsf{size}^{\delta}(\llbracket q \rrbracket_G, \mathcal{U}^t) = \Omega(n)$.*

Complexity of query answering. We formulate a decision problem analogous to the classical boolean query answering problem (for atemporal databases), in such a way that it remains defined even if $\llbracket q \rrbracket_G$ does not admit a finite representation in $\mathcal{U}^t$.

Intuitively, decide whether a tuple $\mathbf{u}$ in $\mathcal{U}^t$ represents a set of answers (in $\mathcal{U}$), and whether the time interval in $\mathbf{u}$ is maximal. Formally, define the (partial) order $\sqsubseteq$ over $\mathcal{U}^t$ as $\langle o_1, o_2, \tau_1, d_1 \rangle \sqsubseteq \langle o_3, o_4, \tau_2, d_2 \rangle$ iff $\langle o_1, o_2, d_1 \rangle = \langle o_3, o_4, d_2 \rangle$ and $\tau_1 \subseteq \tau_2$. And let $\mathsf{unfold}(u)$ denote the unfolding of $\mathbf{u}$. We say that $\mathbf{u}$ is a *compact answer* to q over G if $\mathbf{u} \in \max_{\sqsubseteq}\{\mathbf{u}' \in \mathcal{U}^t \mid \mathsf{unfold}(\mathbf{u}') \subseteq \llbracket q \rrbracket_G\}$. We can now define our problem:

COMPACT ANSWER^t	
Input:	TG G, TRPQ q, tuple $\mathbf{u} \in \mathcal{U}^t$
Decide:	$\mathbf{u}$ is a compact answer to q over G

We show in the appendix that the results proven in [3] for answering TRPQs in $\mathcal{U}$ transfer to our setting, even in the case where $\llbracket q \rrbracket_G$ cannot be finitely represented in $\mathcal{U}^t$:

▶ **Proposition 4.** *COMPACT ANSWER^t is PSPACE-complete.*

We also emphasize that hardness is proven with a graph of fixed size, except for the active temporal domain $\mathcal{T}_G$.

We observe that, for this problem, complexity is driven by the size of the input time intervals, and there is no reason a priori to assume that intervals in the graphs are larger than the ones in the query. This is why the traditional distinction made in database theory between data and combined complexity is arguably less relevant here. Hence, we focus on combined complexity and leave a finer-grained analysis for future work.

4.2 Folding distances ($\mathcal{U}^d$)

This representation is symmetric to the previous one, using now intervals for distances (rather than time points). In other words, the universe $\mathcal{U}^d$ of tuples is defined as

$$\mathcal{U}^d = \{\langle o_1, o_2, t, \delta \rangle \mid o_1, o_2 \in N \cup E, t \in \mathcal{T}, \delta \in \mathsf{intv}(\mathcal{T})\}$$

And the unfolding of $\langle o_1, o_2, t, \delta \rangle$ is $\{\langle o_1, o_2, t, d \rangle \mid d \in \delta\}$.

Inductive representation. Similarly to what we did above for $\mathcal{U}^t$, we define in the appendix a representation $(\!|q|\!)^d_G$ of $\llbracket q \rrbracket_G$ in $\mathcal{U}^d$ by induction on q, and prove that it is correct. We highlight here operators for which this definition is less obvious. The first of these cases is temporal navigation, with

$$(\!|\mathrm{T}_\delta|\!)^d_G = \{\langle o, o, t, ((\delta + t) \cap \mathcal{T}_G) - t \rangle \mid o \in N \cup E, t \in \mathcal{T}_G, (\delta + t) \cap \mathcal{T}_G \neq \emptyset\}$$

which may be infinite over dense time. Alternatively, one can evaluate queries of the form q/T_δ inductively as

$$(\!|q/\mathrm{T}_\delta|\!)^d_G = \{\langle o_1, o_2, t, (\delta' + \delta) \cap \mathcal{T}_G \rangle \mid \langle o_1, o_2, t, \delta' \rangle \in (\!|q|\!)^d_G, (t + (\delta' + \delta)) \cap \mathcal{T}_G \neq \emptyset\}$$

which preserves finiteness, and symmetrically for queries of the form T_δ/q.

Finally, answers to a temporal join can be represented as

$$(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^d_G = \Big\{\langle o_1, o_3, t_1, \delta_2 + t_2 - t_1\rangle \mid \exists o_2\colon \langle o_1, o_2, t_1, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^d_G \land \langle o_2, o_3, t_2, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^d_G \land t_2 \in t_1 + \delta_1\Big\}$$

Finiteness over dense time. If G is over dense time, then $[\![q]\!]_G$ may not be finitely representable in $\mathcal{U}^d$. For instance, consider the query $q = (\texttt{temp >= 38})$ over the graph G of Figure 1. Then for every rational number t in $[290, 310]$, $\langle n_1, n_1, t, 0\rangle \in [\![q]\!]_G$, and no tuple in $\mathcal{U}^d$ can represent more than one of these tuples. We call a query *grounded* if it does not consist exclusively of joins (a.k.a. /) and temporal navigation operators (a.k.a. T_δ). From the definition of $(\!|q|\!)^d_G$, the only sources of non-finiteness are $\neg$, the base cases *pred* and $< k$, and the operator T_δ for non-grounded queries.

Compactness. For the same reasons as above with $\mathcal{U}^t$, a finite set $U \subseteq \mathcal{U}^d$ is compact iff all time intervals for the same o_1, o_2 and t within U are coalesced. And if $V \subseteq \mathcal{U}$ can be finitely represented in $\mathcal{U}^d$, then it also has a unique compact representation in $\mathcal{U}^d$.

Size of compact answers. For star-free (and grounded) queries, the results are symmetric to those for $\mathcal{U}^t$:

▶ **Proposition 5.** *Let q be a grounded star-free TRPQ and G a TG such that $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^d$. Then $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^d) = \Omega(n)$ and $\mathsf{size}^\delta([\![q]\!]_G, \mathcal{U}^d) = O(1)$.*

Complexity of query answering. We define for $\mathcal{U}^d$ a decision problem analogous to Compact Answer^t, and show (with almost identical proofs) that it is PSpace-complete.

4.3 Folding time points and distances ($\mathcal{U}^{td}$)

This representation generalizes the two previous ones, using intervals for time points and distances. The universe $\mathcal{U}^{td}$ is

$$\mathcal{U}^{td} = \{\langle o_1, o_2, \tau, \delta\rangle \mid o_1, o_2 \in N \cup E \text{ and } \tau, \delta \in \mathsf{intv}(\mathcal{T})\},$$

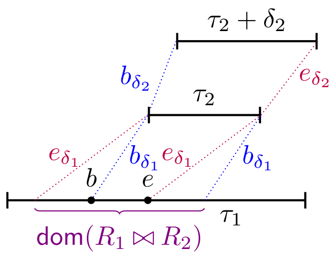
and the unfolding of $\langle o_1, o_2, \tau, \delta\rangle$ is $\{\langle o_1, o_2, t, d\rangle \mid t \in \tau, d \in \delta\}$.

Inductive representation. Again, in the appendix, we define and prove correctness of a representation $(\!|q|\!)^{td}_G$ of $[\![q]\!]_G$ in $\mathcal{U}^{td}$, and we focus here on two operators. for temporal join, we define $(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^{td}_G$ as:

$$\bigcup \{\mathbf{u}_1 \boxtimes \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^{td}_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^{td}_G\}$$

where $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ is defined as follows, and illustrated with Figure 6.

Figure 6 Temporal join for two tuples in $\mathcal{U}^{td}$. For simplicity, in this example, τ_2 and $\mathsf{range}(R_1) \cap \mathsf{dom}(R_2)$ coincide.



Let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1 \rangle$ and $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2 \rangle$. If $o_2 \neq o_3$, then $\mathbf{u}_1 \boxtimes \mathbf{u}_2 = \emptyset$. Otherwise, each tuple in $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ is of the form $\langle o_1, o_4, \tau, \delta \rangle$ for some τ and δ. For $i \in \{1,2\}$, let R_i be the binary relation over $\mathcal{T}$ specified by the time points and distances in $\mathbf{u}_i$, i.e., $R_i = \{(t, t+d) \mid t \in \tau_i, d \in \delta_i\}$. And let $R_1 \bowtie R_2$ denote $\{(t_1, t_3) \mid (t_1, t_2) \in R_1 \text{ and } (t_2, t_3) \in R_2 \text{ for some } t_2\}$. Then the intervals in the set $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ should intuitively represent this relation $R_1 \bowtie R_2$. For each time point $t \in \mathsf{dom}(R_1 \bowtie R_2)$, let δ_t denote the maximal interval s.t. $(\delta_t + t) \subseteq \tau_2$. We define $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ as $\{\langle o_1, o_4, [t,t], \delta_t + \delta_2 \rangle \mid t \in \mathsf{dom}(R_1 \bowtie R_2)\}$.

To complete the definition, we characterize δ_t and $\mathsf{dom}(R_1 \bowtie R_2)$ in terms of $\mathbf{u}_1$ and $\mathbf{u}_2$. For readability, we assume below that τ_1, τ_2 and δ_1 are all closed-closed intervals. First, define τ_2' as $\mathsf{range}(R_1) \cap \mathsf{dom}(R_2)$. Then from the definitions of R_1 and R_2, we get $\tau_2' = (\tau_1 + \delta_1) \cap \tau_2$. If $\tau_2' = \emptyset$, then $\mathsf{dom}(R_1 \bowtie R_2) = \emptyset$, otherwise $\mathsf{dom}(R_1 \bowtie R_2) = (\tau_2' \ominus \delta_1) \cap \tau_1$. Finally, let $b = b_{\tau_2'} - b_{\delta_1}$ and $e = e_{\tau_2'} - e_{\delta_1}$. Then for every $t \in \mathsf{dom}(R_1 \bowtie R_2)$,

$$\delta_t = [b_{\delta_1} + \max(0, b - t), e_{\delta_1} - \max(0, t - e)]$$

Note that if $b < e$, then $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ is not compact, because $\delta_t = \delta_1$ for every $t \in [b, e]$, as can be observed in Figure 6. Compactness can be recovered by modifying the definition of $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ accordingly (for the specific case where $o_2 = o_3$ and $b < e$), as $\{\langle o_1, o_4, [t,t], \delta_t + \delta_2 \rangle \mid t \in \mathsf{dom}(R_1 \bowtie R_2) \setminus [b,e]\} \cup \{\langle o_1, o_4, [b,e], \delta_1 + \delta_2 \rangle\}$.

We define $(\!|\mathrm{T}_\delta|\!)_G^{td}$ in a similar fashion, as $\bigcup_{o \in N \cup E} \{\langle o, o, \mathcal{T}_G, \delta \rangle \boxtimes \langle o, o, \mathcal{T}_G, [0,0] \rangle\}$.

Polygon cover. Finiteness and compactness in $\mathcal{U}^{td}$ can be related to the well-known problem of covering a rectilinear polygon with rectangles. Let $U \subseteq \mathcal{U}^{td}$ and $V \subseteq \mathcal{U}$ be two sets of tuples that share the same objects o_1 and o_2. Then U unfolds as V iff they intuitively cover the same area in the Euclidean plane of times per distances, i.e. if

$$\bigcup \{\tau \times \delta \mid \langle o_1, o_2, \tau, \delta \rangle \in U\} = \{(t, d) \mid \langle o_1, o_2, t, d \rangle \in V\}$$

Given a rectilinear polygon and a number k, deciding whether there is a set of at most k (possibly overlapping) rectangles that exactly cover the polygon is known to be NP-complete [10, 4]. However, finding a cover with a minimal number of non-overlapping rectangles is tractable [24, 16]. As an illustration, in Figure 7a, the first polygon has a unique minimal cover with (two) overlapping rectangles, and two minimal covers with (three) non-overlapping rectangles.

Finiteness over dense time. If $V \subseteq \mathcal{U}$ can be finitely represented in $\mathcal{U}^t$ (resp., $\mathcal{U}^d$), then clearly, it can also be finitely represented in $\mathcal{U}^{td}$ (but the converse may not hold). However, $[\![q]\!]_G$ may not be finitely representable in $\mathcal{U}^{td}$.

▶ **Example 6.** For instance, consider a TG G over dense time with a single node o and no edge, let p be a boolean predicate such that $\mathsf{val}(o, p) = \{[0,1]\}$, and let q be the query $p/\mathrm{T}_{[0,1]}/p$. Then $[\![q]\!]_G = \{\langle o, o, t, d \rangle \mid t \in [0,1] \text{ and } d \in [0, 1-t]\}$, which forms a triangle (precisely, with coordinates $(0,0), (0,1)$ and $(1,1)$) in our plane, as shown in Figure 7c. And obviously, this area cannot be exactly covered by finitely many rectangles.

Compactness. There may be several compact representations in $\mathcal{U}^{td}$ for the same $V \subseteq \mathcal{U}$. For instance, the minimal number of rectangle needed to cover an "L"-shaped polygon is two, and there are several such covers, as illustrated with Figure 7b. This argument easily generalizes to discrete time. Besides, for any rectilinear polygon P, a query q (with only unions) and graph G can be constructed in polynomial time so that $[\![q]\!]_G$ covers exactly P. Therefore, minimizing the representation of $[\![q]\!]_G$ in $\mathcal{U}^{td}$ (when finite) is intractable.

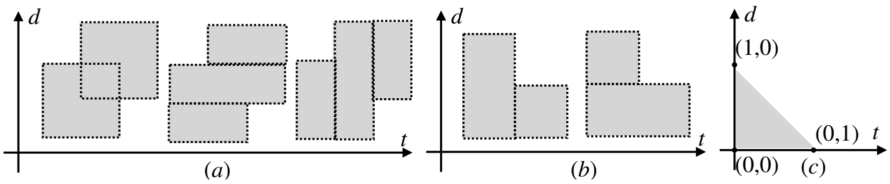


Figure 7 Polygons and coverings for times per distances

An important difference between this representation and the two previous ones is that a subset U of $\mathcal{U}^{td}$ may be *redundant*, meaning that two tuples $\mathbf{u}_1$ and $\mathbf{u}_2$ in U may have overlapping unfoldings. For instance, in Figure 7a, the first cover is minimal but redundant, whereas the two other covers are non-redundant but of size 3. If we require tuples to be non-redundant, then the representations of answers to a query may be less concise but more practical: transforming a given representation into a minimal non-redundant one is tractable (as discussed above), and non-redundant data may be better suited for downstream tasks, such as aggregation. However, uniqueness is not regained, as shown in Fig. 7a.

Size of compact answers. If $V \subseteq \mathcal{U}$ can be finitely represented in $\mathcal{U}^{td}$, then trivially, a compact representation of V in $\mathcal{U}^{td}$ must be smaller than the compact representation of V in $\mathcal{U}^{t}$ (resp. $\mathcal{U}^{d}$), if the latter exists. So compact answers under this representation must be smaller than under the two previous ones. However, maybe surprisingly, for star-free queries, the size of a compact representation may still be affected by the size of time intervals in q:

▶ **Proposition 7.** *Let q be a star-free TRPQ and G a TG such that $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^{td}$. Then $\mathsf{size}^{\tau}([\![q]\!]_G, \mathcal{U}^{td}) = O(1)$ and $\mathsf{size}^{\delta}([\![q]\!]_G, \mathcal{U}^{td}) = \Omega(n)$.*

Complexity of query answering. We define a problem analogous to Compact Answer^t for this representation, using element-wise set inclusion between pairs of intervals for the order $\sqsubseteq$, i.e., $\langle o_1, o_2, \tau_1, \delta_1\rangle \sqsubseteq \langle o_3, o_4, \tau_2, \delta_2\rangle$ iff $\langle o_1, o_2\rangle = \langle o_3, o_4\rangle$, $\tau_1 \subseteq \tau_2$ and $\delta_1 \subseteq \delta_2$. The problem is again PSpace-hard, and in PSpace if we allow redundant tuples.

4.4 Folding and shifting ($\mathcal{U}^s$)

We now define a fourth, more complex representation, which guarantees that a finite representation of $[\![q]\!]_G$ exists whose size is independent of the time intervals present in either q or G. The rationale is illustrated with Figure 6 (already discussed in the previous section), which shows why the temporal join of two tuples $\mathbf{u}_1$ and $\mathbf{u}_2 \in \mathcal{U}^{td}$ may not have a finite representation in $\mathcal{U}^{td}$. The key observation is that in this figure, the relation $R_1 \bowtie R_2$ is fully specified by $\mathsf{dom}(R_1 \bowtie R_2)$, δ_1, δ_2, b and e. Building on this observation, we define a representation that extends each tuple of the previous one with two values for b and e, so that each time point $t \in \tau$ can now be associated with a different range of distances, which we call δ_t.[3] Let $\langle o_1, o_2, \tau, \delta\rangle \in \mathcal{U}^{td}$, let $t \in \tau$, and let $b, e \in \mathcal{T}$. The following may be a (nonempty) interval:

$$_{\delta}\lfloor\ b_\delta + \max(0, b - t)\ ,\ e_\delta - \max(0, t - e)\ \rfloor_\delta$$

If this is the case, we use δ_t to denote this interval. Otherwise, δ_t is undefined.
We can now define our universe $\mathcal{U}^s$: a tuple $\langle o_1, o_2, \tau, \delta, b, e\rangle$ is in $\mathcal{U}^s$ iff

[3] For conciseness, this notation omits b and e, which should be clear from the context.

- $o_1, o_2 \in N \cup E$, $\tau, \delta \in \mathsf{intv}(\mathcal{T})$ and $b, e \in \mathcal{T}$, and
- δ_t is defined for every $t \in \tau$.

The unfolding of $\langle o_1, o_2, \tau, \delta, b, e\rangle$ is $\{\langle o_1, o_2, t, d\rangle \mid t \in \tau, d \in \delta_t\}$.

Inductive representation. In the appendix, we define a representation $(\!|q|\!)^s_G$ of $[\![q]\!]_G$ in $\mathcal{U}^s$ by induction on q, and proves that it is correct. For *pred*, the representation is:

$$(\!|pred|\!)^s_G = \{\langle o, o, \tau, [0,0], b_\tau, e_\tau\rangle \mid \tau \in \mathsf{val}(o, pred)\}$$

For temporal join, $(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^s_G$ is defined as:

$$\{\mathbf{u}_1 \bowtie \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^s_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^s_G, \mathbf{u}_1 \sim \mathbf{u}_2\}$$

where $\mathbf{u}_1 \sim \mathbf{u}_2$ and $\mathbf{u}_1 \bowtie \mathbf{u}_2$ are defined as follows.
For $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1, b_1, e_1\rangle$ and $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2, e_2, b_2\rangle$, let

$$\delta'_1 = {}_{\delta_1}\lfloor\, b_{\delta_1} + \max(0, b_1 - b_{\tau_1}), e_{\delta_1} - \max(0, e_{\tau_1} - e_1)\,\rfloor_{\delta_1}, \text{ and}$$
$$\tau = (((\tau_1 + \delta'_1) \cap \tau_2) \ominus \delta'_1) \cap \tau_1.$$

Then the relation $\sim \subseteq \mathcal{U}^s \times \mathcal{U}^s$ is defined as $\mathbf{u}_1 \sim \mathbf{u}_2$ iff $o_2 = o_3$ and $\tau \neq \emptyset$.
And if $\mathbf{u}_1 \sim \mathbf{u}_2$, then $\mathbf{u}_1 \bowtie \mathbf{u}_2$ is defined as $\langle o_1, o_4, \tau, \delta_1 + \delta_2, b, e\rangle$, with $b = \max(b_1, b_2 - b_{\delta_1})$ and $e = \min(e_1, e_2 - e_{\delta_1})$.

Similarly, we define $(\!|\mathrm{T}_\delta|\!)^s_G$ as $\{\langle o, o, \mathcal{T}_G, \delta, b_{\mathcal{T}_G}, e_{\mathcal{T}_G}\rangle \bowtie \langle o, o, \mathcal{T}_G, [0,0], b_{\mathcal{T}_G}, e_{\mathcal{T}_G}\rangle \mid o \in N \cup E\}$.

Finiteness over dense time. Finiteness follows from the definition of $(\!|q|\!)^s_G$, observing that the operator q_1/q_2 produces at most one tuple per pair $(\mathbf{u}_1, \mathbf{u}_2) \in (\!|q_1|\!)^s_G \times (\!|q_2|\!)^s_G$.

Compactness. Over dense time, we can use once again the Euclidean plane of time per distance to show that a set V of tuples in $\mathcal{U}$ may have several compact representations in $\mathcal{U}^s$ (a proof can be found in the appendix). As for the cost of minimizing a set of tuples, observe that any rectangle $\tau \times \delta$ in our plane is exactly covered by some tuple in $\mathcal{U}^s$. Therefore the NP-hardness result for $\mathcal{U}^{td}$ also translates to this setting, i.e., a rectilinear polygon can be covered with at most k rectangles iff there is a set of at most k tuples in $\mathcal{U}^s$ that cover exactly this area.

Size of compact answers. For star-free queries, immediately from the definition of $(\!|q|\!)^s_G$, the number of tuples in $(\!|q|\!)^s_G$ is not affected by the size of intervals present in G or q (even though $(\!|q|\!)^s_G$ is not necessarily compact).

▶ **Proposition 8.** *For a star-free TRPQ q and TG G, $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^s) = \mathsf{size}^\delta([\![q]\!]_G, \mathcal{U}^s) = O(1)$*

5 Related work

Temporal relational DBs. In temporal relational DBs, tuples are most commonly associated with a *single* time interval, viewed as a compact representations of time points at which the tuple holds [5]. The coalescing operator, which merges value-equivalent tuples over consecutive or overlapping time intervals, has received a lot of attention. Böhlen et al. [6] showed that coalescing can be implemented in SQL, and provided a comprehensive analysis of various coalescing algorithms and their performance. Later on, Al-Kateb et al. [1] investigated coalescing in the attribute timestamped CME temporal relational model, before Zhou et al. [31] exploited SQL:2003's analytical functions for the computation of coalescing. Their technique are the state-of-the-art, requiring a single scan over the ordered input and can be computed in $\mathcal{O}(n \log n)$. Also relevant to our work is the efficient computation of temporal joins over intervals. There has been a long line of research on temporal joins [19], ranging from partition-based [12, 9], index-based [15, 23], and sorting based [28, 7] techniques. Recently, in [13] it has been shown that a temporal join with the overlap predicate can be transformed into a sequence of two range joins. Our inductive representations of answers require temporal joins (cf. Section 4.1 and Figure 5) and range joins (cf. Section 4.2).

Temporal graphs. Temporal graph models vary in terms of temporal semantics, time representation (time point, interval), timestamped entities (graphs, nodes, edges, or attribute-value assignments), and whether they represent evolution of topology alone, or also of attributes. A sequence of snapshots is the simplest representation, in which a state of a graph is associated with either a time point or an interval during which it was in that state [17, 29]. Among recent proposals (and aside from [2], on which this paper builds), Byun et al. [8] developed ChronoGraph, which is both a temporal graph model and a graph traversal language, with dedicated aggregation techniques; Johnson et al. [22] developed Nepal, a query language scalable for large networks; Debrouvier et al. [11] introduced T-GQL, a Cypher-like query language for TPGs; Moffitt et al. [27] suggested an algebraic framework for analyzing temporal graphs, and Labouseur et al. [26] developed the graph DB system G$*$ for storing and managing dynamic graphs in distributed environments. To our knowledge, the problem we are addressing, of producing compact answers to a TRPQ, is new.

6 Conclusions

We defined and studied four alternative ways to produce compact answers to a TRPQ over a TG, which vary in terms of conciseness and potential usage. We believe this is a step forward towards integrating path queries and temporal navigation. Notably, the last of these representations ensures that answers to a query can always be finitely represented, and that the number of results is independent of the length of input time intervals. Among open questions, one may investigate the properties of non-redundancy under this last representation, in particular whether tractability for minimizing compact answers is regained.

References

1 Mohammed Al-Kateb, Essam Mansour, and Mohamed E. El-Sharkawi. CME: A temporal relational model for efficient coalescing. In *Proc. of the 12th Int. Symp. on Temporal Representation and Reasoning (TIME)*, pages 83–90. IEEE Computer Society, 2005.

2 Marcelo Arenas, Pedro Bahamondes, Amir Aghasadeghi, and Julia Stoyanovich. Temporal regular path queries. In *Proc. of the 38th IEEE Int. Conf. on Data Engineering (ICDE)*, pages 2412–2425. IEEE Computer Society, 2022.

3 Marcelo Arenas, Pedro Bahamondes, and Julia Stoyanovich. Temporal regular path queries: Syntax, semantics, and complexity. CoRR Technical Report arXiv:2107.01241, arXiv.org e-Print archive, 2021. Available at https://arxiv.org/abs/2107.01241. URL: https://arxiv.org/abs/2107.01241.

4 Larry Aupperle, Harold E. Conn, J. Mark Keil, and Joseph O'Rourke. *Covering Orthogonal Polygons with Squares.* Johns Hopkins University, Department of Computer Science, 1988.

5 Michael H. Böhlen, Anton Dignös, Johann Gamper, and Christian S. Jensen. Temporal data management – An overview. In *Tutorial Lectures of the 7th European Summer School on Business Intelligence and Big Data (eBISS)*, volume 324 of *Lecture Notes in Business Information Processing*, pages 51–83. Springer, 2017.

6 Michael H. Böhlen, Richard T. Snodgrass, and Michael D. Soo. Coalescing in temporal databases. In *Proc. of the 22nd Int. Conf. on Very Large Data Bases (VLDB)*, pages 180–191, 1996.

7 Panagiotis Bouros, Nikos Mamoulis, Dimitrios Tsitsigkos, and Manolis Terrovitis. In-memory interval joins. *Very Large Database J.*, 30(4):667–691, 2021.

8 Jaewook Byun, Sungpil Woo, and Daeyoung Kim. Chronograph: Enabling temporal graph traversals for efficient information diffusion analysis over time. *IEEE Trans. on Knowledge and Data Engineering*, 32(3):424–437, 2020.

9 Francesco Cafagna and Michael H. Böhlen. Disjoint interval partitioning. *Very Large Database J.*, 26(3):447–466, 2017.

10 Joseph C. Culberson and Robert A. Reckhow. Covering polygons is hard. *J. of Algorithms*, 17(1):2–44, 1994.

11 Ariel Debrouvier, Eliseo Parodi, Matías Perazzo, Valeria Soliani, and Alejandro Vaisman. A model and query language for temporal graph databases. *Very Large Database J.*, 30(5):825–858, 2021.

12 Anton Dignös, Michael H. Böhlen, and Johann Gamper. Overlap interval partition join. In *Proc. of the 35th ACM Int. Conf. on Management of Data (SIGMOD)*, pages 1459–1470, 2014.

13 Anton Dignös, Michael H. Böhlen, Johann Gamper, Christian S. Jensen, and Peter Moser. Leveraging range joins for the computation of overlap joins. *Very Large Database J.*, 31(1):75–99, 2022.

14 Anton Dignös, Boris Glavic, Xing Niu, Johann Gamper, and Michael H. Böhlen. Snapshot semantics for temporal multiset relations. *Proc. of the VLDB Endowment*, 12(6):639–652, 2019.

15 Jost Enderle, Matthias Hampel, and Thomas Seidl. Joining interval data in relational databases. In *Proc. of the 25th ACM Int. Conf. on Management of Data (SIGMOD)*, pages 683–694, 2004.

16 David Eppstein. Graph-theoretic solutions to computational geometry problems. In *Revised Papers the 35th Int. Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 5911 of *Lecture Notes in Computer Science*, pages 1–16, 2009.

17 Arash Fard, Amir Abdolrashidi, Lakshmish Ramaswamy, and John A Miller. Towards efficient query processing on massive time-evolving graphs. In *Proc. of the 8th Int. Conf. on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pages 567–574. IEEE Computer Society, 2012.

18 Nadime Francis, Alastair Green, Paolo Guagliardo, Leonid Libkin, Tobias Lindaaker, Victor Marsault, Stefan Plantikow, Mats Rydberg, Petra Selmer, and Andrés Taylor. Cypher: An evolving query language for property graphs. In *Proc. of the 39th ACM Int. Conf. on Management of Data (SIGMOD)*, pages 1433–1445, 2018.

19 Dengfeng Gao, Christian S. Jensen, Richard T. Snodgrass, and Michael D. Soo. Join operations in temporal databases. *Very Large Database J.*, 14(1):2–29, 2005.

20 Steve Harris and Andy Seaborne. SPARQL 1.1 query language. W3C Recommendation, World Wide Web Consortium, March 2013. Available at http://www.w3.org/TR/sparql11-query.

21 Christian S. Jensen and Richard T. Snodgrass. Bitemporal relation. In *Encyclopedia of Database Systems*. Springer, 2nd edition, 2018.

22 Theodore Johnson, Yaron Kanza, Laks VS Lakshmanan, and Vladislav Shkapenyuk. Nepal: a path query language for communication networks. In *Proc. of the 1st ACM SIGMOD Workshop on Network Data Analytics*, pages 1–8, 2016.

23 Martin Kaufmann, Amin Amiri Manjili, Panagiotis Vagenas, Peter M. Fischer, Donald Kossmann, Franz Färber, and Norman May. Timeline index: a unified data structure for processing queries on temporal data in SAP HANA. In *Proc. of the 34th ACM Int. Conf. on Management of Data (SIGMOD)*, pages 1173–1184, 2013.

24 J. Mark Keil. Minimally covering a horizontally convex orthogonal polygon. In *Proc. of the 2nd Annual ACM SIGACT/SIGGRAPH Symposium on Computational Geometry (SCG)*, pages 43–51, 1986.

25 Krishna G. Kulkarni and Jan-Eike Michels. Temporal features in SQL: 2011. *SIGMOD Record*, 41(3):34–43, 2012.

26 Alan G. Labouseur, Jeremy Birnbaum, Paul W. Olsen, Sean R. Spillane, Jayadevan Vijayan, Jeong-Hyon Hwang, and Wook-Shin Han. The G* graph database: Efficiently managing large distributed dynamic graphs. *Distributed and Parallel Databases*, 33:479–514, 2015.

27 Vera Zaychik Moffitt and Julia Stoyanovich. Temporal graph algebra. In *Proc. of the 16th Int. Symp. on Database Programming Languages (DBPL)*, pages 1–12, 2017.

28 Danila Piatov, Sven Helmer, and Anton Dignös. An interval join optimized for modern hardware. In *Proc. of the 32th IEEE Int. Conf. on Data Engineering (ICDE)*, pages 1098–1109. IEEE Computer Society, 2016.

29 Chenghui Ren, Eric Lo, Ben Kao, Xinjie Zhu, and Reynold Cheng. On querying historical evolving graph sequences. *Proc. of the VLDB Endowment*, 4(11):726–737, 2011.

30 Richard T. Snodgrass, editor. *The TSQL2 Temporal Query Language*. Kluwer, 1995.

31 Xin Zhou, Fusheng Wang, and Carlo Zaniolo. Efficient temporal coalescing query support in relational database systems. In *Proc. of the 17th Int. Conf. on Database and Expert Systems Applications (DEXA)*, volume 4080 of *Lecture Notes in Computer Science*, pages 676–686. Springer, 2006.

A Appendix

The structure adopted in this appendix differs from one followed in the article. Results here are grouped by topic (inductive representation, complexity, etc.), rather than by format ($\mathcal{U}^t$, $\mathcal{U}^d$, etc.). This allows us to factorize some proofs, and emphasize what differs from one case to the other.

The most technical results are the correctness of the inductive representation of compact answers in $\mathcal{U}^s$ (in particular for the join operator), proven in Section A.2.4.2, and to a lesser extent the analogous result for $\mathcal{U}^{td}$, proven in Section A.2.3.2.

On the other hand, complexity proofs (in Section A.3) leverage results already proven in [3].

Results pertaining to the size of compact answers (in Section A.4), follow either from the corresponding inductive representations (for the upper bounds), or from simple examples (for the lower bounds), similar to the ones already provided in the body of the article.

Most arguments that pertain to compactness and cost of coalescing answers are already provided in Section 4, so we only complete these when necessary, in Section A.5.

Finally, for finiteness over dense time, all negative results (i.e. non-finiteness) are illustrated in the article, and all positive results (i.e. finiteness) follow from the definitions of the inductive representations (and the fact that they are correct).

Table of Content

A.1 Notation

Let $\langle o_1, o_2, \tau, \delta, b, e\rangle \in \mathcal{U}^s$, and let $t \in \tau$.
In the article, we defined the interval δ_t for each t as

$$_{\delta}\lfloor\ b_\delta + \max(0, b - t)\ ,\ e_{\delta_i} - \max(0, t - e)\ \rfloor_{\delta}$$

In this appendix, we will use $\delta(t)$ instead of δ_t.
This notation will allow us to write $\delta_1(t)$ when several tuples are involved.
Note that the time points b and e in this notaton are still omitted, for conciseness, because they should be clear from the context.

A.2 Inductive characterizations

Let q be a TRPQ and G a TG.
Then $[\![q]\!]_G$ is the set of anwers to q over G (represented as tuples in $\mathcal{U}$).
In this section, we provide the full definition of the four inductive representations of $[\![q]\!]_G$ discussed in the article, in $\mathcal{U}^t$, $\mathcal{U}^d$, $\mathcal{U}^{td}$ and $\mathcal{U}^s$ respectively, and prove that they are correct.
These representations are denoted as $(\!|q|\!)^t_G$, $(\!|q|\!)^d_G$, $(\!|q|\!)^{td}_G$ and $(\!|q|\!)^s_G$ respectively.

A.2.1 In $\mathcal{U}^t$

A.2.1.1 Definition

If q is a TRPQ and $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ a TG, then the representation $(\!|q|\!)^t_G$ of $[\![q]\!]_G$ in $\mathcal{U}^t$ is defined inductively as follows:

$$
\begin{aligned}
(\!|< k|\!)^t_G &= \begin{cases} \{\langle o, o, [b_{\mathcal{T}_G}, k), 0\rangle \mid o \in (N \cup E)\} \text{ if } k \leq e_{\mathcal{T}_G}, \text{ and} \\ \emptyset \text{ otherwise}\end{cases}\\
(\!|\mathit{pred}|\!)^t_G &= \{\langle o, o, \tau, 0\rangle \mid o \in (N \cup E), \tau \in \mathsf{val}(o, \mathit{pred})\}\\
(\!|\mathrm{T}_\delta|\!)^t_G &= \{\langle o, o, [t_1, t_1], t_2 - t_1\rangle \mid o \in (N \cup E), t_1 \in \mathcal{T}_G, t_2 \in (\delta + t_1) \cap \mathcal{T}_G\}\\
(\!|\mathrm{F}|\!)^t_G &= \{\langle v, e, \mathcal{T}_G, 0\rangle \mid \mathsf{src}(e) = v\} \cup \{\langle e, v, \mathcal{T}_G, 0\rangle \mid \mathsf{tgt}(e) = v\}\\
(\!|\mathrm{B}|\!)^t_G &= \{\langle v, e, \mathcal{T}_G, 0\rangle \mid \mathsf{tgt}(e) = v\} \cup \{\langle e, v, \mathcal{T}_G, 0\rangle \mid \mathsf{src}(e) = v\}\\
(\!|(?\mathsf{path})|\!)^t_G &= \{\langle o, o, \tau, 0\rangle \mid \langle o, o', \tau, d\rangle \in (\!|\mathsf{path}|\!)^t_G \text{ for some } o' \in N \cup E, d \in \mathcal{T}\}\\
(\!|\mathsf{test}_1 \vee \mathsf{test}_2|\!)^t_G &= (\!|\mathsf{test}_1|\!)^t_G \cup (\!|\mathsf{test}_2|\!)^t_G\\
(\!|\mathsf{test}_1 \wedge \mathsf{test}_2|\!)^t_G &= \{\langle o, o, \tau_1 \cap \tau_2, 0\rangle \mid \langle o, o, \tau_1, 0\rangle \in (\!|\mathsf{test}_1|\!)^t_G, \langle o, o, \tau_2, 0\rangle \in (\!|\mathsf{test}_2|\!)^t_G, \tau_1 \cap \tau_2 \neq \emptyset\}\\
(\!|\neg\mathsf{test}|\!)^t_G &= \bigcup_{o \in N \cup E} \Big\{\langle o, o, \tau, 0\rangle \mid \tau \in \mathsf{compl}\left(\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\}, \mathcal{T}_G\right)\Big\}\\
(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^t_G &= \Big\{\langle o_1, o_3, ((\tau_1 + d_1) \cap \tau_2) - d_1, d_1 + d_2\rangle \mid \\
&\qquad \exists o_2 \colon \langle o_1, o_2, \tau_1, d_1\rangle \in (\!|\mathsf{path}_1|\!)^t_G \wedge \langle o_2, o_3, \tau_2, d_2\rangle \in (\!|\mathsf{path}_2|\!)^t_G \wedge (\tau_1 + d_1) \cap \tau_2 \neq \emptyset\Big\}\\
(\!|\mathsf{path}_1 + \mathsf{path}_2|\!)^t_G &= (\!|\mathsf{path}_1|\!)^t_G \cup (\!|\mathsf{path}_2|\!)^t_G\\
(\!|\mathsf{path}[m, n]|\!)^t_G &= \bigcup_{k=m}^{n} (\!|\mathsf{path}^k|\!)^t_G\\
(\!|\mathsf{path}[m, \_]|\!)^t_G &= \bigcup_{k \geq m} (\!|\mathsf{path}^k|\!)^t_G
\end{aligned}
\qquad (628)
$$

We observe that when q is of the form $(\mathsf{path}_1 + \mathsf{path}_2)$, $(\mathsf{path}[m, \_])$ and $(\mathsf{path}[m, n])$, the defnition of $(\!|q|\!)^t_G$ is nearly identical to the one of $[\![q]\!]_G$.
This also holds for the three representations below.

A.2.1.2 Correctness

We start with a lemma:

▶ **Lemma 9.** *Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TPG and let q an expression for the symbol* test *in the grammar of Section 3.*
Then:

- *each tuples in $[\![q]\!]_G$ is of the form $\langle o_1, o_2, t, 0\rangle$ for some o_1, o_2 and t,*

- *each tuples in* $(\!|q|\!)^t_G$ *is of the form* $\langle o_1, o_2, \tau, 0\rangle$ *for some* o_1, o_2 *and* τ*.*

Proof. Immediate from the definitions of $[\![q]\!]_G$ and $(\!|q|\!)^t_G$. ◀

The following result states that the representation $(\!|q|\!)^t_G$ is correct:

▶ **Proposition 10.** *Let* $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ *be a TPG and* q *a TRPQ. Then the unfolding of* $(\!|q|\!)^t_G$ *is* $[\![q]\!]_G$*.*

Proof.
Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TG, and let q be a TRPQ.
We show below that:

(I) for any $\langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, there is a $\tau \in \mathsf{intv}(\mathcal{T})$ such that
 (a) $\langle o_1, o_2, \tau, d\rangle \in (\!|q|\!)^t_G$, and
 (b) $t \in \tau$,

(II) for any $\langle o_1, o_2, \tau, d\rangle \in (\!|q|\!)^t_G$ for any $t \in \tau$,
$\langle o_1, o_2, t, d\rangle$ is in $[\![q]\!]_G$.

We proceed by induction on the structure of q.
If q is of the form *pred*, $< k$, F, B, $(\mathsf{test} \vee \mathsf{test})$, $(\mathsf{path} + \mathsf{path})$, $\mathsf{path}[m, n]$ or $\mathsf{path}[m, \_]$, then I and II immediately follow from the definitions of $[\![q]\!]_G$ and $(\!|q|\!)^t_G$.
So we focus below on the five remaining cases:

- $q = \mathrm{T}_\delta$.
  From the above definitions, we have:

$$\begin{aligned}
[\![q]\!]_G &= \{\langle o, o, t, d\rangle \mid o \in (N \cup E), t \in \mathcal{T}_G, d \in \delta, t + d \in \mathcal{T}_G\} \\
(\!|q|\!)^t_G &= \{\langle o, o, [t_1, t_1], t_2 - t_1\rangle \mid o \in (N \cup E), t_1 \in \mathcal{T}_G, t_2 \in (\delta + t_1) \cap \mathcal{T}_G\}
\end{aligned}$$

  - For I, let $\mathbf{v} = \langle o, o, t, t + d\rangle \in [\![q]\!]_G$.
    And let $\mathbf{u} = \langle o, o, [t, t], d\rangle$ in $\mathcal{U}^t$.
    For Ia we show that $\mathbf{u} \in (\!|q|\!)^t_G$.
    From $\mathbf{v} \in [\![q]\!]_G$, we get $o \in N \cup E$ and $t \in \mathcal{T}_G$.
    Besides, because $\mathbf{v} \in [\![q]\!]_G$ still,

$$t + d \in \mathcal{T}_G \tag{1}$$

    and

$$\begin{aligned}
d &\in \delta && (2)\\
t + d &\in t + \delta && (3)
\end{aligned}$$

    So from (1) and (3)

$$t + d \in (\delta + t) \cap \mathcal{T}_G \tag{4}$$

    So there is a t_2 (namely $t + d$) such that $d = t_2 - t$ and $t_2 \in t + \delta \cap \mathcal{T}_G$.
    Together with the definition of $(\!|q|\!)^t_G$, this implies $\mathbf{u} \in (\!|q|\!)^t_G$, which concludes the proof for Ia.
    And trivially, $t \in [t, t]$, so Ib is verified as well.

- For II, let $\mathbf{u} = \langle o, o, [t, t], d\rangle \in (\!|q|\!)^t_G$.
  From $\mathbf{u} \in (\!|q|\!)^t_G G$, we get $o \in N \cup E$ and $t \in \mathcal{T}_G$.
  So to conclude the proof, it is sufficient to show that *(i)* $d \in \delta$ and *(ii)* $t + d \in \mathcal{T}_G$.
  Because $\mathbf{u} \in (\!|q|\!)^t_G G$ still, we have

$$d = t_2 - t \text{ for some } t_2 \in (\delta + t) \cap \mathcal{T}_G \tag{5}$$

  From (5), we get $t_2 = t + d$.
  Therefore from (5) still,

$$t + d \in (\delta + t) \cap \mathcal{T}_G \tag{6}$$

  which proves *(ii)* .
  And from (6), we also get

$$\begin{aligned} t + d &\in \delta + t \\ t + d - t &\in (\delta + t) - t \\ d &\in \delta \end{aligned}$$

  which proves *(i)* .

- $q = \mathsf{test}_1 \wedge \mathsf{test}_2$.
From the above definitions, we have:

$$\begin{aligned} [\![q]\!]_G =& [\![\mathsf{test}_1]\!]_G \cap [\![\mathsf{test}_2]\!]_G \\ (\!|q|\!)^t_G =& \{\langle o, o, \tau_1 \cap \tau_2, 0\rangle \mid \langle o, o, \tau_1, 0\rangle \in (\!|\mathsf{test}_1|\!)^t_G, \langle o, o, \tau_2, 0\rangle \in (\!|\mathsf{test}_2|\!)^t_G, \tau_1 \cap \tau_2 \neq \emptyset\} \end{aligned}$$

- For I, let $\mathbf{v} = \langle o, o, t, d\rangle \in [\![q]\!]_G$.
  From Lemma 9, $d = 0$.
  And from the definition of $[\![q]\!]_G$, $\mathbf{v} \in [\![\mathsf{test}_1]\!]_G \cap [\![\mathsf{test}_2]\!]_G$.
  So by IH, there are intervals τ_1 and τ_2 s.t. $\langle o, o, \tau_i, 0\rangle \in (\!|\mathsf{test}_i|\!)^t_G$ for $i \in \{1, 2\}$ and $t \in \tau_1 \cap \tau_2$.
  Together with the definition of $(\!|q|\!)^t_G$, this proves I.
- For II, let $\langle o, o, \tau, d\rangle \in (\!|q|\!)^t_G$.
  Then from Lemma 9, $d = 0$.
  And from the definition of $(\!|q|\!)^t_G$, there are two intervals τ_1 and τ_2 s.t. $\tau = \tau_1 \cap \tau_2$ and $\langle o, o, \tau_i, 0\rangle \in (\!|\mathsf{test}_i|\!)^t_G$ for $i \in \{1, 2\}$.
  Now take any $t \in \tau$.
  Then $t \in \tau_i$ for $i \in \{1, 2\}$.
  So by IH, $\langle o, o, t, 0\rangle \in [\![\mathsf{test}_i]\!]_G$ for each $i \in \{1, 2\}$.
  Together with the defintiion of $[\![q]\!]_G$, this proves II.

- $q = (?\mathsf{path})$.
From the above definitions, we have:

$$\begin{aligned} [\![q]\!]_G &= \{\langle o, o, t, 0\rangle \mid \langle o, o', t, t + d\rangle \in [\![\mathsf{path}]\!]_G \text{ for some } o' \in N \cup E, d \in \mathcal{T}\} \\ (\!|q|\!)^t_G &= \{\langle o, o, \tau, 0\rangle \mid \langle o, o', \tau, d\rangle \in (\!|\mathsf{path}|\!)^t_G \text{ for some } o' \in N \cup E, d \in \mathcal{T}\} \end{aligned}$$

- For I, let $\langle o, o, t, 0\rangle \in [\![q]\!]_G$.
  From the definition of $[\![q]\!]_G$, there are o' and d such that $\langle o, o', t, t+d\rangle \in [\![\mathsf{path}]\!]_G$.
  So by IH, there is a τ s.t. $t \in \tau$ and $\langle o, o', \tau, d\rangle \in (\!|\mathsf{path}|\!)^t_G$.
  Therefore $\langle o, o, \tau, 0\rangle \in (\!|q|\!)^t_G$, from the definition of $(\!|q|\!)^t_G$.
- For II, let $\langle o, o, \tau, 0\rangle \in (\!|q|\!)^t_G$.
  From the definition of $(\!|q|\!)^t_G$, there are o' and d s.t. $\langle o, o', \tau, d\rangle \in (\!|\mathsf{path}|\!)^t_G$.
  Now take any $t \in \tau$.
  By IH, $\langle o, o', t, t+d\rangle \in [\![\mathsf{path}]\!]_G$.
  Therefore $\langle o, o, t, 0\rangle \in [\![q]\!]_G$, from the definition of $[\![q]\!]_G$.

- $q = \neg\mathsf{test}$.
  From the above definitions, we have:

$$[\![q]\!]_G = (\{\langle o, o\rangle \mid o \in N \cup E\} \times \mathcal{T}_G \times \{0\}) \setminus [\![\mathsf{test}]\!]_G$$
$$(\!|q|\!)^t_G = \bigcup_{o \in N \cup E} \left\{ \langle o, o, \tau, 0\rangle \mid \tau \in \mathsf{compl}\left(\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\}, \mathcal{T}_G\right) \right\}$$

- For I, let $\mathbf{v} = \langle o, o, t, 0\rangle \in [\![q]\!]_G$.
  From the definition of $[\![q]\!]_G$, $\mathbf{v} \notin [\![\mathsf{test}]\!]_G$.
  So

$$t \notin \{t' \mid \langle o, o, t', 0\rangle \in [\![\mathsf{test}]\!]_G\} \tag{7}$$

  Now by IH, together with Lemma 9, we get:

$$\langle o, o, t', 0\rangle \in [\![\mathsf{test}]\!]_G \text{ iff } t' \in \tau' \text{ for some } \tau' \text{ s.t. } \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G \tag{8}$$

  So from (7) and (8):

$$t \notin \bigcup\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\}$$

  Therefore

$$t \in \mathcal{T}_G \setminus \bigcup\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\} \tag{9}$$

  So $t \in \tau$ for some $\tau \in \mathsf{compl}\,(\bigcup\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\}, \mathcal{T}_G)$.
  And $\langle o, o, \tau, 0\rangle \in (\!|q|\!)^t_G$, from the definition of $(\!|q|\!)^t_G$.
- For II, let $\langle o, o, \tau, 0\rangle \in (\!|q|\!)^t_G$.
  And take any $t \in \tau$.
  From the definition of $(\!|q|\!)^t_G$:

$$t \in \mathcal{T}_G \setminus \bigcup\{\tau' \mid \langle o, o, \tau', 0\rangle \in (\!|\mathsf{test}|\!)^t_G\}$$

  Together with (8), this implies

$$\langle o, o, t, 0\rangle \notin [\![\mathsf{test}]\!]_G$$

  Therefore $\langle o, o, t, 0\rangle \in [\![q]\!]_G$, from the definition of $[\![q]\!]_G$.

- $q = \mathsf{path}_1/\mathsf{path}_2$.
  From the above definitions, we have:

$$\begin{aligned}
[\![q]\!]_G &= \{\langle o_1, o_3, t, d_1 + d_2\rangle \mid \exists o_2\colon \langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G \wedge \langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G\} \\
(\!|q|\!)^t_G &= \Big\{\langle o_1, o_3, ((\tau_1 + d_1) \cap \tau_2) - d_1, d_1 + d_2\rangle \mid \\
&\qquad \exists o_2\colon \langle o_1, o_2, \tau_1, d_1\rangle \in (\!|\mathsf{path}_1|\!)^t_G \wedge \langle o_2, o_3, \tau_2, d_2\rangle \in (\!|\mathsf{path}_2|\!)^t_G \wedge (\tau_1 + d_1) \cap \tau_2 \neq \emptyset \Big\}
\end{aligned}$$

  - For I, let $\mathbf{v} = \langle o_1, o_3, t, d\rangle \in [\![q]\!]_G$.
    Fom the definition of $[\![q]\!]_G$, there are o_2, d_1 and d_2 such that $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, $\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$ and $d = d_1 + d_2$.
    By IH, because $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, there is a τ_1 such that $t \in \tau_1$ and

$$\langle o_1, o_2, \tau_1, d_1\rangle \in (\!|\mathsf{path}_1|\!)^t_G \tag{10}$$

    And similarly, because $\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$, there is a τ_2 such that $t + d_1 \in \tau_2$ and

$$\langle o_2, o_3, \tau_2, d_2\rangle \in (\!|\mathsf{path}_2|\!)^t_G \tag{11}$$

    From $t \in \tau_1$, we get

$$t + d_1 \in \tau_1 + d_1 \tag{12}$$

    Together with the fact that $t + d_1 \in \tau_2$, this implies

$$\tau_1 + d_1 \cap \tau_2 \neq \emptyset \tag{13}$$

    So from (10), (11), (13) and the definition of $(\!|q|\!)^t_G$,

$$\langle o_1, o_2, ((\tau_1 + d_1) \cap \tau_2) - d_1, d_1 + d_2\rangle \in (\!|q|\!)^t_G$$

    which proves Ia.
    And in order to prove Ib, we only need to show that

$$t \in ((\tau_1 + d_1) \cap \tau_2) - d_1$$

    We know that $t \in \tau_1$, therefore

$$t + d_1 \in \tau_1 + d_1$$

    Together with the fact that $t + d_1 \in \tau_2$, this yields

$$\begin{aligned}
t + d_1 &\in (\tau_1 + d_1) \cap \tau_2 \\
t &\in ((\tau_1 + d_1) \cap \tau_2) - d_1
\end{aligned}$$

  - For II, let $\mathbf{u} = \langle o_1, o_3, \tau, d\rangle \in (\!|q|\!)^t_G$, and let $t \in \tau$.
    We show that $\langle o_1, o_3, t, t + d\rangle \in [\![q]\!]_G$.
    Because $\mathbf{u} \in (\!|q|\!)^t_G$, from the definition of $(\!|q|\!)^t_G$, there are τ_1, τ_2, d_1, d_2 and o_2 s.t.:
    (i) $d = d_1 + d_2$
    (ii) $\tau = ((\tau_1 + d_1) \cap \tau_2) - d_1$
    (iii) $\langle o_1, o_2, \tau_1, d_1\rangle \in (\!|\mathsf{path}_1|\!)^t_G$
    (iv) $\langle o_2, o_3, \tau_2, d_2\rangle \in (\!|\mathsf{path}_2|\!)^t_G$

Since $t \in \tau$, from ii, we have

$$\begin{aligned} t &\in ((\tau_1 + d_1 \cap \tau_2) - d_1 && (14)\\ t + d_1 &\in (((\tau_1 + d_1 \cap \tau_2) - d_1) + d_1 && (15)\\ t + d_1 &\in (\tau_1 + d_1) \cap \tau_2 && (16)\\ t + d_1 &\in \tau_1 + d_1 && (17)\\ t &\in \tau_1 && (18) \end{aligned}$$

From iii, by IH, for any $t' \in \tau_1$

$$\langle o_1, o_2, t' + d_1\rangle \in [\![q]\!]_G$$

In particular, from (18)

$$\langle o_1, o_2, t, t + d_1\rangle \in [\![q]\!]_G \tag{19}$$

And from iv, by IH, for any $t'' \in \tau_2$

$$\langle o_2, o_3, t'', t'' + d_2\rangle \in [\![q]\!]_G$$

In particular, from (16)

$$\langle o_2, o_3, t + d_1, (t + d_1) + d_2\rangle \in [\![q]\!]_G \tag{20}$$

So from (19), (20) and the definition of $[\![q]\!]_G$

$$\langle o_1, o_3, t, t + d_1 + d_2\rangle \in [\![q]\!]_G$$

◀

A.2.2 In $\mathcal{U}^d$

A.2.2.1 Definition

We start with the case where q is an expression for the symbol test in the grammar of Section 3.

$$\begin{aligned} (\!|< k|\!)^d_G &= \{\langle o, o, t, [0,0]\rangle \mid o \in (N \cup E), t \in \mathcal{T}_G, t < k\}\\ (\!|pred|\!)^d_G &= \{\langle o, o, t, [0,0]\rangle \mid t \in \tau \text{ for some } \tau \in \mathsf{val}(o, pred)\}\\ (\!|\mathrm{F}|\!)^d_G &= \{\langle v, e, t, [0,0]\rangle \mid \mathsf{src}(e) = v, t \in \mathcal{T}_G\} \cup \{\langle e, v, t, [0,0]\rangle \mid \mathsf{tgt}(e) = v, t \in \mathcal{T}_G\}\\ (\!|\mathrm{B}|\!)^d_G &= \{\langle v, e, t, [0,0]\rangle \mid \mathsf{tgt}(e) = v, t \in \mathcal{T}_G\} \cup \{\langle e, v, t, [0,0]\rangle \mid \mathsf{src}(e) = v, t \in \mathcal{T}_G\}\\ (\!|(?\mathsf{path})|\!)^d_G &= \{\langle o, o, t, [0,0]\rangle \mid \exists o', \delta\colon \langle o, o', t, \delta\rangle \in (\!|\mathsf{path}|\!)^d_G\}\\ (\!|\mathsf{test}_1 \vee \mathsf{test}_2|\!)^d_G &= (\!|\mathsf{test}_1|\!)^d_G \cup (\!|\mathsf{test}_2|\!)^d_G\\ (\!|\mathsf{test}_1 \wedge \mathsf{test}_2|\!)^d_G &= (\!|\mathsf{test}_1|\!)^d_G \cap (\!|\mathsf{test}_2|\!)^d_G\\ (\!|\neg\mathsf{test}|\!)^d_G &= \Big\{\langle o, o, t, [0,0]\rangle \mid o \in N \cup E, t \in \mathcal{T}_G \setminus \{t' \mid \langle o, o, t', [0,0]\rangle \in (\!|\mathsf{test}|\!)^d_G\}\Big\} \end{aligned}$$

Next, we consider the operators $(\mathsf{path}_1 + \mathsf{path}_2)$, $(\mathsf{path}[m, \_])$ and $(\mathsf{path}[m, n])$. For these cases, $(\!|q|\!)^{td}_G$ is once again defined analogously to $[\![q]\!]_G$, in terms of temporal join (a.k.a. $\mathsf{path}_1/\mathsf{path}_2$) and set union.
We only write the definitions here for the sake of completeness:

$$\begin{aligned} (\!|\mathsf{path}_1 + \mathsf{path}_2|\!)^{td}_G &= (\!|\mathsf{path}_1|\!)^{td}_G \cup (\!|\mathsf{path}_2|\!)^{td}_G\\ [\![\mathsf{path}[m, n]]\!]_G &= \bigcup_{k=m}^{n} (\!|\mathsf{path}^k|\!)^{td}_G\\ [\![\mathsf{path}[m, \_]]\!]_G &= \bigcup_{k \geq m} (\!|\mathsf{path}^k|\!)^{td}_G \end{aligned}$$

The only remaining operators are temporal join ($\mathsf{path}_1/\mathsf{path}_2$) and temporal navigation (T_δ), already defined in the article.
We reproduce here these two definition for convenience:

$$
\begin{aligned}
(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^d_G = \ & \Big\{\langle o_1, o_3, t_1, \delta_2 + t_2 - t_1\rangle \mid \\
& \qquad \exists o_2 \colon \langle o_1, o_2, t_1, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^d_G \wedge \langle o_2, o_3, t_2, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^d_G \wedge t_2 \in t_1 + \delta_1\Big\} \\
(\!|\mathrm{T}_\delta|\!)^d_G = \ & \{\langle o, o, t, ((\delta + t) \cap \mathcal{T}_G) - t\rangle \mid o \in N \cup E, t \in \mathcal{T}_G, (\delta + t) \cap \mathcal{T}_G \neq \emptyset\}
\end{aligned}
$$

We also reproduce the alternative characterization of $(\!|\mathrm{T}_\delta|\!)^d_G$ provided in the article, as a unary operator:

$$(\!|q/\mathrm{T}_\delta|\!)^d_G = \{\langle o_1, o_2, t, (\delta' + \delta) \cap \mathcal{T}_G\rangle \mid \langle o_1, o_2, t, \delta'\rangle \in (\!|q|\!)^d_G, (t + (\delta' + \delta)) \cap \mathcal{T}_G \neq \emptyset\}$$

A.2.2.2 Correctness

The following result states that the representation $(\!|q|\!)^d_G$ is correct:

▶ **Proposition 11.** *Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TPG and q a TRPQ. Then the unfolding of $(\!|q|\!)^d_G$ is $[\![q]\!]_G$.*

Proof.
Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TG, and let q be a TRPQ.
We show below that:

- **(I)** for any $\langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, there is a $\delta \in \mathsf{intv}(\mathcal{T})$ such that
  - **(a)** $\langle o_1, o_2, t, \delta\rangle \in (\!|q|\!)^d_G$, and
  - **(b)** $d \in \delta$,
- **(II)** for any $\langle o_1, o_2, t, \delta\rangle \in (\!|q|\!)^d_G$ for any $d \in \delta$, $\langle o_1, o_2, t, d\rangle$ is in $[\![q]\!]_G$.

We proceed once again by induction on the structure of q.
If q is of the form pred, $< k$, F, B, ($\mathsf{test} \vee \mathsf{test}$), ($\mathsf{test} \wedge \mathsf{test}$), $\neg\mathsf{test}$, ($\mathsf{path} + \mathsf{path}$), $\mathsf{path}[m, n]$ or $\mathsf{path}[m, \_]$, then I and II immediately follow from the definitions of $[\![q]\!]_G$ and $(\!|q|\!)^d_G$.
If q is of the form ($?\mathsf{path}$), then the proof is nearly identical to one already provided for $(\!|(?\mathsf{path})|\!)^t_G$.
So we focus below on the two remaining cases:

- $q = \mathrm{T}_\delta$.
  From the above definitions, we have:

$$
\begin{aligned}
[\![q]\!]_G &= \{\langle o, o, t, d\rangle \mid o \in (N \cup E), t \in \mathcal{T}_G, d \in \delta, t + d \in \mathcal{T}_G\} \\
(\!|q|\!)^d_G &= \{\langle o, o, t, ((\delta + t) \cap \mathcal{T}_G) - t\rangle \mid o \in N \cup E, t \in \mathcal{T}_G, (\delta + t) \cap \mathcal{T}_G \neq \emptyset\}
\end{aligned}
$$

  - For I, let $\mathbf{v} = \langle o, o, t, d\rangle \in [\![q]\!]_G$.
    And let $\mathbf{u} = \langle o, o, t, ((\delta + t) \cap \mathcal{T}_G) - t\rangle$ in $\mathcal{U}^d$.
    For Ia we show that $\mathbf{u} \in (\!|q|\!)^d_G$.
    From $\mathbf{v} \in [\![q]\!]_G$, we get $o \in N \cup E$ and $t \in \mathcal{T}_G$.
    Besides, because $\mathbf{v} \in [\![q]\!]_G$ still,

$$t + d \in \mathcal{T}_G \tag{21}$$

    and

$$d \in \delta \tag{22}$$
$$t + d \in t + \delta \tag{23}$$

So from (21) and (23)

$$t + d \in (\delta + t) \cap \mathcal{T}_G \quad (24)$$
$$(\delta + t) \cap \mathcal{T}_G \neq \emptyset \quad (25)$$

Together with the definition of $(\!|q|\!)^d_G$, this implies $\mathbf{u} \in (\!|q|\!)^d_G$, which concludes the proof for Ia.
Finally, from (24), we get

$$t + d - t \in ((\delta + t) \cap \mathcal{T}_G) - t \quad (26)$$
$$d \in ((\delta + t) \cap \mathcal{T}_G) - t \quad (27)$$

which proves Ib.

- For II, let $\mathbf{u} = \langle o, o, t, ((\delta + t) \cap \mathcal{T}_G) - t\rangle \in (\!|q|\!)^d_G$, and let $d \in ((\delta + t) \cap \mathcal{T}_G) - t$.
From $\mathbf{u} \in (\!|q|\!)^d_G G$, we get $o \in N \cup E$ and $t \in \mathcal{T}_G$.
So to conclude the proof, it is sufficient to show that *(i)* $d \in \delta$ and *(ii)* $t + d \in \mathcal{T}_G$.
By assumption, we have

$$d \in ((\delta + t) \cap \mathcal{T}_G) - t \quad (28)$$
$$d + t \in (\delta + t) \cap \mathcal{T}_G \quad (29)$$
$$d + t \in \mathcal{T}_G \quad (30)$$

which proves *(ii)* .
And from (29), we also get

$$d + t \in \delta + t$$
$$d + t - t \in (\delta + t) - t$$
$$d \in \delta$$

which proves *(i)* .

- $q = \mathsf{path}_1/\mathsf{path}_2$.
From the above definitions, we have:

$$\begin{aligned} [\![q]\!]_G &= \{\langle o_1, o_3, t, d_1 + d_2\rangle \mid \exists o_2 \colon \langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G \wedge \langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G\} \\ (\!|q|\!)^d_G &= \left\{\langle o_1, o_3, t_1, \delta_2 + t_2 - t_1\rangle \mid \exists o_2 \colon \langle o_1, o_2, t_1, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^d_G \wedge \langle o_2, o_3, t_2, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^d_G \wedge t_2 \in t_1 + \delta_1\right\} \end{aligned}$$

- For I, let $\mathbf{v} = \langle o_1, o_3, t, d\rangle \in [\![q]\!]_G$.
Fom the definition of $[\![q]\!]_G$, there are o_2, d_1 and d_2 such that $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, $\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$ and $d = d_1 + d_2$.
By IH, because $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, there is a δ_1 such that $d_1 \in \delta_1$ and

$$\langle o_1, o_2, t, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^d_G \quad (31)$$

And similarly, because $\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$, there is a δ_2 such that $d_2 \in \delta_2$ and

$$\langle o_2, o_3, t + d_1, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^d_G \quad (32)$$

Next, since $d \in \delta_1$

$$t + d_1 \in t + \delta_1 \quad (33)$$

So from (31), (32), (33) and the definition of $(\!|q|\!)^d_G$ (replacing t_1 with t and t_2 with $t+d_1$), we get

$$\langle o_1, o_2, t, \delta_2 + (t+d_1) - t\rangle \in (\!|q|\!)^d_G$$

which proves Ia.
And in order to prove Ib, we only need to show that $d \in \delta_2 + (t+d_1) - t$, or in other words that

$$d \in \delta_2 + d_1$$

We know that

$$d_2 \in \delta_2 \quad (34)$$
$$d_2 + d_1 \in \delta_2 + d_1 \quad (35)$$

Together with the fact that $d = d_1 + d_2$, this concludes the proof for Ib.

- For II, let $\mathbf{u} = \langle o_1, o_3, t_1, \delta\rangle \in (\!|q|\!)^d_G$, and let $d \in \delta$.

  Because $\mathbf{u} \in (\!|q|\!)^d_G$, from the definition of $(\!|q|\!)^d_G$, there are δ_1, δ_2, t_2 and o_2 s.t.:

  (i) $\delta = \delta_2 + t_2 - t_1$
  (ii) $t_2 \in t_1 + \delta_1$
  (iii) $\langle o_1, o_2, t_1, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^d_G$
  (iv) $\langle o_2, o_3, t_2, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^d_G$

  From i and ii, we get

$$\begin{aligned}\delta &= \delta_2 + (t_1 + \delta_1) - t_1\\ &= \delta_2 + \delta_1\end{aligned}$$

  Together with $d \in \delta$, this implies that there are $d_1 \in \delta_1$ and $d_2 \in \delta_2$ such that $d = d_1 + d_2$.
  Next, because $d_1 \in \delta_1$, from iii, by IH

$$\langle o_1, o_2, t_1, t_1 + d_1\rangle \in [\![q]\!]_G \quad (36)$$

  And similarly, because $d_2 \in \delta_2$, from iv

$$\langle o_2, o_3, t_2, t_2 + d_2\rangle \in [\![q]\!]_G \quad (37)$$

  So from (36), (37) and the definition of $[\![q]\!]_G$

$$\langle o_1, o_3, t_1, d_1 + d_2\rangle \in [\![q]\!]_G \quad (38)$$

  Together with the fact that $d = d_1 + d_2$, this concludes the proof for II.

◀

A.2.3 In $\mathcal{U}^{td}$

A.2.3.1 Definition

We start with the case where q is an expression for the symbol test in the grammar of Section 3.

As a consequence of Lemma 9, $(\!|q|\!)^{td}_G$ can be trivially defined out of $(\!|q|\!)^{t}_G$ by replacing the distance 0 with the interval $[0,0]$, i.e.

$$(\!|\mathsf{test}|\!)^{td}_G = \{\langle o, o, \tau, [0,0]\rangle \mid \{\langle o, o, \tau, 0\rangle \in (\!|\mathsf{test}|\!)^{t}_G\}$$

Next, if q is of the form $(\mathsf{path}_1 + \mathsf{path}_2)$, $(\mathsf{path}[m, \_])$ or $(\mathsf{path}[m,n])$, then the definition of $(\!|q|\!)^{td}_G$ is once again nearly identical to the one of $[\![q]\!]_G$:

$$\begin{aligned}
(\!|\mathsf{path}_1 + \mathsf{path}_2|\!)^{td}_G &= (\!|\mathsf{path}_1|\!)^{td}_G \cup (\!|\mathsf{path}_2|\!)^{td}_G \\
[\![\mathsf{path}[m,n]]\!]_G &= \bigcup_{k=m}^{n} (\!|\mathsf{path}^k|\!)^{td}_G \\
[\![\mathsf{path}[m,_]]\!]_G &= \bigcup_{k\geq m} (\!|\mathsf{path}^k|\!)^{td}_G
\end{aligned}$$

The only remaining operators are temporal join $(\mathsf{path}_1/\mathsf{path}_2)$ and temporal navigation (T_δ), already defined in the article, and reproduced here for convenience:

$$\begin{aligned}
(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^{td}_G &= \bigcup\{\mathbf{u}_1 \boxtimes \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^{td}_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^{td}_G\} \\
(\!|\mathrm{T}_\delta|\!)^{td}_G &= \bigcup_{o \in N \cup E} \{\langle o, o, \mathcal{T}_G, \delta\rangle \boxtimes \langle o, o, \mathcal{T}_G, [0,0]\rangle\}
\end{aligned}$$

where $\mathbf{u}_1 \boxtimes \mathbf{u}_2$ is defined as follows.
Let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ and $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2\rangle$.
Define τ_2' as

$$\tau_2' = (\tau_1 + \delta_1) \cap \tau_2$$

If $o_2 \neq o_3$ or $\tau_2' = \emptyset$, then $\mathbf{u}_1 \boxtimes \mathbf{u}_2 = \emptyset$.
Otherwise, let:

$$\tau = (\tau_2' \ominus \delta_1) \cap \tau_1$$
$$b = b_{\tau_2'} - b_{\delta_1}$$
$$e = e_{\tau_2'} - e_{\delta_1}$$

And for every $t \in \tau$, let

$$\delta(t) = {}_{\delta_1}\lfloor\, b_{\delta_1} + \max(0, b-t),\ e_{\delta_1} - \max(0, t-e)\,\rfloor_{\delta_1}$$

Then

$$\mathbf{u}_1 \boxtimes \mathbf{u}_2 = \{\langle o_1, o_4, [t,t], \delta(t) + \delta_2\rangle \mid t \in \tau\}$$

A.2.3.2 Correctness

We start with a lemma:

▶ **Lemma 12.** *Let* $\alpha, \beta \in \mathsf{intv}(\mathcal{T})$*. Then*

$$\beta \ominus \alpha = \{t \mid (t+\alpha) \cap \beta \neq \emptyset\}$$

Next, if $\mathbf{u} = \langle o_1, o_2, \tau, \delta\rangle \in \mathcal{U}^{td}$, we call *temporal relation induced by* $\mathbf{u}$ the set

$$\{(t, t+d) \mid t \in \tau, d \in \delta\}$$

We also define the binary operator $\bowtie$: $(\mathcal{T} \times \mathcal{T}) \times (\mathcal{T} \times \mathcal{T}) \to (\mathcal{T} \times \mathcal{T})$ as in the article, i.e.

$$R_1 \bowtie R_2 = \{t_1, t_3 \mid (t_1, t_2) \in R_1 \text{ and } (t_2, t_3) \in R_2 \text{ for some } t_2\}$$

We can now formulate the following lemma:

▶ **Lemma 13.** *Let* $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ *and* $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2\rangle$ *be two tuples in* $\mathcal{U}^{td}$ *such that* $o_2 = o_3$*. And for* $i \in \{1, 2\}$*, let* R_i *denote the temporal relation induced by* $\mathbf{u}_i$*. Then*

$$R_1 \bowtie R_2 = \bigcup_{\langle o_1, o_2, \tau, \delta\rangle \in \mathbf{u}_1 \overline{\bowtie} \mathbf{u}_2} \{(t, t+d) \mid t \in \tau, d \in \delta\}$$

Proof. $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ and $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2\rangle$ be two tuples in $\mathcal{U}^{td}$ such that $o_2 = o_3$. And for $i \in \{1, 2\}$, let R_i denote the temporal relation induced by $\mathbf{u}_i$.

We show that:

- **(I)** **(a)** If $\tau_2' = \emptyset$, then $\mathsf{dom}(R_1 \bowtie R_2) = \emptyset$,
  - **(b)** otherwise $\tau = \mathsf{dom}(R_1 \bowtie R_2)$,
- **(II)** for each $t \in \tau$,

$$t + \delta(t) + \delta_2 = \{t' \mid (t, t') \in R_1 \bowtie R_2\}$$

We start with I.
From the defnition of "+" (applied to two intervals):

$$\tau_1 + \delta_1 = \{t + d \mid t \in \tau_1, d_1 \in \delta_1\} \tag{39}$$

So from the definition of R_1

$$\tau_1 + \delta_1 = \mathsf{range}(R_1) \tag{40}$$

Since $\tau_2 = \mathsf{dom}(R_2)$, this implies

$$(\tau_1 + \delta_1) \cap \tau_2 = \mathsf{range}(R_1) \cap \mathsf{dom}(R_2) \tag{41}$$
$$\tau_2' = \mathsf{range}(R_1) \cap \mathsf{dom}(R_2) \tag{42}$$

If $\mathsf{range}(R_1) \cap \mathsf{dom}(R_2) = \emptyset$, then $\mathsf{dom}(R_1 \bowtie R_2) = \emptyset$, immediately from the definition of $\bowtie$, which concludes the proof of Ia.
Otherwise, from Lemma 12,

$$\tau_2' \ominus \delta_1 = \{t \mid (t + \delta_1) \cap \tau_2' \neq \emptyset\} \tag{43}$$

So from (42)

$$\begin{aligned}
\tau_2' \ominus \delta_1 &= \{t \mid (t + \delta_1) \cap \mathsf{range}(R_1) \cap \mathsf{dom}(R_2) \neq \emptyset\}\\
(\tau_2' \ominus \delta_1) \cap \tau_1 &= \{t \in \tau_1 \mid (t + \delta_1) \cap \mathsf{range}(R_1) \cap \mathsf{dom}(R_2) \neq \emptyset\}\\
(\tau_2' \ominus \delta_1) \cap \tau_1 &= \mathsf{dom}(R_1 \bowtie R_2)\\
\tau &= \mathsf{dom}(R_1 \bowtie R_2)
\end{aligned}$$

which proves Ib.

Now for II, let $t \in \tau$.
We show below that *(i)* $t + \delta(t) = \{t' \mid (t, t') \in R_1$ and $t' \in \mathsf{range}(R_1) \cap \mathsf{dom}(R_2)\}$.
Together with the definition of $\bowtie$ (and the fact that $t + \delta(t)$ is an interval), this proves II.

We only prove the result for the case where τ, τ_2' and δ_1 are closed-closed intervals (the proof for the other 63 cases is symmetric).
First, from Ib and the assumption that $t \in \tau$, we have $t \in \tau_1$. So from the definition of R_1,

$$t + \delta_1 = \{t' \mid (t, t') \in R_1\} \tag{44}$$

Together with (42), this means that *(i)* is equivalent to *(ii)* $t + \delta(t) = \{(t + \delta_1) \cap \tau_2'\}$. So in order to prove II (and conclude our proof), it is sufficient to prove *(ii)* .

Now since $t \in \tau$, from Ib and the definition of τ_2', we have $(t + \delta(t)) \cap \tau_2' \neq \emptyset$. And since $\delta(t)$ and τ_2' are intervals, $(t + \delta(t)) \cap \tau_2'$ is an also an interval. So in order to prove *(ii)* , it is sufficient to show that $t + b_{\delta(t)}$ (resp. $t + e_{\delta(t)}$) is the smallest (resp. greatest) value in $(t + \delta_1) \cap \tau_2'$. We only prove the result for $t + b_{\delta(t)}$ (the proof for $t + e_{\delta(t)}$) is symmetric. We consider two cases.

- If $b \leq t$, then

$$b_{\tau_2'} - b_{\delta_1} \leq t \quad \text{from the definition of } b \tag{45}$$
$$b_{\tau_2'} - b_{\delta_1} + b_{\delta_1} \leq t + b_{\delta_1} \tag{46}$$
$$b_{\tau_2'} \leq t + b_{\delta_1} \tag{47}$$

  And because $t \in \tau$

$$t \leq e_{\tau} \tag{48}$$
$$t \leq e_{\tau_2'} - b_{\delta_1} \quad \text{from the definition of } \tau \tag{49}$$
$$t + b_{\delta_1} \leq e_{\tau_2'} - b_{\delta_1} + b_{\delta_1} \tag{50}$$
$$t + b_{\delta_1} \leq e_{\tau_2'} \tag{51}$$

  So from (47) and (51)

$$t + b_{\delta_1} \in \tau_2' \tag{52}$$

  Next, since $b \leq t$ (by assumption), we have

$$b - t \leq 0$$
$$\max(0, b - t) = 0$$

  So from the definition of $\delta(t)$

$$b_{\delta(t)} = b_{\delta_1} \tag{53}$$

  Therefore $t + b_{\delta(t)}$ is the smallest value in $t + \delta_1$.
  So from (52), it is also the smallest value in $t + \delta_1 \cap \tau_2'$, which concludes the proof for this case.
- If $b > t$, then

$$b - t > 0 \tag{54}$$
$$\max(0, b - t) = b - t \tag{55}$$

  So from the definition of $\delta(t)$

$$b_{\delta(t)} = b_{\delta_1} + b - t \tag{56}$$

  Besides, from (54)

$$b - t + b_{\delta_1} > b_{\delta_1} \tag{57}$$

So from (56) and (57)

$$b_{\delta(t)} > b_{\delta_1} \tag{58}$$

Next, since $t \in \tau$

$$b_\tau \leq t \tag{59}$$

And from the definition of τ

$$b_{\tau'_2} - e_{\delta_1} \leq b_\tau \tag{60}$$

So from (59) and (60)

$$\begin{aligned}
b_{\tau'_2} - e_{\delta_1} &\leq t && (61)\\
b_{\tau'_2} - t &\leq e_{\delta_1} && (62)\\
b_{\tau'_2} - t + b_{\delta_1} - b_{\delta_1} &\leq e_{\delta_1} && (63)\\
b_{\delta_1} + (b_{\tau'_2} - b_{\delta_1}) - t &\leq e_{\delta_1} && (64)\\
b_{\delta_1} + b - t &\leq e_{\delta_1} && \text{from the definition of } b \quad (65)\\
b_{\delta(t)} &\leq e_{\delta_1} && \text{from (56)} \quad (66)
\end{aligned}$$

Therefore from (58) and (66)

$$\begin{aligned}
b_{\delta(t)} &\in \delta_1 && (67)\\
t + b_{\delta(t)} &\in t + \delta_1 && (68)
\end{aligned}$$

Finally, from (56) still,

$$\begin{aligned}
t + b_{\delta(t)} &= t + b_{\delta_1} + b - t && (69)\\
&= t + b_{\delta_1} + b_{\tau'_2} - b_{\delta_1} - t && \text{from the definition of } b \quad (70)\\
&= b_{\tau'_2} && (71)
\end{aligned}$$

So $t + b_{\delta(t)}$ is the smallest value in τ'_2.
Together with (68), this concludes the proof for this case.

◀

The following result states that the representation $(\!|q|\!)^{td}_G$ is correct:

▶ **Proposition 14.** *Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TPG and q a TRPQ. Then the unfolding of $(\!|q|\!)^{td}_G$ is $[\![q]\!]_G$.*

Proof.
Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TG, and let q be a TRPQ.
We show below that:

- **(I)** for any $\langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, there are $\tau, \delta \in \mathsf{intv}(\mathcal{T})$ such that
  - **(a)** $\langle o_1, o_2, \tau, \delta\rangle \in (\!|q|\!)^{td}_G$,
  - **(b)** $t \in \tau$, and
  - **(c)** $d \in \delta$.
- **(II)** for any $\langle o_1, o_2, \tau, \delta\rangle \in (\!|q|\!)^{td}_G$ for any $(t, d) \in \tau \times \delta$, $\langle o_1, o_2, t, d\rangle$ is in $[\![q]\!]_G$.

We proceed once again by induction on the structure of q.

If q is of the form *pred*, F, B, $(\mathsf{test} \vee \mathsf{test})$, $(\mathsf{path} + \mathsf{path})$, $\mathsf{path}[m, n]$ or $\mathsf{path}[m, \_]$, then I and II immediately follow from the definitions of $[\![q]\!]_G$ and $(\!|q|\!)^{td}_G$.

If q is of the form $\mathsf{test} \wedge \mathsf{test}$, $\neg\mathsf{test}$ or $(?\mathsf{path})$, then the proof is nearly identical to the one already provided for $(\!|q|\!)^{t}_G$.

So we focus below on the two remaining cases:

- $q = \mathsf{path}_1/\mathsf{path}_2$.

  From the above definitions, we have:

$$\begin{aligned} [\![q]\!]_G &= \{\langle o_1, o_3, t, d_1 + d_2\rangle \mid \exists o_2 \colon \langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G \wedge \langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G\} \\ (\!|\mathsf{path}_1/\mathsf{path}_2|\!)^{td}_G &= \bigcup\{\mathbf{u}_1 \bowtie \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^{td}_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^{td}_G\} \end{aligned}$$

  - For I, let $\mathbf{v} = \langle o_1, o_3, t, d\rangle \in [\![q]\!]_G$.

    Fom the definition of $[\![q]\!]_G$, there are o_2, d_1 and d_2 such that $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, $\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$ and $d = d_1 + d_2$.

    By IH, because $\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G$, there are τ_1 and δ_1 such that $t \in \tau_1$, $d_1 \in \delta_1$ and

$$\langle o_1, o_2, \tau_1, \delta_1\rangle \in (\!|\mathsf{path}_1|\!)^{td}_G \tag{72}$$

    Let R_1 be the temporal relation induced by this tuple $\langle o_1, o_2, \tau_1, \delta_1\rangle$.

    Since $t \in \tau_1$ and $d_1 \in \delta_1$, we have

$$(t, t + d_1) \in R_1 \tag{73}$$

    Similarly, because $\langle o_2, o_3, t+d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G$, there are τ_2 and δ_2 such that $t+d_1 \in \tau_2$, $d_2 \in \delta_2$ and

$$\langle o_2, o_3, \tau_2, \delta_2\rangle \in (\!|\mathsf{path}_2|\!)^{td}_G \tag{74}$$

    Let R_2 be the temporal relation induced by this tuple $\langle o_2, o_3, \tau_2, \delta_2\rangle$.

    Since $t + d_1 \in \tau_2$ and $d_2 \in \delta_2$, we have

$$(t + d_1, t + d_1 + d_2) \in R_2 \tag{75}$$

    So from (73), (75) and Lemma 13, there are τ and δ such that $\langle o_1, o_3, \tau, \delta\rangle \in u_1 \bowtie u_2$, $t \in \tau$ and $d_1 + d_2 = d \in \delta$, which concludes the proof for I.

  - For II, let $\mathbf{u} = \langle o_1, o_3, t_1, \delta\rangle \in (\!|q|\!)^{td}_G$, and let $(t, d) \in \tau \times \delta$.

    Because $\mathbf{u} \in (\!|q|\!)^{td}_G$, from the definition of $(\!|q|\!)^{td}_G$, there are $\mathbf{u}_1$ and $\mathbf{u}_2$ s.t.:

    (i) $\mathbf{u} \in \mathbf{u}_1 \bowtie \mathbf{u}_2$

    (ii) $\mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^{td}_G$

    (iii) $\mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^{td}_G$

    Let R_i be the temporal relation induced by u_i for $i \in \{1, 2\}$.

    From i, and Lemma 13,

$$(t, t + d) \in R_1 \bowtie R_2 \tag{76}$$

    Now let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ and $\mathbf{u}_2 = \langle o_2, o_3, \tau_2, \delta_2\rangle$ for some $o_2, \tau_1, \tau_2, \delta_1$ and δ_2.

    From (76) and the definition of $\bowtie$, there must be d_1 and d_2 s.t. $d = d_1 + d_2$, $t \in \tau_1, d_1 \in \delta_1$, $t + d_1 \in \tau_2$ and $d_2 \in \delta_2$.

    So from ii, and iii, by IH

$$\langle o_1, o_2, t, d_1\rangle \in [\![\mathsf{path}_1]\!]_G \tag{77}$$

$$\langle o_2, o_3, t + d_1, d_2\rangle \in [\![\mathsf{path}_2]\!]_G \tag{78}$$

So from (77), (78) and the definition of $[\![q]\!]_G$

$$\langle o_1, o_3, t, d_1 + d_2\rangle \in [\![q]\!]_G,$$

which concludes the proof for II. ◀

A.2.4 In $\mathcal{U}^s$

A.2.4.1 Definition

If q is an expression for the symbol test in the grammar of Section 3, then the definition of $(\!|q|\!)^s_G$ is nearly identical to the one of $(\!|q|\!)^{td}_G$, extending each tuple $\{\langle o, o, \tau, [0, 0]\rangle$ with b_τ and e_τ, i.e.

$$(\!|\mathsf{test}|\!)^s_G = \{\langle o, o, \tau, [0, 0], b_\tau, e_\tau\rangle \mid \{\langle o, o, \tau, [0, 0]\rangle \in (\!|\mathsf{test}|\!)^{td}_G\}$$

Next, if q is of the form $(\mathsf{path}_1 + \mathsf{path}_2)$, $(\mathsf{path}[m, \_])$ or $(\mathsf{path}[m, n])$, then the definition of $(\!|q|\!)^{td}_G$ is once again nearly identical to the one of $[\![q]\!]_G$:

$$\begin{aligned} (\!|\mathsf{path}_1 + \mathsf{path}_2|\!)^s_G = &\quad (\!|\mathsf{path}_1|\!)^s_G \cup (\!|\mathsf{path}_2|\!)^s_G \\ [\![\mathsf{path}[m, n]]\!]_G = &\quad \bigcup_{k=m}^{n} (\!|\mathsf{path}^k|\!)^s_G \\ [\![\mathsf{path}[m, \_]]\!]_G = &\quad \bigcup_{k\geq m} (\!|\mathsf{path}^k|\!)^s_G \end{aligned}$$

So the only remaining operator are temporal join $(\mathsf{path}_1/\mathsf{path}_2)$ and temporal navigation (T_δ), already defined in the article. We reproduce here these two definition for convenience:

$$\begin{aligned} (\!|\mathsf{path}_1/\mathsf{path}_2|\!)^s_G = &\quad \{\mathbf{u}_1 \bowtie \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^s_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^s_G, \mathbf{u}_1 \sim \mathbf{u}_2\} \\ (\!|\mathrm{T}_\delta|\!)^s_G = &\quad \{\langle o, o, \mathcal{T}_G, \delta, b_{\mathcal{T}_G}, e_{\mathcal{T}_G}\rangle \bowtie \langle o, o, \mathcal{T}_G, [0, 0], b_{\mathcal{T}_G}, e_{\mathcal{T}_G}\rangle \mid o \in N \cup E\} \end{aligned}$$

where $\mathbf{u}_1 \sim \mathbf{u}_2$ and $\mathbf{u}_1 \bowtie \mathbf{u}_2$ are defined as follows.
Let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1, b_1, e_1\rangle$ and $\mathbf{u}_2 = \langle o_3, o_4, \tau_2, \delta_2, e_2, b_2\rangle$.
Define

$$\delta'_1 = {}_{\delta_1}\lfloor\, b_{\delta_1} + \max(0, b_1 - b_{\tau_1}), e_{\delta_1} - \max(0, e_{\tau_1} - e_1)\,\rfloor_{\delta_1}$$

and

$$\tau = (((\tau_1 + \delta'_1) \cap \tau_2) \ominus \delta'_1) \cap \tau_1$$

Then $\mathbf{u}_1 \sim \mathbf{u}_2$ iff $o_2 = o_3$ and $\tau \neq \emptyset$.
If $\mathbf{u}_1 \sim \mathbf{u}_2$, then $\mathbf{u}_1 \bowtie \mathbf{u}_2 = \langle o_1, o_4, \tau, \delta_1 + \delta_2, b, e\rangle$, with

$$b = \max(b_1, b_2 - b_{\delta_1})$$
$$e = \min(e_1, e_2 - e_{\delta_1})$$

A.2.4.2 Correctness

We start with two lemmas:

▶ **Lemma 15.** *Let $\mathbf{u} = \langle o_1, o_2, \tau, \delta, b, e\rangle \in \mathcal{U}^s$. Then for any $t_1, t_2 \in \tau$ s.t. $t_1 \leq t_2$:*

$$t_1 + b_{\delta(t_1)} \leq t_2 + b_{\delta(t_2)} \text{ and}$$
$$t_1 + e_{\delta(t_1)} \leq t_2 + e_{\delta(t_2)}$$

▶ **Lemma 16.** *Let* $\mathbf{u} = \langle o_1, o_2, \tau, \delta, b, e\rangle \in \mathcal{U}^s$*. And let* τ' *denote the interval* $(b_\tau + b_{\delta(b_\tau)}, e_\tau + e_{\delta(e_\tau)})$*. Then for any* $t' \in \tau'$*, there is a* $t \in \tau$ *s.t.* $t' \in t + \delta(t)$*.*

Next, similarly to what we did above for $\mathcal{U}^{td}$, if $\mathbf{u} = \langle o_1, o_2, \tau, \delta, b, e\rangle \in \mathcal{U}^s$, we call *temporal relation induced by* $\mathbf{u}$ the set

$$\{(t, t + d) \mid t \in \tau, d \in \delta(t)\}$$

We can now formulate a result analogous to Lemma 13:

▶ **Lemma 17.** *Let* $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}^s$*, and for* $i \in \{1, 2\}$*, let* R_i *denote the temporal relation induced by* $\mathbf{u}_i$*. If* $\mathbf{u}_1 \sim \mathbf{u}_2$ *and* $\mathbf{u}_1 \boxtimes \mathbf{u}_2 = \langle o_1, o_3, \tau, \delta, b, e\rangle$*, then*

$$R_1 \bowtie R_2 = \{(t, t + d) \mid t \in \tau, d \in \delta(t)\}$$

Proof. Let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ and $\mathbf{u}_2 = \langle o_2, o_3, \tau_2, \delta_2\rangle$.
As explained in Section A.1, for $i \in \{1, 2\}$ and $t \in \tau_i$, we use $\delta_i(t)$ for the interval

$${}_{\delta_i}\lfloor\ b_{\delta_i} + \max(0, b_i - t)\ ,\ e_{\delta_i} - \max(0, t - e_i)\ \rfloor_{\delta_i}$$

We need to prove that *(i)* $\tau = \mathsf{dom}(R_1 \bowtie R_2)$ and that *(ii)* for each $t \in \tau$,

$$t + \delta(t) = \{t' \mid (t, t') \in R_1 \bowtie R_2\}$$

The proof of *(i)* is nearly identical to the one provided above for Lemma 13.
For *(ii)* , let $t \in \tau$.
We only provide a proof for the case where τ, δ_1 and δ_2 are closed-closed intervals (the proof for the other 63 cases is symmetric).
Since $t \in \tau$, from the defnition of τ, $t \in \tau_1$.
Therefore from the definition of R_1,

$$t + \delta_1(t) = \{t' \mid (t, t') \in R_1\} \tag{79}$$

So from *(i)* and the fact that $t \in \tau$

$$t + \delta_1(t) \cap \mathsf{dom}(R_2) \neq \emptyset \tag{80}$$

Now let a (resp. z) denote the smallest (resp. largest) value in $t + \delta_1(t) \cap \mathsf{dom}(R_2)$.
Then from (79), a (resp. z) is also the smallest value s.t. $(t, a) \in R_1$ and $a \in \mathsf{dom}(R_2)$ (resp. the largest value s.t. $(t, z) \in R_1$ and $z \in \mathsf{dom}(R_2)$).

Next, from Lemma 15, for any $x \in [a, z]$, we have

$$a + b_{\delta_2(a)} \leq x + b_{\delta_2(x)} \tag{81}$$

and

$$x + e_{\delta_2(x)} \leq z + e_{\delta_2(z)} \tag{82}$$

Now let a' and z' denote $a + b_{\delta_2(a)}$ and $z + e_{\delta_2(z)}$ respectively.
From (81) and the definition of R_2, a' is the smallest value s.t. $(x, a') \in R_2$ for some $x \in [a, b]$.
And similarly, from (82) and the definition of R_2, z' is the largest value s.t. $(x, z') \in R_2$ for some $x \in [a, b]$.
Together with the defnition of a (resp. of z), this implies that a' (resp. z') is also the smallest (resp. largest) value s.t. $(t, a') \in R_1 \bowtie R_2$ (resp. $(t, z') \in R_1 \bowtie R_2$).
To conclude the proof, we show that:

1. $(t, x) \in R_1 \bowtie R_2$ for each $x \in [a', z']$, and
2. $t + \delta(t) = [a', z']$.

We start with 1.
Consider the tuple $\mathbf{u}' = \langle o_2, o_3, [a, b], \delta_2, b_2, e_2 \rangle \in \mathcal{U}^s$, and let R' be the temporal relation induced by $\mathbf{u}'$.
Then from he definitions of u' and $\mathbf{u}_2$:

$$R' \subseteq R_2 \tag{83}$$

Now take any $x \in [a', z']$.
From Lemma 16 and the definitions of a' and z', there is a $w \in [a, b]$ such that $x \in \delta_2(w)$.
Therefore

$$(w, x) \in R'$$

So from (83)

$$(w, x) \in R_2 \tag{84}$$

Finally, since $[a, b] = t + \delta_1(t)$ and $w \in [a, b]$,

$$(t, w) \in R_1 \tag{85}$$

Together with (84), this implies

$$(t, x) \in R_1 \bowtie R_2$$

which concludes the proof for 1.

For 2, we only prove that $t + b_{\delta_t} = a'$ (the proof that $t + e_{\delta_t} = z'$ is symmetric).
Following the definition of b, we consider 2 cases:
1. $b_1 < b_2 - b_{\delta_1}$
2. $b_1 \geq b_2 - b_{\delta_1}$

In Case 1, we have

$$\begin{aligned} b_1 <& \quad b_2 - b_{\delta_1} && (86)\\ \max(b_1, b_2 - b_{\delta_1}) =& \quad b_2 - b_{\delta_1} && (87)\\ b =& \quad b_2 - b_{\delta_1} \quad \text{from the definition of } b && (88) \end{aligned}$$

And (in Case 1 still):

$$\begin{aligned} b_1 <& \quad b_2 - b_{\delta_1} && (89)\\ 0 <& \quad b_2 - b_{\delta_1} - b_1 && (90)\\ \max(0, b_2 - b_{\delta_1} - b_1) =& \quad b_2 - b_{\delta_1} - b_1 && (91) \end{aligned}$$

Then we consider two subcases:
(i) $t < b_2 - b_{\delta_1}$
(ii) $t \geq b_2 - b_{\delta_1}$

In Case i, we get

$$
\begin{aligned}
t <{} & b_2 - b_{\delta_1} && && (92)\\
0 <{} & b_2 - b_{\delta_1} - t && && (93)\\
\max(0, b_2 - b_{\delta_1} - t) ={} & b_2 - b_{\delta_1} - t && && (94)
\end{aligned}
$$

Now from the definition of δ_t,

$$
\begin{aligned}
b_{\delta_t} ={} & b_{\delta_1} + b_{\delta_2} + \max(0, b - t) && && (95)\\
={} & b_{\delta_1} + b_{\delta_2} + \max(0, b_2 - b_{\delta_1} - t) && \text{from (88)} && (96)\\
={} & b_{\delta_1} + b_{\delta_2} + b_2 - b_{\delta_1} - t && \text{from (94)} && (97)\\
={} & b_{\delta_2} + b_2 - t && && (98)\\
b_{\delta_t} + t ={} & b_{\delta_2} + b_2 - t + t && && (99)\\
={} & b_{\delta_2} + b_2 && && (100)
\end{aligned}
$$

Next, from the definition of a'

$$
\begin{aligned}
a' ={} & b_{\delta_2(a)} + a && && (101)\\
={} & b_{\delta_2} + \max(0, b_2 - a) + a && && (102)
\end{aligned}
$$

And, from the definition of a

$$
\begin{aligned}
a ={} & b_{\delta_1(t)} + t && && (103)\\
={} & b_{\delta_1} + \max(0, b_1 - t) + t && && (104)
\end{aligned}
$$

Then we have two further subcases:

(I) $t \geq b_1$, or

(II) $t < b_1$

In case I:

$$
\begin{aligned}
t \geq{} & b_1 && && (105)\\
0 \geq{} & b_1 - t && && (106)\\
\max(0, b_1 - t) ={} & 0 && && (107)\\
a ={} & b_{\delta_1} + t && \text{from (104)} && (108)\\
\max(0, b_2 - a) ={} & \max(0, b_2 - b_{\delta_1} - t) && && (109)\\
={} & b_2 - b_{\delta_1} - t && \text{from (94)} && (110)\\
={} & b_2 - a && \text{from (108)} && (111)
\end{aligned}
$$

In case II:

$$\begin{aligned}
t <{} & b_1 && (112)\\
0 <{} & b_1 - t && (113)\\
\max(0, b_1 - t) ={} & b_1 - t && (114)\\
a ={} & b_{\delta_1} + b_1 - t + t & \text{from (104)} \quad & (115)\\
={} & b_{\delta_1} + b_1 && (116)\\
\max(0, b_2 - a) ={} & \max(0, b_2 - b_{\delta_1} - b_1) && (117)\\
={} & b_2 - b_{\delta_1} - b_1 & \text{from (91)} \quad & (118)\\
={} & b_2 - a & \text{from (116)} \quad & (119)\\
& && (120)
\end{aligned}$$

So in both cases I and II, we get

$$\max(0, b_2 - a) = b_2 - a$$

Thefore from (102)

$$\begin{aligned}
a' ={} & b_{\delta_2} + b_2 - a + a && (121)\\
={} & b_{\delta_2} + b_2 && (122)\\
={} & t + b_{\delta_t} & \text{from (100)} \quad & (123)
\end{aligned}$$

which concludes the proof for Case 1 i.

We continue with Case 1 ii.
From ii:

$$\begin{aligned}
t \geq{} & b_2 - b_{\delta_1} && (124)\\
0 \geq{} & b_2 - b_{\delta_1} - t && (125)\\
\max(0, b_2 - b_{\delta_1} - t) ={} & 0 && (126)
\end{aligned}$$

Now from the definition of δ_t:

$$\begin{aligned}
b_{\delta_t} ={} & b_{\delta_1} + b_{\delta_2} + \max(0, b - t) && (127)\\
={} & b_{\delta_1} + b_{\delta_2} + \max(0, b_2 - b_1 - t) & \text{from (88)} \quad & (128)\\
={} & b_{\delta_1} + b_{\delta_2} & \text{from (126)} \quad & (129)\\
b_{\delta_t} + t ={} & b_{\delta_1} + b_{\delta_2} + t && (130)
\end{aligned}$$

Next, from 1 and ii, by transitivity, we get

$$\begin{aligned}
b_1 \leq{} & t && (131)\\
\max(0, b_1 - t) ={} & 0 && (132)
\end{aligned}$$

And from the definition of a

$$
\begin{aligned}
a =&\quad b_{\delta_1(t)} + t && (133)\\
=&\quad b_{\delta_1} + \max(0, b_1 - t) + t && (134)\\
=&\quad b_{\delta_1} + t && \text{from (132)} \quad (135)\\
\geq&\quad b_{\delta_1} + b_2 - b_{\delta_1} && \text{from Case ii} \quad (136)\\
\geq&\quad b_2 && (137)\\
0 \geq&\quad b_2 - a && (138)\\
\max(0, b_2 - a) =&\quad 0 && (139)
\end{aligned}
$$

Therefore from (102) and (139)

$$
\begin{aligned}
a' =&\quad b_{\delta_2} + a && (140)\\
=&\quad b_{\delta_2} + b_{\delta_1} + t && \text{from (135)} \quad (141)\\
=&\quad b_{\delta_t} + t && \text{from (100)} \quad (142)
\end{aligned}
$$

which concludes the proof for Case 1 ii.

We continute with Case 2.
In this case, we get

$$
\begin{aligned}
b_1 \geq&\quad b_2 - b_{\delta_1} && (143)\\
\max(b_1, b_2 - b_{\delta_1}) =&\quad b_1 && (144)\\
b =&\quad b_1 && \text{from the definition of } b \quad (145)
\end{aligned}
$$

And from 2 still, we derive

$$
\begin{aligned}
b_1 \geq&\quad b_2 - b_{\delta_1} && (146)\\
0 \geq&\quad b_2 - b_{\delta_1} - b_1 && (147)\\
\max(0, -b_{\delta_1} - b_1) =&\quad 0 && (148)
\end{aligned}
$$

As well as

$$
\begin{aligned}
b_1 \geq&\quad b_2 - b_{\delta_1} && (149)\\
b_1 + b_{\delta_1} \geq&\quad b_2 && (150)
\end{aligned}
$$

Next, we distinguish two subcases, namely
(a) $t < b_1$ and
(b) $t \geq b_1$
We start with Case a.
In this case,

$$
\begin{aligned}
t <&\quad b_1 && (151)\\
0 <&\quad b_1 - t && (152)\\
\max(0, b_1 - t) =&\quad b_1 - t && (153)
\end{aligned}
$$

And from the definition of δ_t:

$$\begin{aligned}
b_{\delta_t} =&\ b_{\delta_1} + b_{\delta_2} + \max(0, b - t) && && (154)\\
=&\ b_{\delta_1} + b_{\delta_2} + \max(0, b_1 - t) && \text{from (145)} && (155)\\
=&\ b_{\delta_1} + b_{\delta_2} + b_1 - t && \text{from (153)} && (156)\\
b_{\delta_t} + t =&\ b_{\delta_1} + b_{\delta_2} + b_1 - t + t && && (157)\\
=&\ b_{\delta_1} + b_{\delta_2} + b_1 && && (158)
\end{aligned}$$

Next, from the definition of a

$$\begin{aligned}
a =&\ b_{\delta_1(t)} + t && && (159)\\
=&\ \max(0, b_1 - t) + b_{\delta_1} + t && && (160)\\
=&\ b_1 - t + b_{\delta_1} + t && \text{from (153)} && (161)\\
=&\ b_1 + b_{\delta_1} && && (162)
\end{aligned}$$

So from (150)

$$\begin{aligned}
a \geq&\ b_2 && && (163)\\
0 \geq&\ b_2 - a && && (164)\\
\max(0, b_2 - a) =&\ 0 && && (165)\\
b_{\delta_2} + \max(0, b_2 - a) =&\ b_{\delta_2} && && (166)\\
b_{\delta_2(a)} =&\ b_{\delta_2} && && (167)\\
b_{\delta_2(a)} + a =&\ b_{\delta_2} + a && && (168)\\
a' =&\ b_{\delta_2} + a && \text{from the definition of } a' && (169)\\
a' =&\ b_{\delta_2} + b_1 + b_{\delta_1} && \text{from (162)} && (170)\\
a' =&\ b_{\delta_t} + t && \text{from (158)} && (171)
\end{aligned}$$

which concludes the proof for Case 2 a.

We end with Case 2 b. In this case,

$$\begin{aligned}
t \geq&\ b_1 && && (172)\\
0 \geq&\ b_1 - t && && (173)\\
\max(0, b_1 - t) =&\ 0 && && (174)
\end{aligned}$$

And from the definition of δ_t:

$$\begin{aligned}
b_{\delta_t} =&\ b_{\delta_1} + b_{\delta_2} + \max(0, b - t) && && (175)\\
=&\ b_{\delta_1} + b_{\delta_2} + \max(0, b_1 - t) && \text{from (145)} && (176)\\
=&\ b_{\delta_1} + b_{\delta_2} && \text{from (174)} && (177)\\
b_{\delta_t} + t =&\ b_{\delta_1} + b_{\delta_2} + t && && (178)
\end{aligned}$$

Next, from the definition of a

$$\begin{aligned}
a =&\ b_{\delta_1(t)} + t && && (179)\\
=&\ \max(0, b_1 - t) + b_{\delta_1} + t && && (180)\\
=&\ b_{\delta_1} + t && \text{from (174)} && (181)
\end{aligned}$$

Now from b

$$
\begin{aligned}
b_1 + \leq{}& \ t && && (182)\\
b_1 + b_{\delta_1} \leq{}& \ t + b_{\delta_1} && && (183)\\
b_1 + b_{\delta_1} \leq{}& \ a && \text{from (181)} && (184)\\
b_2 \leq{}& \ a && \text{from (150), by transitivity} && (185)\\
b_2 - a \leq{}& \ 0 && && (186)\\
\max(0, b_2 - a) ={}& \ 0 && && (187)\\
b_{\delta_2} + \max(0, b_2 - a) ={}& \ b_{\delta_2} && && (188)\\
b_{\delta_2(a)} ={}& \ b_{\delta_2} && && (189)\\
b_{\delta_2(a)} + a ={}& \ b_{\delta_2} + a && && (190)\\
a' ={}& \ b_{\delta_2} + a && \text{from the definition of } a' && (191)\\
={}& \ b_{\delta_2} + b_{\delta_1} + t && \text{from (181)} && (192)\\
={}& \ b_{\delta_t} + t && \text{from (178)} && (193)
\end{aligned}
$$

which concludes the proof for Case 2 b. ◀

The following result states that the representation $(\!|q|\!)^s_G$ is correct:

▶ **Proposition 18.** *Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TPG and q a TRPQ. Then the unfolding of $(\!|q|\!)^s_G$ is $[\![q]\!]_G$.*

Proof. Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TG, and let q be a TRPQ.
To prove the result, it is sufficient to show that:

- **(I)** for any $\langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, there are $\tau, \delta \in \mathsf{intv}(\mathcal{T})$ and $b, e \in \mathcal{T}$ such that
  - **(a)** $\langle o_1, o_2, \tau, \delta, b, e\rangle \in (\!|q|\!)^s_G$,
  - **(b)** $t \in \tau$, and
  - **(c)** $d \in \delta(t)$ (where $\delta(t)$ is defined in terms of t, δ, b and e, as explained above).
- **(II)** for any $\langle o_1, o_2, \tau, \delta, b, e\rangle \in (\!|q|\!)^s_G$ for any $t \in \tau$ and $d \in \delta(t)$, $\langle o_1, o_2, t, d\rangle$ is in $[\![q]\!]_G$.

Again, the proof is by induction on the structure of q.
If q is of the form *pred*, $< k$, F, B, $(\mathsf{test} \vee \mathsf{test})$, $(\mathsf{path} + \mathsf{path})$, $\mathsf{path}[m, n]$ or $\mathsf{path}[m, \_]$, then I and II immediately follow from the definitions of $[\![q]\!]_G$ and $(\!|q|\!)^s_G$.
If q is of the form $\mathsf{test} \wedge \mathsf{test}$, $\neg\mathsf{test}$ or $(?\mathsf{path})$, then the proof is nearly identical to the one already provided for $(\!|q|\!)^t_G$.
And if q is of the form T_δ or $\mathsf{path}_1/\mathsf{path}_2$, then the proof is nearly identical to the one already provided for $(\!|q|\!)^{td}_G$, using Lemma 17 instead of 13. ◀

A.3 Complexity of query answering

We provide in this section complexity results for query answering under the different compact representations studied in the article. The proofs leverage results proven in [2] for non-compact answers. We emphasize that for hardness, the reductions use a graph of fixed size, with the only exception of the temporal domain $\mathcal{T}$.

We start by reproducing the decision problem investigated in [2], which will be intrumental:

Answer	
Input:	TG G over discrete time, TRPQ q, tuple $\mathbf{u} \in \mathcal{U}$
Decide:	$\mathbf{u} \in [\![q]\!]_G$

Next, we define a decision problem for each compact representation, analogous to the problem Compact Answer^t defined in the article.
Let x be one of $[t]$, $[d]$, $[t,d]$ or $([t,d],b,e)$. If $\mathbf{u} \subseteq \mathcal{U}^x$, we use $\mathsf{unfold}(\mathbf{u})$ for the unfolding of $\mathbf{u}$, and we define the partial order $\sqsubseteq_x$ over $\mathcal{U}^x$ as

$$\mathbf{u}_1 \sqsubseteq_x \mathbf{u}_2 \text{ iff } \mathsf{unfold}(\mathbf{u}_1) \subseteq \mathsf{unfold}(\mathbf{u}_2)$$

With this definition, we can decline the notion of compact answer (defined in the article) in four flavors, as follows:

▶ **Definition 19** (Compact answer). *Let G be a TG, let q be a TRPQ and let $\mathbf{u} \in \mathcal{U}^x$. $\mathbf{u}$ is a* compact answer *to q over G (in $\mathcal{U}^x$) if $\mathbf{u} \in \max_{\sqsubseteq_x}\{\mathbf{u}' \in \mathcal{U}^x \mid \mathsf{unfold}(\mathbf{u}') = [\![q]\!]_G\}$*

And we decline the associated decision problem analogously:

Compact Answer^x	
Input:	TG G, TRPQ q, tuple $\mathbf{u} \in \mathcal{U}^x$
Decide:	$\mathbf{u}$ is a compact answer to q over G (in $\mathcal{U}^x$)

Our proofs are structured as follows:

- For membership, we leverage the fact that Answer is in PSpace, which was also proven in [2]: more precisely, we show in Section A.3.1 that Compact Answer^t, Compact Answer^d, and Compact Answer^{td} can each be reduced to a finite number of independent calls to an oracle for Answer.
- For hardness, the results follow immediately from PSpace-hardness for Answer, which was proven in [2]. We show this in Section A.3.2, with a simple reduction from Answer to each of the 5 other problems.

A.3.1 Membership

If $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ is a *TG* and q a TRPQ, we use $\mathsf{boundaries}(G,q)$ for the set of all interval boundaries that appear in G and q, i.e.

$$\mathsf{boundaries}(G,q) = \bigcup\Big\{\{b_\delta, e_\delta\} \mid \mathrm{T}_\delta \text{ appears in } q\Big\} \cup \{b_{\mathcal{T}_G}, e_{\mathcal{T}_G}\} \cup \bigcup\Big\{\{b_\tau, e_\tau\} \mid \tau \in \mathsf{val}(o,p) \text{ for some } o \in N \cup E \text{ and } p \in \mathit{Pred}\Big\}$$

Note that $\mathsf{boundaries}(G,q)$ is finite.

Next, if $Q \subseteq \mathbb{Q}$, we use Q^{+-} to denote the smallest superset of Q that is closed under addition and subtraction.
We can now make the two following observations:

▶ **Lemma 20.** *Let G be a TG, let q be a TRPQ, let $\mathbf{u} = \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, let $Q =$ $\mathsf{boundaries}(G,q) \cup \{d\}$, and let τ be the largest interval s.t. $t \in \tau$ and $\langle o_1, o_2, t', d\rangle \in [\![q]\!]_G$ for all $t' \in \tau$. Then*

$$b_\tau \in Q^{+-} \text{ and } e_\tau \in Q^{+-}$$

▶ **Lemma 21.** *Let G be a TG, let q be a TRPQ, let $\mathbf{u} = \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, let $Q = \mathsf{boundaries}(G, q) \cup \{t\}$ and let δ be the largest interval s.t. $d \in \delta$ and $\langle o_1, o_2, t, d'\rangle \in [\![q]\!]_G$ for all $d' \in \delta$. Then*

$$b_\delta \in Q^{+-} \text{ and } e_\delta \in Q^{+-}$$

Next, let $\sqsubseteq$ denote set inclusion lifted to pairs of intervals, i.e.

$$(\tau_1, \delta_1) \sqsubseteq (\tau_2, \delta_2) \text{ iff } \tau_1 \subseteq \tau_2 \text{ and } \delta_1 \subseteq \delta_2$$

The following is an immediate consequence of Lemmas 20 and 21:

▶ **Corollary 22.** *Let G be a TG, let q be a TRPQ, let $\mathbf{u} = \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G$, let $Q = \mathsf{boundaries}(G, q) \cup \{t, d\}$, let $P = \{(t, d) \mid \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G\}$, and let $(\tau, \delta) \in \max_\sqsubseteq\{(\tau', \delta') \in \mathsf{intv}(\mathcal{T}) \times \mathsf{intv}(\mathcal{T}) \mid t \in \tau \text{ and } d \in \delta\}$. Then*

$$\{b_\tau, e_\tau, b_\delta, e_\delta\} \subseteq Q^{+-}$$

We can now prove our membership results:

▶ **Proposition 23.** *Compact Answer^t is in PSpace.*

Proof.
Let G be a TG, let q be a TRPQ, and let $\mathbf{u} = \langle o_1, o_2, \tau, d\rangle \in \mathcal{U}^t$.
We use Q for $\mathsf{boundaries}(G, q) \cup \{d\}$, and T for the set defined by

$$T = \{t \mid \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G\}$$

We also use k to denote the product of the denominators of all numbers in Q, i.e.

$$k = \Pi\{j \mid \frac{i}{j} \in Q \text{ for some i } \in \mathbb{Z}\}$$

Note that k (encoded in binary) can be computed in time polynomial (therefore using space polynomial) in the cumulated sizes of G, q and $\mathbf{u}$.
We also use $\frac{1}{k}\mathbb{Z}$ (resp. $\frac{1}{2k}\mathbb{Z}$) for the set of all multiples of $\frac{1}{k}$ (resp. $\frac{1}{2k}$), i.e

$$\frac{1}{k}\mathbb{Z} = \{\frac{i}{k} \mid i \in \mathbb{Z}\}$$

and

$$\frac{1}{2k}\mathbb{Z} = \{\frac{i}{2k} \mid i \in \mathbb{Z}\}$$

Note that

$$Q^{+-} \subset \frac{1}{k}\mathbb{Z} \subset \frac{1}{2k}\mathbb{Z}$$

Now let τ' be the largest interval such that $\tau \subseteq \tau' \subseteq T$.
Recall that by assumption, $\tau \neq \emptyset$.
Under this assumption, $\langle G, q, u\rangle$ is an instance of Compact Answer^t iff $\tau = \tau'$.
We show that $\tau = \tau'$ can be decided using space polynomial in the cumulated size of (the encodings of) G, q and u.

First, fom Lemma 20, we observe that $b_\tau \notin \frac{1}{k}\mathbb{Z}$ or $e_\tau \notin \frac{1}{k}\mathbb{Z}$ implies $\tau \neq \tau'$.
And $b_\tau \in \frac{1}{k}\mathbb{Z}$ (resp. $e_\tau \in \frac{1}{k}\mathbb{Z}$) can be decided in time polynomial in the encoding of b_τ (resp. e_τ).
So we can focus on the case where $b_\tau \in \frac{1}{k}\mathbb{Z}$ and $e_\tau \in \frac{1}{k}\mathbb{Z}$.
We use b_{inf} for the largest element of $(\frac{1}{2k}\mathbb{Z}) \setminus \tau$ that satisfies $b_{\mathsf{inf}} \leq b_\tau$.
And similarly we use e_{sup} for the smallest element of $(\frac{1}{2k}\mathbb{Z}) \setminus \tau$ that satisfies $e_\tau \leq e_{\mathsf{inf}}$.
Observe that b_{inf} and e_{sup} can be computed using space polynomial in (the encoding of) τ.

We show below that for any (nonempty) interval α with boundaries in $\frac{1}{k}\mathbb{Z}$,

$$\alpha \subseteq T \text{ iff } \alpha \cap \frac{1}{2k}\mathbb{Z} \subseteq T \tag{194}$$

Therefore in order to decide whether $\tau = \tau'$, it is sufficient to decide whether
(I) $\tau \cap \frac{1}{2k}\mathbb{Z} \subseteq T$, and
(II) $\{b_{\mathsf{inf}}, e_{\mathsf{sup}}\} \cap T = \emptyset$
Now observe that:
- I can be decided with a finite number of independent calls to a procedure for Answer, and
- II can be decided with two calls to such a procedure.

And it was shown in [2] that Answer is in PSpace.

To complete the proof, we show that (194) holds.
The right direction ($\alpha \subseteq T$ implies $\alpha \cap \frac{1}{2k}\mathbb{Z} \subseteq T$) is trivial.
For the left direction, assume by contradiction that $\alpha \cap \frac{1}{2k}\mathbb{Z} \subseteq T$ but $\alpha \not\subseteq T$.
Take any $t \in \alpha \setminus T$.
Since $\alpha \cap \frac{1}{2k}\mathbb{Z} \subseteq T$ and $t \notin T$, we have

$$t \notin \frac{1}{2k}\mathbb{Z} \tag{195}$$

Next, since α has boundaries in $\frac{1}{k}\mathbb{Z}$,

$$\alpha \cap \frac{1}{2k}\mathbb{Z} \neq \emptyset \tag{196}$$

(for instance, $b_\alpha + \frac{1}{2k} \in \alpha \cap \frac{1}{2k}\mathbb{Z}$).
Together with (195), this implies that there is a t' in $\alpha \cap \frac{1}{2k}\mathbb{Z}$ s.t. either $t' < t$ or $t < t'$.
Let us assume w.l.o.g. that the former holds (the proof for the latter case is symmetric).
And let t_{inf} be the largest value that satisfies $t_{\mathsf{inf}} \in \alpha \cap \frac{1}{2k}\mathbb{Z}$ and $t_{\mathsf{inf}} < t$.
Then

$$t - t_{\mathsf{inf}} < \frac{1}{2k} \tag{197}$$

Now recall that by assumption, $\alpha \cap \frac{1}{2k}\mathbb{Z} \subseteq T$.
Therefore $t_{\mathsf{inf}} \in T$.
So from Lemma 20, there is a β with boundaries in $\frac{1}{k}\mathbb{Z}$ s.t. $\beta \subseteq T$ and $t_{\mathsf{inf}} \in \beta$.
Then we have two cases, either $e_\beta \neq t_{\mathsf{inf}}$ or $e_\beta = t_{\mathsf{inf}}$:
- ($e_\beta \neq t_{\mathsf{inf}}$).
  In this case, since $e_\beta \in \frac{1}{k}\mathbb{Z}$, and $t_{\mathsf{inf}} \in \frac{1}{2k}\mathbb{Z}$,

  $$\frac{1}{2k} \leq e_\beta - t_{\mathsf{inf}}$$

Together with (197), this yields (by transitivity)

$$t - t_{\mathsf{inf}} < e_\beta - t_{\mathsf{inf}} \tag{198}$$
$$t < e_\beta \tag{199}$$

Now since $t_{\mathsf{inf}} \in \beta$,

$$b_\beta \leq t_{\mathsf{inf}} \tag{200}$$

Together with $t_{\mathsf{inf}} < t$, this implies

$$b_\beta < t \tag{201}$$

Together with (199), this yields

$$t \in \beta$$

Since $\beta \subseteq T$, this implies $t \in T$, which contradicts the definition of t.

- $(e_\beta = t_{\mathsf{inf}})$.
In this case, since β has boundaries in $\frac{1}{k}\mathbb{Z}$,

$$t_{\mathsf{inf}} \in \frac{1}{k}\mathbb{Z} \tag{202}$$

And because $t \in \alpha$ and $t_{\mathsf{inf}} < t$

$$t_{\mathsf{inf}} < t \leq e_\alpha \tag{203}$$
$$t_{\mathsf{inf}} < e_\alpha \tag{204}$$

Together with (202) and $e_\alpha \in \frac{1}{k}\mathbb{Z}$, this implies

$$\frac{1}{k} \leq e_\alpha - t_{\mathsf{inf}} \tag{205}$$

Now let $t_{\mathsf{sup}} = t_{\mathsf{inf}} + \frac{1}{2k}$.
From (205), we get

$$t_{\mathsf{sup}} < e_\alpha \tag{206}$$

Next, since $t_{\mathsf{inf}} \in \alpha$ and $t_{\mathsf{inf}} < t_{\mathsf{sup}}$

$$b_\alpha < t_{\mathsf{sup}} \tag{207}$$

So from (206) and (207)

$$t_{\mathsf{sup}} \in \alpha$$

So from Lemma 20, there is a β' with boundaries in $\frac{1}{k}\mathbb{Z}$ s.t. $\beta' \subseteq T$ and $t_{\mathsf{sup}} \in \beta'$.
Next, from (197) and the definition of t_{sup}

$$t_{\mathsf{sup}} - t < \frac{1}{2k} \tag{208}$$

So with an argument symmetric to the one used above to show $t \in \beta$, we get $t \in \beta'$, which once again contradicts $t \notin T$.

◀

▶ **Proposition 24.** *Compact Answer^d is in PSpace*

Proof.
The proof is symmetric to the one provided above for Proposition 23, using Lemma 21 instead of Lemma 20. ◀

▶ **Proposition 25.** *Compact Answer^{td} is in PSpace.*

Proof.
The proof is analogous to the one provided above for Proposition 23, using Corollary 22 instead of Lemma 20.
More precisely, let G be a TG, let q be a TRPQ, and let $\mathbf{u} = \langle o_1, o_2, \tau, \delta\rangle \in \mathcal{U}^{td}$.
We use Q for $\mathsf{boundaries}(G, q) \cup \{t, d\}$, and P for the set defined by

$$P = \{(t, d) \mid \langle o_1, o_2, t, d\rangle \in [\![q]\!]_G\}$$

We also define k, $\frac{1}{k}\mathbb{Z}$ and $\frac{1}{2k}\mathbb{Z}$ identically as in the proof of Proposition 23.
Then analogously to what we showed in this proof, for any pair of intervals (α_1, α_2) with boundaries in $\frac{1}{k}\mathbb{Z}$,

$$\alpha_1 \times \alpha_2 \subseteq P \text{ iff } (\alpha_1 \cap \frac{1}{2k}\mathbb{Z}) \times (\alpha_2 \cap \frac{1}{2k}\mathbb{Z}) \subseteq P$$

So with a similar argument, deciding whether $\langle o_1, o_2, \tau, \delta\rangle$ is a compact answer to q over G can be reduced to deciding

- $\{b_\tau, e_\tau, b_\delta, e_\delta\} \subseteq \frac{1}{k}\mathbb{Z}$,
- $(\tau \cap \frac{1}{2k}\mathbb{Z}) \times (\delta \cap \frac{1}{2k}\mathbb{Z}) \subseteq P$,
- $\big(\{b^\tau_{\mathsf{inf}}, e^\tau_{\mathsf{sup}}\} \times (\delta \cap \frac{1}{2k}\mathbb{Z})\big) \cap P = \emptyset$ and
- $\big((\tau \cap \frac{1}{2k}\mathbb{Z}) \times \{b^\delta_{\mathsf{inf}}, e^\delta_{\mathsf{sup}}\}\big) \cap P = \emptyset$

where b^τ_{inf} is the largest element in $(\frac{1}{2k}\mathbb{Z}) \setminus \tau$ that satisfies $b^\tau_{\mathsf{inf}} \leq b_\tau$, e^τ_{sup} is the smallest element in $(\frac{1}{2k}\mathbb{Z}) \setminus \tau$ that satisfies $e_\tau \leq e^\tau_{\mathsf{sup}}$, and b^δ_{inf} and e^δ_{sup} are defined analogously. ◀

A.3.2 Hardness

▶ **Proposition 26.** *Compact Answer^t is PSpace-hard*

Proof. The proof is a straightforward reduction from Answer.
Let $G = \langle N, E, \mathsf{conn}, \mathcal{T}_G, \mathsf{val}\rangle$ be a TG, let q be a TRPQ and let $\mathbf{u} = \{o_1, o_2, t, d\} \in \mathcal{U}$.
W.l.o.g., let us assume that $\{o_1, o_2\} \subseteq N$ (the proof for the 3 other cases is symmetric).

Now let $P \subseteq \mathit{Pred}$ be the set of predicates defined by $p \in \mathit{Pred}$ iff there is an o s.t. $\mathsf{val}(o, p) \neq \emptyset$.
Take two fresh predicates $p_1, p_2 \in \mathit{Pred} \setminus P$, two fresh nodes $n_1, n_2 \notin N$ and two fresh edges $e_1, e_2 \notin E$.
And let $G' = \langle N \cup \{n_1, n_2\}, E \cup \{e_1, e_2\}, \mathsf{conn}', \mathcal{T}_G, \mathsf{val}'\rangle$ be the TG identical to G, except for the functions conn' and val', defined by

- $\mathsf{conn}'(e) = \mathsf{conn}(e)$ for all $e \in E$,
- $\mathsf{conn}'(e_1) = (n_1, o_1)$,
- $\mathsf{conn}'(e_2) = (o_2, n_2)$,
- $\mathsf{val}'(o, p) = \mathsf{val}(o, p)$ for all $(o, p) \in (N \cup E) \times (\mathit{Pred} \setminus \{p_1, p_2\})$,
- $\mathsf{val}'(n_1, p_1) = \{[t, t]\}$, and

- $\mathsf{val}'(n_2, p_2) = \{[t+d, t+d]\}$

Let q' be the TRPQ defined by

$$q' = p_1/\mathrm{F}/q/\mathrm{F}/p_2$$

Then immediately from the semantics of TRPQs:

$$\mathbf{u} \in [\![q]\!]_G \text{ iff } [\![q']\!]_{G'} = \{\langle n_1, n_2, t, d\rangle\} \tag{209}$$

Now consider the tuple $\mathbf{u}' = \{n_1, n_2, [t,t], d\} \in \mathcal{U}^t$.
Then from (209), $\mathbf{u} \in [\![q]\!]_G$ iff $\mathbf{u}'$ is a compact answer to q over G in $\mathcal{U}^t$.

Clearly, the input $\langle G', q, \mathbf{u}'\rangle$ to Compact Answer^t can be computed in time polynomial in the size of (the encodings of) G, q and $\mathbf{u}$.
And it was shown in [2] that Answer is PSpace-complete.

◀

▶ **Proposition 27.** *Compact Answer^d, Compact Answer^{td} and Compact Answer^s are PSpace-hard.*

Proof. The proofs are nearly identical to the one provided above for Compact Answer. The graph G' is defined identically in all cases, so that the reductions only differ w.r.t. to the tuple $\mathbf{u}'$.
This tuple is defined as follows:

- $\{n_1, n_2, t, [d,d]\}$ for Compact Answer^d,
- $\{n_1, n_2, [t,t], [d,d]\}$ for Compact Answer^{td},
- $\{n_1, n_2, [t,t], [d,d], t, t\}$ for Compact Answer^s.

◀

A.4 Size of compact answers

▶ **Proposition 3.** *Let q be a star-free TRPQ and G a TG such that $[\![q]\!]_G$ can be finitely represented in $\mathcal{U}^t$. Then $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^t) = O(1)$ and $\mathsf{size}^\delta([\![q]\!]_G, \mathcal{U}^t) = \Omega(n)$.*

Proof. $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^t) = O(1)$ follows from the definition of $(\!|q|\!)^t_G$ (in Section A.2.1.1) and Proposition 10, by induction on the structure of q:

- In the base case, i.e. when q is of the form $\mathit{pred}, < k, \mathrm{T}_\delta, \mathrm{F}, \mathrm{B}$ or $?\mathsf{path}$, the number of tuples in $(\!|q|\!)^t_G$ is independent of the size of the intervals present in G.
- For all binary operators, (i.e. when q is of the form $\mathsf{test}_1 \wedge \mathsf{test}_2, \mathsf{test}_1 \vee \mathsf{test}_2, \mathsf{path}_1/\mathsf{path}_2$ or $\mathsf{path}_1 + \mathsf{path}_2$), the cardinality of $(\!|q|\!)^t_G$ is bounded by some function of the size of the operands.
  Precisely, if k_1 (resp. k_2) is the cardinality of $(\!|q_1|\!)^t_G$ (resp. $(\!|q_2|\!)^t_G$), then the cardinality of $(\!|q_1/q_2|\!)^t_G$ (resp. $(\!|q_1+q_2|\!)^t_G, q_1 \wedge q_2, q_1 \vee q_2$) is bounded by $k_1 \cdot k_2$ (resp. $k_1 + k_2$, $k_1 \cdot k_2$, $k_1 + k_2$).
  And by IH, the cardinality of $(\!|q_i|\!)^t_G$ is independent of the size of intervals in G.
  This argument also applies to the case where q is of the form $\mathsf{path}[m,n]$, since it is equivalent in this case to a finite union of joins.
- If q is of the form $\neg q'$, then the cardinality of $(\!|q|\!)^t_G$ is bounded by $2k$, where k is the cardinality of $(\!|q'|\!)^t_G$, since the complement in $\mathcal{T}_G$ of an interval can always be represented with at most two intervals.
  And once again, by IH, k is independent of the size of intervals in G.

To show that $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^t) = \Omega(n)$, we use a graph G over discrete time, with a single node o and no edge.
For some $n \in \mathbb{N}$, consider the TRPQ $q = \mathrm{T}_{[0,n]}$, and assume that the (discrete) domain $\mathcal{T}_G$ of G is $[0, n]$.
Then

$$(\!|q|\!)^t_G = \{\langle o, o, [0, 0], d\rangle \mid d \in [0, n]\}$$

Each tuple in this set has a different value for the distance d, therefore $(\!|q|\!)^t_G$ is the compact representation of $\llbracket q \rrbracket_G$ in $\mathcal{U}^t$.
And this set has cardinality n. ◀

▶ **Proposition 5.** *Let q be a grounded star-free TRPQ and G a TG such that $\llbracket q \rrbracket_G$ can be finitely represented in $\mathcal{U}^d$. Then $\mathsf{size}^\tau(\llbracket q \rrbracket_G, \mathcal{U}^d) = \Omega(n)$ and $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^d) = O(1)$.*

Proof. $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^d) = O(1)$ follows from the definition of $(\!|q|\!)^d_G$ (in Section A.2.2.1) and Proposition 11, with and argument is nearly identical to the one that we provided to show that $\mathsf{size}^\tau(\llbracket q \rrbracket_G, \mathcal{U}^t) = O(1)$, in the proof of Proposition 3.
To show that $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^t) = \Omega(n)$, we use once again a graph G over discrete time, with a single node o and no edge.
Let $p \in \mathit{Pred}$, and for $n \in \mathbb{N}$, assume that p holds only at node o in the interval $[0, n]$, i.e. $\mathsf{val}(o, p) = \{[0, n]\}$, and assume that the active temporal domain $\mathcal{T}_G$ is also $[0, n]$.
Now consider the query $q = p$. Then

$$(\!|q|\!)^d_G = \{\langle o, o, t, [0, 0]\rangle \mid t \in [0, n]\}$$

Each tuple in this set has a different value for the distance d, therefore $(\!|q|\!)^d_G$ is the compact representation of $\llbracket q \rrbracket_G$ in $\mathcal{U}^d$.
And this set has cardinality n. ◀

▶ **Proposition 7.** *Let q be a star-free TRPQ and G a TG such that $\llbracket q \rrbracket_G$ can be finitely represented in $\mathcal{U}^{td}$. Then $\mathsf{size}^\tau(\llbracket q \rrbracket_G, \mathcal{U}^{td}) = O(1)$ and $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^{td}) = \Omega(n)$.*

Proof. To show why $\mathsf{size}^\delta(\llbracket q \rrbracket_G, \mathcal{U}^d) = \Omega(n)$, we use an example analogous to the the one used to show non-finiteness (illustrated with Figure 7c).
Consider a TG G over discrete time with a single node o and no edge.
For some $n \in \mathbb{N}$, consider the TRPQ $q = \mathrm{T}_{[0,n]}$, and assume that the domain $\mathcal{T}_G$ of G is $[0, n]$.
Then $\llbracket q \rrbracket_G = \{\langle o, o, t, d\rangle \mid t \in [0, n] \text{ and } d \in [0, n - t]\}$.
For instance, if $x = 2$, then

$$\llbracket q \rrbracket_G = \{\langle o, o, 0, 0\rangle, \langle o, o, 0, 1\rangle, \langle o, o, 0, 2\rangle, \langle o, o, 1, 1\rangle, \langle o, o, 1, 2\rangle, \langle o, o, 2, 2\rangle\}$$

Now consider again the Euclidean plane $\mathbb{Z} \times \mathbb{Z}$, and the polygon defined by the points $(0, 0), (0, n), (n, 0)$ and $\{(t, n - t + 1) \mid t \in [1, n]\}$.
Intuitively, this is a "near triangle" (visually, analogous to an upper approximation of the integral of a linear function).
The minimal number of tuples in $\mathcal{U}^{td}$ needed to represent $\llbracket q \rrbracket_G$ is the minimal number of rectangles needed to cover this polygon, and this number is n.

To show that $\mathsf{size}^\tau(\llbracket q \rrbracket_G, \mathcal{U}^{td}) = O(1)$ we use an argument similar to the one that we provided to show that $\mathsf{size}^\tau(\llbracket q \rrbracket_G, \mathcal{U}^t) = O(1)$, in the proof of Proposition 3.

We proceed once again again by induction on the structure of q.
If q is an expression for the symbol test in the grammar of Section 3, then the argument is identical to the one provided for $\mathcal{U}^t$.
This is is also the case if q is of the form $F, B, \mathsf{path} + \mathsf{path}$ or $\mathsf{path}[m, n]$.
So we focus here on the two remaining operators, namely the cases where q is of the form T_δ or $\mathsf{path}_1/\mathsf{path}_2$.
Recall that

$$(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^{td}_G \bigcup \{\mathbf{u}_1 \overline{\bowtie}\, \mathbf{u}_2 \mid \mathbf{u}_1 \in (\!|\mathsf{path}_1|\!)^{td}_G, \mathbf{u}_2 \in (\!|\mathsf{path}_2|\!)^{td}_G\}$$

and

$$(\!|\mathrm{T}_\delta|\!)^{td}_G = \bigcup_{o \in N \cup E} \{\langle o, o, \mathcal{T}_G, \delta\rangle \overline{\bowtie}\, \langle o, o, \mathcal{T}_G, [0, 0]\rangle\}$$

Let $\mathbf{u}_1 = \langle o_1, o_2, \tau_1, \delta_1\rangle$ and $\mathbf{u}_2 = \langle o_1, o_2, \tau_2, \delta_2\rangle$ be two tuples in $\mathcal{U}^{td}$. For simplicity, we assume that all intervals here are closed-closed, and we focus on the discrete case (but the argument can be easily adapted to the other cases).
We show that the cardinality of $\mathbf{u}_1 \overline{\bowtie}\, \mathbf{u}_2$ is bounded by a function of the cardinality of δ_1.
Together with the definitions of $(\!|\mathsf{path}_1/\mathsf{path}_2|\!)^{td}_G$ and $(\!|\mathrm{T}_\delta|\!)^{td}_G$, this proves our claim.
We use R_1, R_2, $R_1 \bowtie R_2$, b and e here with the same meaning as in Section 4.3.
We already saw in Section 4.3 that the cardinality of $\mathbf{u}_1 \overline{\bowtie}\, \mathbf{u}_2$ is the cardinality of the set $\mathsf{dom}(R_1 \bowtie R_2) \setminus [b, e] + 1$ if $b \leq e$, and the cardinality of $\mathsf{dom}(R_1 \bowtie R_2)$ otherwise.
In the former case, as illustrated with Figure 6, the set $\mathsf{dom}(R_1 \bowtie R_2) \setminus [b, e]$ is the union of the two intervals $[b_{\mathsf{dom}(R_1 \bowtie R_2)}, b]$ and $[e, e_{\mathsf{dom}(R_1 \bowtie R_2)}]$, and the cardinality of each of these is bounded by $e_{\delta_1} - b_{\delta_1}$, which is indeed the cardinality of δ_1.
For the case where $e < b$, $\mathsf{dom}(R_1 \bowtie R_2)$ is also the union of these two intervals (which now overlap), and it can be easily seen from Figure 6 that $e_{\delta_1} - b_{\delta_1}$ is still an upper bound on the cardinality of each of them.

◀

▶ **Proposition 8.** *For a star-free TRPQ q and TG G,* $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^s) = \mathsf{size}^\delta([\![q]\!]_G, \mathcal{U}^s) = O(1)$

Proof. Both results follow from the definition of $(\!|q|\!)^s_G$ (in Section A.2.4.1) and Proposition 18, with arguments nearly identical to the one that we provided to show that $\mathsf{size}^\tau([\![q]\!]_G, \mathcal{U}^t) = O(1)$, in the proof of Proposition 3. ◀

A.5 Compactness

Regarding the cost of coalescing, all our results are already justified in the body of the article.
Regarding (non-)unicity of compact representations, all arguments are also provided, with the exception of non-unicity for the fourth representations (i.e. in $\mathcal{U}^s$).

We show that a set V of tuples in $\mathcal{U}$ that share the same objects o_1 and o_2 may have several compact representations in $\mathcal{U}^s$.
Over dense time, consider once again the Euclidean plane $\mathcal{T}_G \times \mathbb{Q}$.
Observe that:

(i) any rectangle $\tau \times \delta$ in this plane is exactly covered by some tuple in $\mathcal{U}^s$ (namely $\langle o_1, o_2, \tau, \delta, b_\tau, e_\tau\rangle$), and

(ii) the area covered by a tuple in $\mathcal{U}^s$ forms either a rectangle, or a polygon with some non-square angles.

Now assume that the area covered by V forms an L-shaped polygon.
From ii, this area cannot be exactly covered by a single tuple in $\mathcal{U}^s$, and from i there are more than one pair of tuples that exactly cover it (as illustrated by Figure 7b).
The same argument can easily be adapted to discrete time.