THE EXPRESSIVE POWER OF THE TEMPORAL QUERY LANGUAGE $L_{\mathcal{H}}$

Matei Popovici[1,2], Lorina Negreanu[3]

The paper investigates the expressive power of the temporal query language $L_{\mathcal{H}}$. We show that First-Order Logic is unable to formulate queries such as temporal connectivity, which can be naturally expressed by $L_{\mathcal{H}}$. The paper describes in detail our application of the Ehrenfeucht-Fraïssé method, which is used to examine limitations in the expressive power of First-Order Logic.

Keywords: mathematical logic, finite model theory, FOL, Ehrenfeucht-Fraïssé games
MSC2010: 03C13.

1. Introduction

Temporal knowledge representation and reasoning methods require *expressive power* — which ensures usefulness in non-trivial applications as well as *reasonable complexity*, which guarantees efficient implementation. The temporal language $L_{\mathcal{H}}$ [7, 4, 8, 5] is an attempt to find a balance between expressive power and complexity. $L_{\mathcal{H}}$ serves as a means for expressing temporal constraints over the properties of a domain, such as:'*x is a device which has been operating during the same time device y was operating*".

In this paper, we show that $L_{\mathcal{H}}$ is expressive enough to allow defining *temporal connectivity queries*, which in turn cannot be defined in First-Order Logic (FOL). The methodology which we use relies on Ehrenfeucht-Fraïssé-games. This formal result is of practical interest, because it shows that $L_{\mathcal{H}}$ cannot be embedded in a relational database schema relying on FOL.

The rest of the paper is structured as follows: in Section 2 we introduce temporal graphs and in Section 3 — the language $L_{\mathcal{H}}$. In Section 4 we examine the expressive power of $L_{\mathcal{H}}$. To this end, we show that $L_{\mathcal{H}}$ is expressive enough to capture connectivity constraints. In contrast, in Section 5 we show FOL is unable to express such constraints. Finally, in Section 6, we conclude.

---

[1]Lecturer, Department of Computer Science, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, E-mail: matei.popovici@cs.pub.ro

[2]Postdoctoral researcher, Research Institute of University of Bucharest (ICUB), Romania

[3]Senior Lecturer, Department of Computer Science, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, E-mail: lorina.negreanu@cs.pub.ro

2. Temporal graphs

2.1. Introduction

The language $L_{\mathcal{H}}$ and it's models — temporal graphs, have been extensively described in previous work [4, 8, 5, 7]. In what follows, we present them in brief. Temporal graphs are structures encoding the evolution of a domain. They encode *temporal moments* (formally denoted hypernodes), *system actions* (action nodes) and *time-dependent properties* (or qualify edges).

2.2. Formal definition

A *temporal graph* (short **t**-graph) is a triple $\mathcal{H} = (V, E, H)$ where (V, E) is a directed graph: (i) $a \in V$ are action nodes and (ii) $(a, b) \in E$ are quality edges; H is a partition over the set V. The elements $h \in H$ are called hypernodes. We say two action nodes a, b are *simultaneous* iff $a, b \in h$. We occasionally write $h \in \mathcal{H}, a \in \mathcal{H}, (a, b) \in \mathcal{H}$ instead of $h \in H, a \in V, (a, b) \in E$.

A *trace* in a **t**-graph is a finite sequence $a_1, \ldots, a_n$ of action nodes, such that, for each two consecutive action nodes $a_i\, a_{i+1}$, exactly one of the following cases holds: (i) $a_i \in h$ and $a_{i+1} \in h$, for some hypernode h (i.e. a_i and a_{i+1} occur at the same time); (ii) $(a_i, a_{i+1}) \in E$ (there exists a quality edge from a_i to a_{i+1}). A trace is said to be *compact* iff, for no sequence $a_i a_{i+1} a_{i+2}$, we have $a_i, a_{i+1}, a_{i+2} \in h$. The length of a trace t, denoted $|t|$, is the number of action nodes containing it. If t and t' are two traces, we denote their concatenation by tt'. We write $[a, b]$ to refer to the trace starting in a and ending b. In Figure 1a, $a_1 a_3 a_6$ and $a_1 a_2 a_3 a_6$ are compact traces in a temporal graph.

A *labelled temporal graph* is a temporal graph together with a labelling $\mathcal{L}$, which assigns for each action node a, and each quality edge (a, b) the label $\mathcal{L}(a)$ and $\mathcal{L}(a, b)$, respectively. A label (for both action nodes and quality edges) is a relation instance of the form $Q(i_1, \ldots, i_n)$, where Q is a relation of arity n and $i_1, \ldots, i_n$ are *individuals*. Figure 1a illustrates a labelled temporal graph, encoding the evolution of a time-dependent domain consisting of devices a and b, a sensor-equipped window *win*, and the outside environment, encoded by e.

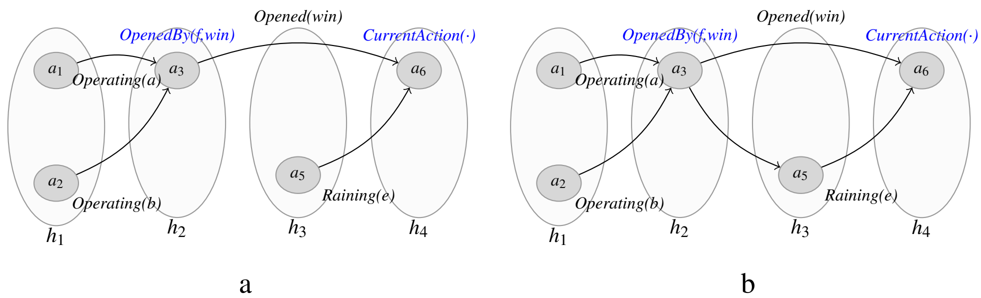

FIGURE 1. The labelled temporal graph $\mathcal{H}_2$

3. The language $L_{\mathcal{H}}$

Definition 3.1 (Precedence)**.** *Let $\mathcal{H}$ be a temporal graph, h, h' be hypernodes from $\mathcal{H}$ and $a \in h, a' \in h'$ be action nodes:*

- h immediately precedes h' *iff there exists a quality edge (a, b) such that $a \in h$ and $b \in h'$. Informally, the condition expresses that there is a property that starts in h and ends in h'.*
- h precedes h' *(denoted $h \succ h'$) iff (i) h immediately precedes h' or (ii) there exists h'' such that h immediately precedes h'' and h'' precedes h'.*
- *a (immediately) precedes a' iff h (immediately) precedes h'.*

Definition 3.2 ($L_{\mathcal{H}}$ syntax)**.** *Let $\mathbb{V}ars$ designate a set of variables, $x \in \mathbb{V}ars$ and R — a relation of arity n. Terms — denoted $t_1, \ldots, t_n$, are either variables or constants. The syntax of $L_{\mathcal{H}}$ is recursively defined as follows:*

$$\varphi ::= R(t_1, \ldots, t_n) \mid R(t_1, \ldots, t_n) \propto \varphi' \mid \neg\varphi' \mid \varphi' \wedge \varphi''$$

where $\propto$ designates any temporal precedence relation from Allen's Interval Algebra [1]*. A formula of the form $R(t_1, \ldots, t_n)$ is called* atomic*. In this paper, we assume, without loss of generality, that $\propto \in \{$ **b** , **a** , **m** $\}$, where* **b** *stands for* before*,* **a** *stands for* after *and* **m** *stands for* meets*.The operator precedence order is: $\neg, \propto, \wedge$.*

Definition 3.3 ($L_{\mathcal{H}}$ semantics)**.** *Let $\varphi \in L_{\mathcal{H}}$ be an atomic formula, $\psi, \psi' \in L_{\mathcal{H}}$ be formulae and $\mathcal{H}$ be a labelled temporal graph. We denote by $||\psi||_{\mathcal{H}}$ the set of quality edges which satisfy ψ in $\mathcal{H}$.*

$$\begin{aligned}
||R(t_1, \ldots, t_n)||_{\mathcal{H}} &= \{(a, b) \in \mathcal{H} \mid \mathcal{L}(a, b) = R(t_1, \ldots, t_n)\} \\
||\varphi\ \boldsymbol{b}\ \psi||_{\mathcal{H}} &= \{(a, b) \in ||\varphi||_{\mathcal{H}} \mid \exists(c, d) \in ||\psi||_{\mathcal{H}} \text{ such that } b \text{ precedes } c\} \\
||\varphi\ \boldsymbol{a}\ \psi||_{\mathcal{H}} &= \{(a, b) \in ||\varphi||_{\mathcal{H}} \mid \exists(c, d) \in ||\psi||_{\mathcal{H}} \text{ such that } d \text{ precedes } a\} \\
||\neg\psi||_{\mathcal{H}} &= \{(a, b) \in \mathcal{H} \mid (a, b) \notin ||\psi||_{\mathcal{H}}\} \\
||\psi \wedge \psi'||_{\mathcal{H}} &= ||\psi||_{\mathcal{H}} \cap ||\psi'||_{\mathcal{H}}
\end{aligned}$$

4. The expressive power of $L_{\mathcal{H}}$

In this section, we introduce the standard concept of query definability [6] and show that the *connectivity query* is definable in $L_{\mathcal{H}}$.

A *vocabulary* is a set σ of symbols R_i, each having assigned a natural number n_i, called *arity*. A *(σ-) structure*, denoted $\mathfrak{R}$, contains a set U, and an assignment which maps: (i) each symbol c in σ of arity 0 to an element of U, (ii) each symbol R of arity $n > 0$, to a relation $R^{\mathfrak{R}}$ of arity n. We also refer to $\mathfrak{R}$ as *relational database*. Given two relational databases $\mathfrak{R}_1$ and $\mathfrak{R}_2$ over the same vocabulary, an isomorphism between $\mathfrak{R}_1$ and $\mathfrak{R}_2$ is a function $f : U_1 \to U_2$ such that: (i) f is bijective and (*ii*) if $(u_1, \ldots, u_n) \in R_i^{\mathfrak{R}_1}$, then $(f(u_1), \ldots, f(u_n)) \in R_i^{\mathfrak{R}_2}$, for all relation symbols R_i. A class $\mathcal{C}$ is a set of databases which is *closed under isomorphisms*, i.e. for any $\mathfrak{R} \in \mathcal{C}$, if $\mathfrak{R}'$ is isomorphic to $\mathfrak{R}$, then $\mathfrak{R}' \in \mathcal{C}$. A *boolean query* is a function $F : \mathcal{C} \to \{0, 1\}$ that is *preserved under isomorphisms*, i.e. if $\mathfrak{R}$ and $\mathfrak{R}'$ are isomorphic, then $F(\mathfrak{R}) = F(\mathfrak{R}')$.

If L is a logical language, then the boolean query F on $\mathcal{C}$ is L-definable iff there is a sentence φ in L such that, for every $\mathfrak{R} \in \mathcal{C}$, we have:

$$F(\mathfrak{R}) = 1 \iff \mathfrak{R} \models_L \varphi$$

Let *CON* be the boolean query which verifies the existence of **t**-graph connectivity (mentioned above), more precisely, $CON(\mathcal{H}) = 1$ iff, from each action node a, there exists a trace to any other action node b from $\mathcal{H}$.

Proposition 4.1 ($L_{\mathcal{H}}$ definability)**.** CON *is $L_{\mathcal{H}}$-definable.*

Proof. We prove the complement of *CON*, denoted $\neg CON$, to be $L_{\mathcal{H}}$-definable. First, we introduce the entailment relation $\models_{L_{\mathcal{H}}}$ with respect to $\|\cdot\|_{.}$, as follows: $\mathcal{H} \models_{L_{\mathcal{H}}} \varphi$ iff $\|\varphi\|_{\mathcal{H}} \neq \emptyset$. Without the loss of generality, we assume each quality edge (a, b) in $\mathcal{H}$ is also labelled *Q(gen)*, where *gen* is some arbitrary individual. This enables us to formulate a nicer formula, which does not require a disjunction over all distinct quality labels. Then, it is clear that $\neg CON(\mathcal{H}) = 1$ iff $\mathcal{H} \models_{L_{\mathcal{H}}} \neg$ *(A(gen) before A(gen) ∧ A(gen) after A(gen))*, by the following argument. The (sub-)formula ψ *=(A(gen) before A(gen) ∧ A(gen) after A(gen))* is satisfied by a quality edge which is both before and after another some quality edge. Thus, $\|\neg\phi\|$ is non-empty iff there exists some quality edge which is neither before nor after another. Thus, $\neg CON$ is definable in $L_{\mathcal{H}}$ by $\neg\psi$. □

5. FOL undefinability of *CON*

In this section, we show that *CON* cannot be defined in First-Order Logic (FOL).

First, we introduce a relational representation of a temporal graph. For convenience, use the term *labelling domain* to refer to σ-structures.

Let $\sigma = (\sigma_A, \sigma_Q)$ designate the vocabulary of the represented domain, where σ_Q is the vocabulary for qualities and σ_A, that for actions. Let Q be an arbitrary symbol from σ_Q, of arity n. We denote by $aug(Q)$, the *augmentation* of Q. $aug(Q)$ is the same symbol Q but with arity $n + 2$. Similarly, the augmentation $aug(A)$ of the symbol $A \in \sigma_A$ of arity n is the symbol A with arity $n + 1$. We denote by $aug(\sigma)$ the vocabulary obtained from σ where all quality and action symbols have been augmented. Let $\mathfrak{D} = (I, Q_1^{\mathfrak{D}}, \ldots, Q_n^{\mathfrak{D}}, A_1^{\mathfrak{D}}, \ldots, A_n^{\mathfrak{D}})$ designate a labelling domain over σ. A temporal graph $\mathcal{H}$, labelled with relation instances from $\mathfrak{D}$ is a structure over vocabulary $aug(\sigma) \cup \{R_H\}$, and having the universe $V \cup I \cup H$:

$$\mathcal{H}_{\mathfrak{D}} = (I \cup V \cup H, R_H^{\mathfrak{D}}, aug(Q_1^{\mathfrak{D}}), \ldots, aug(Q_n^{\mathfrak{D}}), aug(A_1^{\mathfrak{D}}), \ldots, aug(A_n^{\mathfrak{D}}))$$

where: each subset h of A from H is seen as an atomic symbol; if $a \in h$ in $\mathcal{H}$, then $(a, h) \in R_H^{\mathfrak{D}}$; for each quality edge $(a, b) \in E$ having the label $\mathcal{L}(a, b) = Q(i_1, \ldots, i_n)$, we have $(a, b, i_1, \ldots, i_n) \in aug(Q)^{\mathfrak{D}}$; for each action node $a \in A$ having the label $\mathcal{L}(a) = A(i_1, \ldots, i_n)$, we have $(a, i_1, \ldots, i_n) \in aug(A)^{\mathfrak{D}}$. We write $a \in \mathcal{H}_{\mathfrak{D}}$ iff a is part of the universe of $\mathcal{H}_{\mathfrak{D}}$.

$\mathcal{H}_{\mathfrak{D}}$ can be interpreted as a *relational database* where the universe contains individuals, hypernodes and action nodes, $R_H^{\mathfrak{D}}$ is a table where each entry (a, h) assigns to each action node a the hypernode h when a occurs, each entry $(a, b, i_1, \ldots, i_n)$ from any table $Q^{\mathfrak{D}}$ where $Q \in aug(\sigma_Q)$ indicates that individuals $i_1, \ldots, i_n$ have been enrolled in a domain relationship designated Q, and this relationship was initiated by action node a and ceased by action node b. Finally, each entry $(a, i_1, \ldots, i_n)$ from any table $A^{\mathfrak{D}}$ where $A \in aug(\sigma_A)$ indicates that individuals $i_1, \ldots, i_n$ are involved by an action designated A, and represented as a in the temporal graph.

In what follows, we prove that CON (and equally $\neg CON$) are not definable in First-Order Logic. This means that, there is no FOL-formula which is able to distinguish **t**-graphs in the class $\mathcal{C} = \{\mathcal{H} \mid CON(\mathcal{H}) = 1\}$ from any other **t**-graph outside $\mathcal{C}$. The methodology we use is due Ehrenfeucht-Fraïssé [2, 3]. Essentially, the method is comprised of the following steps: (i) fix a natural number k; (ii) build two temporal graphs, such that one is connected, while the other is not; (iii) show that no sentence from $FO[k]$ (First-Order Logic with only k nested quantifiers) can distinguish the structures: i.e. be true in one temporal graph and false in the other. If this result holds for an arbitrarily chosen k, then it naturally extends to FOL. Consider the sentence $\varphi \in FO[k]$ for some $k \in \mathbb{N}$, and assume it is given in prenex normal form:

$$\varphi = Q_1 x_1 Q_2 x_2 \ldots Q_n x_k \psi(x_1, \ldots, x_k)$$

where each Q_i is either $\forall$ or $\exists$. Checking if φ is true in some relational database, in particular, in some **t**-graph, can be seen as a game, played between two opponents: the *Falsifier* and the *Duplicator*. The game consists of k choices of elements from the universe of the database. The objective of the Falsifier is to make φ false, while that of the Duplicator is to make φ true, with each choice. Each existential quantifier ($\exists x_i$) amounts to a a choice of the Duplicator of an element from the universe. Each universal quantifier ($\forall x_i$) amounts for a similar choice of the Falsifier. If the Duplicator wins *no matter how the Falsifier plays*, then φ is true. This is the standard FOL model-checking procedure [6].

Now, consider a different game, this time played on two **t**-graphs instead of just one, where the objective of the Falsifier is: $\mathcal{H}^A \models \varphi$ and $\mathcal{H}^B \not\models \varphi$ i.e. make φ distinguish the **t**-graphs, while that of the Duplicator is: $\mathcal{H}^A \models \varphi$ and $\mathcal{H}^B \models \varphi$ i.e. make the **t**-graphs the same. Again, the game consists of k choices (or rounds). In each round, the Falsifier plays first. He must: (i) first choose between $\mathcal{H}_A$ and $\mathcal{H}_B$, and (ii) choose an element from it's universe. Next, the Duplicator must reply, by choosing another element from the opposite **t**-graph. Unlike the former game, here both players each make k individual choices. Each choice corresponds to a quantifier, and is irrelevant of it's type. Thus, given $\exists x_i \psi(x_i)$, the Falsifier must chose some a_i from the universe, if possible, such that $\psi(a_i)$ is true in $\mathcal{H}^A$ and no matter how the Duplicator replies with some b_i, the formula $\psi(b_i)$ is false in $\mathcal{H}^B$. This means that $\exists x_i \psi(x_i)$ is also false in $\mathcal{H}^B$. The same observation holds for $\forall x_i \psi(x_i)$. On the other hand, if no matter how the Falsifier plays, the Duplicator can respond, in each of the k rounds, such that the formula is true in both **t**-graphs, then the **t**-graphs are indistinguishable.

More formally, given a relational database (in particular, a **t**-graph) $\mathfrak{R}$, a *sub-database* of $\mathfrak{R} = (U, Q_1 \ldots, Q_n)$ is $\mathfrak{R}' = (U', Q_1', \ldots, Q_n')$ where $U' \subseteq U$ and each $Q_i' = \{(a, b) \mid a, b \in U', (a, b) \in Q_i\}$. A *partial isomorphism* from $\mathfrak{R}_1$ to $\mathfrak{R}_2$ is an isomorphism between a sub-database of $\mathfrak{R}_1$ to one of $\mathfrak{R}_2$.

A Ehrenfeucht-Fraïssé-game is characterized by two relational databases, from which the two players choose elements. We shall consider these databases to be **t**-graphs, namely: $\mathcal{H}_A$ and $\mathcal{H}_B$. A *strategy* in a Ehrenfeucht-Fraïssé-game is a sequence $S = I_0, I_1, \ldots, I_k$

of non-empty sets of partial isomorphisms from $\mathcal{H}_A$ to $\mathcal{H}_B$. A *winning strategy* for the Duplicator is a strategy S such that: S has the **forth property**: for every $i < k$, every partial isomorphism $f \in I_i$ and every element $\alpha \in \mathcal{H}_A$, there is a partial isomorphism $g \in I_{i+1}$ such that $\alpha \in \mathrm{dom}(g)$ and $f \subseteq g$; and S has the **back property**: for every $i < k$, every partial isomorphism $f \in I_i$ and every element $\beta \in \mathcal{H}_B$, there is a partial isomorphism $g \in I_{i+1}$ such that $\beta \in \mathrm{rng}(g)$ and $f \subseteq g$.

Each partial isomorphism I_i describes a possible evolution of the first i moves from the game. The forth property ensures the Duplicator can respond validly to any possible play I_i of i moves which are at most k, and to any choice α which the Falsifier may make in the current round, provided that α is chosen from $\mathcal{H}_A$. Similarly, the back property ensures the existence of a valid response to any possible play, when the Falsifier plays an element from $\mathcal{H}_B$. The Duplicator *wins* the k-round Ehrenfeucht-Fraïssé-game iff he has a winning strategy in every round k of the game. We build two **t**-graphs $\mathcal{H}_A$ and $\mathcal{H}_B$, and show that the Duplicator wins the k-round Ehrenfeucht-Fraïssé-game on these two structures. Without loss of generality, we assume the Duplicator and the Falsifier choose action nodes only. The following line of reasoning will still hold in the general case. The construction is as follows: let $k > 0$ be a natural number and let $n = 2 * 2^k$. We build $\mathcal{H}_A^{(k)}$ and $\mathcal{H}_B^{(k)}$ as follows:

$$U^A = \{h_1^A, \ldots, h_n^A\} \cup \{a_1, \ldots, a_{2n}\} \quad U^B = \{h_1^B, \ldots, h_n^B\} \cup \{a_1, \ldots, a_{2n}\}$$
$$R_H^A = \bigcup_{1 \le i \le n} \{(a_{2i-1}, h_i), (a_{2i}, h_i)\} \quad R_H^B = \bigcup_{1 \le i \le n} \{(a_{2i-1}, h_i), (a_{2i}, h_i)\}$$

$$E^A = \bigcup_{1 \le i \le n-1} \{(a_{2i}, a_{2i+1})\} \cup \{(a_{2n}, a_1)\}$$
$$E^B = \bigcup_{1 \le i \le n-1/2} \{(a_{2i}, a_{2i+1})\} \cup \bigcup_{n/2 \le i \le n-1} \{(a_{2i}, a_{2i+1})\} \cup \{(a_n, a_1), (a_{2n}, a_n)\}$$

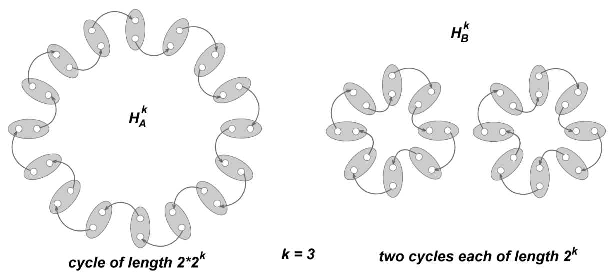

Figure 2. The temporal graphs $\mathcal{H}_A^{(k)}$ and $\mathcal{H}_B^{(k)}$ for $k = 3$

Proposition 5.1. *At any round $1 \le i < k$ from a k-round Ehrenfeucht-Fraïssé-game played on $\mathcal{H}(^k)_A$ and $\mathcal{H}(^k)_B$, there exists a trace $[x_l, x_r]$, both in $\mathcal{H}_A$ or $\mathcal{H}_B$, such that: (i) $\|[x_l, x_r]\| \ge 2^{k-1}$ and (ii) $[x_l, x_r]$ does not contain other previously-chosen elements.*

Proof. The proof is done by induction on i.

Basis: Without the loss of generality, assume x_0 is chosen in $\mathcal{H}_A$ and y_0 in $\mathcal{H}_B$. The case when x_0 is chosen in $\mathcal{H}_B$ is symmetric. Assume the Spoiler plays x_1 in $\mathcal{H}_A$. No matter how the choice is done, we have :$\|[x_1, x_0]\| + \|[x_0, x_1]\| = 2 * 2^k$, thus one of $[x_1, x_0]$, $[x_0, x_1]$, must be of length at least 2^{k-1}.

Induction step: Assume $[x_l, x_r]$ is the trace guaranteed to exist by the induction hypothesis. Let x_{i+1} be an element chosen in current round. If x_{i+1} does not belong in the trace $[x_l, x_r]$, then the property trivially holds, since $\|[x_l, x_r]\| > 2^{k-i} > 2^{k-(i+1)}$. If x_{i+1} is in $[x_l, x_r]$, then we have $\|[x_l, x_{i+1}]\| + \|[x_{i+1}, x_r]\| \geq 2^{k-i} = 2 * 2^{k-(i+1)}$. Thus, at least one of the intervals has a length larger than $2^{k-(i+1)}$. □

Proposition 5.2. *Let $x_0, \ldots, x_i$ be the elements chosen in $\mathcal{H}_A^{(k)}$, and $y_0, \ldots, y_i$ be the elements chosen in $\mathcal{H}_B^{(k)}$, in the first $i \geq 1$ rounds of a k-round Ehrenfeucht-Fraïssé game. Then, for any $0 \leq j, l \leq i$, the following holds:*

- *if $\|[y_j, y_l]\| < 2^{k-i}$ then $\|[x_j, x_l]\| = \|[y_j, y_l]\|$;*
- *if $\|[y_j, y_l]\| \geq 2^{k-i}$ then $\|[x_j, x_l]\| \geq 2^{k-i}$;*

Proof. The induction is done on i. Assume the Falsifier chooses from $\mathcal{H}_B^{(k)}$. The case when the choice is done in $\mathcal{H}_A^{(k)}$ is symmetric.

Basis. The initial choice pair is x_0, y_0. y_1 is the choice of the Falsifier. If $\|[y_0, y_1]\| < 2^{k-1}$ (i.e. y_0 and y_1 are in the same "ring" in $\mathcal{H}_B^{(1)}$), the Duplicator can chose x_1 in $\mathcal{H}_A^{(1)}$ such that $\|[x_0, x_1]\| = \|[y_0, y_1]\|$, since the unique "ring" in $\mathcal{H}_A^{(1)}$ is big enough.

If $\|[y_0, y_1]\| \geq 2^{k-1}$ then either the nodes are adjacent ($\|[y_0, y_1]\| = 1$), situation which falls in the above category, or y_0 and y_1 are in different rings. If this is so, we can simply chose x_1 such that $\|[x_0, x_1]\| = 2^{k-1}+1$. Since the number of action nodes in $\mathcal{H}_A^{(1)}$ is $2*2^k$, and no other nodes were previously selected, x_1 can be indeed chosen. Moreover, $\|[x_1, x_0]\| = 2 * 2^k - 2^{k-1} + 1 \geq 2^{k-1}$.

Induction step. We distinguish three cases, depending on how y_{i+1} is positioned in $\mathcal{H}_B$, w.r.t. the previous choices.

a) y_{i+1} is the first selected node from a ring of $\mathcal{H}_B^{(k)}$. Thus, for all previous choices y_j such that $0 \geq j \geq i$, $\|[y_{i+1}, y_j]\| = \infty$. By Proposition 5.1, there exist x_u and x_v in $\mathcal{H}_A^{(k)}$, such that $[x_u, x_v]$ has size greater or equal than 2^{k-i}, and there are no previously-chosen elements in the trace. Then, we can select x_{i+1} in the trace $[x_u, x_v]$ such that $\|[x_u, x_{i+1}]\| \geq 2^{k-(i+1)}$ and $\|[x_{i+1}, x_v]\| \geq 2^{k-(i+1)}$.

b) y_{i+1} is the second selected node from a ring of $\mathcal{H}_B^{(k)}$. Let y_j designate the first such node. We have that, for all $0 \leq l \leq i$, and $l \neq j$, $\|[y_j, y_l]\| \geq 2^{k-i}$, thus, by induction hypothesis, $\|[x_j, x_l]\| \geq 2^{k-i}$ and also $\|[x_l, x_j]\| \geq 2^{k-i}$. Thus, we have enough space to choose x_{i+1}, such that the required condition holds.

c) y_{i+1} falls between some y_j and y_l, in $\mathcal{H}_B^{(k)}$. Then:

(i) $\|[y_j, y_l]\| < 2^{k-i}$. By the induction hypothesis, $\|[x_j, x_l]\| = \|[y_j, y_l]\|$. We can simply chose x_{i+1} such that $\|[x_j, x_{i+1}]\| = \|[y_j, y_{i+1}]\|$ and $\|[x_{i+1}, x_l]\| = \|[y_{i+1}, y_l]\|$.

(ii) if $\|[y_j, y_{i+1}]\| < 2^{k-(i+1)}$, then $\|[y_{i+1}, y_l]\| \geq 2^{k-(i+1)}$. Thus, we can chose x_{i+1} such that $\|[x_j, x_{i+1}]\| = \|[x_{i+1}, x_l]\| \geq 2^{k-(i+1)}$. The case $\|[y_{i+1}, y_l]\| < 2^{k-(i+1)}$ is similar. If $\|[y_j, y_{i+1}]\| \geq 2^{k-(i+1)}$ and $\|[y_{i+1}, y_l]\| \geq 2^{k-(i+1)}$, simply choose x_{i+1} at the middle of $\|[x_j, x_l]\|$. The induction hypothesis guarantees there is enough space to ensure $\|[x_j, x_{i+1}]\| \geq 2^{k-(i+1)}$ and $\|[x_{i+1}, x_l]\| \geq 2^{k-(i+1)}$.

□

Theorem 5.1. *The Duplicator wins the k-round game, played in* $\mathcal{H}_A^{(k)}$ *and* $\mathcal{H}_B^{(k)}$.

Proof. Suppose all k rounds have been played and the choices were $x_0, \ldots, x_n$ and $y_0, \ldots, y_n$, and that $\mathcal{H}_a$, and $\mathcal{H}_b$ are the subgraphs of $\mathcal{H}_A^{(k)}$ and $\mathcal{H}_B^{(k)}$, containing only $x_0, \ldots, x_n$ and $y_0, \ldots, y_n$, respectively. Assume some y_i and y_j are adjacent. By Proposition 5.2, it follows that x_i and x_j are also adjacent. Thus, $\mathcal{H}_a$ and $\mathcal{H}_b$ are isomorphic. Assume y_i and y_j are at a distance larger than 1. Then x_i and x_j are also at a distance larger than 1. Again, $\mathcal{H}_a$ and $\mathcal{H}_b$ are isomorphic. □

Theorem 5.2 (Ehrenfeucht-Fraïssé [2, 3])**.** *The following statements are equivalent:*

- *for any* $\varphi \in FO[k]$, $\varphi \models \mathfrak{R}_A$ *iff* $\varphi \models \mathfrak{R}_B$*;*
- *the Duplicator wins the k-round Ehrenfeucht-Fraïssé-game, played on* $\mathfrak{R}_A$ *and* $\mathfrak{R}_B$.

Proposition 5.3 (Corrolary to 5.1,5.2)**.** CON *is not FOL-definable.*

6. Conclusion

An immediate consequence of our result is that relational databases cannot be used in order to perform temporal reasoning over temporal graphs.

We believe this result to be fundamental for the endeavours of [4, 8, 5, 7]*, since it shows that the previously-mentioned approaches are not mere syntactic sugars for already established (temporal) reasoning procedures.*

REFERENCES

[1] *James F. Allen.* Planning as Temporal Reasoning. In James F. Allen, Richard Fikes, and Erik Sandewall, editors, *KR'91: Principles of Knowledge Representation and Reasoning*, pages 3–14. Morgan Kaufmann, San Mateo, California, 1991.

[2] *A. Ehrenfeucht.* An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141, 1961.

[3] *Fraïseé, R.* Sur quelques classifications des systemes de relations. *Universite d'Alger, Publications Scientifiques*, Serie A, 1:35–182, 1954.

[4] *Cristian Giumale, Lorina Negreanu, Mihnea Muraru, and Matei Popovici.* Modeling ontologies for time-dependent applications. *Symbolic and Numeric Algorithms for Scientific Computing, International Symposium on*, 0:202–208, 2010.

[5] *Cristian Giumale, Lorina Negreanu, Mihnea Muraru, Matei Popovici, Alexandru Agache, and Ciprian Dobre.* Modeling with fluid qualities. In *CSCS-18.* 18th International Conference on Control Systems and Computer Science, 2011.

[6] *Leonid Libkin. Elements Of Finite Model Theory (Texts in Theoretical Computer Science. An Eatcs Series).* Springer Verlag, 2004.

[7] *Matei Popovici.* Using evolution graphs for describing topology-aware prediction models in large clusters. In Michael Fisher, Leon van der Torre, Mehdi Dastani, and Guido Governatori, editors, *Computational Logic in Multi-Agent Systems*, volume 7486 of *Lecture Notes in Computer Science*, pages 94–109. Springer Berlin / Heidelberg, 2012.

[8] *Matei Popovici, Mihnea Muraru, Alexandru Agache, Cristian Giumale, Lorina Negreanu, and Ciprian Dobre.* A modeling method and declarative language for temporal reasoning based on fluid qualities. In *Proceedings of the 19th international conference on Conceptual structures for discovering knowledge*, ICCS'11, pages 215–228, Berlin, Heidelberg, 2011. Springer-Verlag.