Proc. 23rd Int. Workshop on Description Logics (DL2010), CEUR-WS 573, Waterloo, Canada, 2010.

Temporal Conceptual Modelling with *DL-Lite*

A. Artale,[1] R. Kontchakov,[2] V. Ryzhikov,[1] and M. Zakharyaschev[2]

[1] KRDB Research Centre
Free University of Bozen-Bolzano
I-39100 Bolzano, Italy
lastname@inf.unibz.it

[2] Dept. of Comp. Science and Inf. Sys.
Birkbeck College
London WC1E 7HX, UK
{roman,michael}@dcs.bbk.ac.uk

1 Introduction

Conceptual modelling formalisms such as the Entity-Relationship model (ER) and Unified Modelling Language (UML) have become a *de facto* standard in database design by providing visual means to describe application domains in a declarative and reusable way. On the other hand, both ER and UML turned out to be closely connected with description logics that are underpinned by formal semantics and thus capable of providing services for effective reasoning over conceptual models; see, e.g., [11, 4].

Temporal conceptual data models (TCMs) [18, 25] have been introduced in the context of temporal databases [20, 15, 13]. In this case, apart from the classical constructs—such as inheritance between classes and relationships, cardinality constraints restricting participation in relationships, and disjointness and covering constraints—temporal constructs are used to capture the temporal behaviour of various components of conceptual schemas. Such constructs can be grouped into 3 categories. *Timestamping constraints* discriminate between those classes, relationships and attributes that change over time and those that are time-invariant [28, 18, 16, 6, 25]. *Evolution constraints* control how domain elements evolve over time by 'migrating' from one class to another [19, 23, 26, 25, 3]. We distinguish between *qualitative* evolution constraints describing generic temporal behaviour, and *quantitative* ones specifying the exact moment of migration. *Temporal cardinality constraints* restrict the number of times an instance of a class participates in a relationship. *Snapshot* cardinality constraints do it at each moment of time, while *lifespan* cardinality constraints impose restrictions over the entire existence of the instance as a member of the class [27, 22].

Temporal conceptual data models can be encoded in various temporal description logics (TDLs), which have been designed and investigated since the seminal paper [24] with the aim of understanding the computational price of introducing a temporal dimension in DLs; see [21] for a recent survey. A general conclusion one can draw from the obtained results is that—as far as there is nontrivial interaction between the temporal and DL components—TDLs based on full-fledged DLs like $\mathcal{ALC}$ turn out to be too complex for effective reasoning (see the end of the introduction for details).

The aim of this paper is to tailor 'minimal' TDLs that are capable of representing various aspects of TCMs and investigate their computational behaviour. First of all, as the DL component we choose the 'light-weight' *DL-Lite* logic

$\mathit{DL\text{-}Lite}^{\mathcal{N}}_{bool}$, which was shown to be adequate for capturing conceptual models without relationship inheritance[1] [4], and its fragment $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{core}$ with most primitive concept inclusions, which are nevertheless enough to represent almost all types of constraints (apart from covering). To discuss our choice of the temporal constructs, consider a toy TCM describing a company.

For the timestamping constraint 'employee is a *snapshot class*' (by the standard TCM terminology, such a class never changes in time) one can use the axiom $\mathsf{Employee} \sqsubseteq \boxast \mathsf{Employee}$ with the temporal operator $\boxast$ 'always.' Likewise, the constraint 'manager is a *temporary class*' in the sense that each of its instances must leave the class, the axiom $\mathsf{Manager} \sqsubseteq \diamondast \neg \mathsf{Manager}$ is required, where $\diamondast$ means 'some time.' Both of these axioms are regarded as *global*, i.e., applicable to all time points. Note that to express $\diamondast$ using more standard temporal constructs, we need both 'some time in the past' $\Diamond_P$ and 'some time in the future' $\Diamond_F$: e.g., $\diamondast = \Diamond_P \Diamond_F$. To encode a snapshot n-ary relationship, one can reify it into a snapshot class with n auxiliary *rigid*—i.e., time-independent—roles; for a temporary relationship, the reifying class is temporary and the roles are *local* [9, 7]. The qualitative evolution constraints 'each manager was once an employee' and 'a manager will always remain a manager' can be expressed by the axioms $\mathsf{Manager} \sqsubseteq \Diamond_P \mathsf{Employee}$ and $\mathsf{Manager} \sqsubseteq \Box_F \mathsf{Manager}$, while 'an approved project keeps its status until a later date when it actually starts' can be expressed using the 'until' operator: $\mathsf{ApprovedProject} \sqsubseteq \mathsf{ApprovedProject}\,\mathcal{U}\,\mathsf{Project}$. The quantitative evolution constraint 'each project must be finished in 3 years' requires the next-time operator $\bigcirc_F$: $\mathsf{Project} \sqsubseteq \bigcirc_F \bigcirc_F \bigcirc_F \mathsf{FinishedProject}$. The snapshot cardinality constraint 'an employee can work on at most 2 projects at each moment of time' can be expressed as $\mathsf{Employee} \sqsubseteq\, \leq 2\, \mathsf{worksOn}$, while the lifespan constraint 'over the whole career, an employee can work on at most 5 projects' requires temporal operators on roles: $\mathsf{Employee} \sqsubseteq\, \leq 5\, \diamondast \mathsf{worksOn}$. Note that 'temporalised' roles of the form $\diamondast R$ and $\boxast R$ are always rigid. To represent a temporal database instance of a TCM, we use assertions like $\bigcirc_P \mathsf{Manager}(\mathsf{bob})$ for 'Bob was a manager last year' and $\bigcirc_F \mathsf{manages}(\mathsf{bob}, \mathsf{cronos})$ for 'Bob will manage project Cronos next year.' As usual, n-ary tables are represented via reification.

These considerations lead us to TDLs based on $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{bool}$ and $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{core}$ and interpreted over the flow of time $(\mathbb{Z}, <)$, in which (1) the future and past temporal operators can be applied to concepts; (2) roles can be declared local or rigid; (3) the 'undirected' temporal operators 'always' and 'some time' can be applied to roles; (4) the concept inclusions (TBox axioms) are global and the database (ABox) assertions are specified to hold at particular moments of time.

To our surprise, the most expressive TDL based on $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{bool}$ and featuring all of (1)–(4) turns out to be undecidable. As follows from the proof of Theorem 5 below, it is a subtle interaction of functionality constraints on temporalised roles with the next-time operator and full Booleans on concepts that causes undecidability. This 'negative' result motivates consideration of various fragments of our full TDL by restricting not only the DL but also the temporal component. The table below illustrates the expressive power of the resulting fragments in the context of TCMs. We also note that both $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{bool}$ and $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{core}$ with global

[1] $\mathit{DL\text{-}Lite}^{\mathcal{N}}_{bool}$ with relationship inclusions regains the full expressive power of $\mathcal{ALC}$.

axioms can capture snapshot cardinality constraints, while lifespan cardinality constraints require temporalised roles of the form $\diamondast R$ and $\boxast R$.

concept temporal operators	timestamping	evolution	
		qualitative	quantitative
$\mathcal{U}/\mathcal{S}$	$+$	$+$	$+$
$\Box_{F/P}, \bigcirc_{F/P}$	$+$	$+$	$+$
$\Box_{F/P}$	$+$	$+$	$-$
$\boxast, \bigcirc_{F/P}$	$+$	$-$	$+$
$\boxast$	$+$	$-$	$-$

The next table summarises the complexity results obtained in this paper for satisfiability of temporal knowledge bases formulated in our TDLs.

concept temporal operators	local & rigid roles only		temporalised roles
	DL-Lite$_{bool}^{\mathcal{N}}$	*DL-Lite*$_{core}^{\mathcal{N}}$	*DL-Lite*$_{bool}^{\mathcal{N}}$
$\mathcal{U}/\mathcal{S}$	PSPACE Thm. 1	PSPACE [8]	undec. Thm. 5
$\Box_{F/P}, \bigcirc_{F/P}$	PSPACE Thm. 2 (*ii*)	NP Thm. 3	undec. Thm. 5
$\Box_{F/P}$	NP Thm. 2 (*i*)	NP [8]	?
$\boxast, \bigcirc_{F/P}$	PSPACE Thm. 2 (*ii*)	NP Thm. 3	undec. Thm. 5
$\boxast$	NP Thm. 2 (*i*)	NLOGSPACE Thm. 4	NP Thm. 6

Apart from the undecidability result of Theorem 5, quite surprising is NP-completeness of the temporal extension of *DL-Lite*$_{core}^{\mathcal{N}}$ with the operators $\Box_F$ and $\bigcirc_F$ (and their past counterparts) on concepts provided by Theorem 3. Indeed, if full Booleans are available, even the propositional temporal logic with these operators is PSPACE-complete. Moreover, if the 'until' operator $\mathcal{U}$ is available in the temporal component, disjunction is expressible even with *DL-Lite*$_{core}^{\mathcal{N}}$ as the underlying DL, and the logic becomes PSPACE-complete [8]. In all other cases, the complexity of TDL reasoning coincides with the maximal complexity of reasoning in the component logics (despite nontrivial interaction between them, as none of our TDLs is a fusion of its components). It is also of interest to observe the dramatic increase of complexity caused by the addition of $\bigcirc_F$ to the logic in the lower right corner of the table (from NP to undecidability).

To put this paper in the more general context of temporal description logics, we note first that our TDLs extend those in [8] with the past-time operators $\mathcal{S}$, $\Box_P$, $\Diamond_P$, $\bigcirc_P$ over $\mathbb{Z}$ (which are essential for capturing timestamping constraints), universal modalities $\boxast$ and $\diamondast$, and temporalised roles. Temporal operators on *DL-Lite* axioms and concepts in the presence of rigid roles were investigated in [7], where it was shown that the resulting temporalisations of *DL-Lite*$_{bool}^{\mathcal{N}}$ and *DL-Lite*$_{horn}^{\mathcal{N}}$ are EXPSPACE-complete. Temporal extensions of the standard DL $\mathcal{ALC}$ feature the following computational behaviour: $\mathcal{ALC}$ with temporal operators on axioms, rigid concepts and roles is 2EXPTIME-complete [10]. It is EXPSPACE-complete if temporal operators on concepts and axioms are allowed but no rigid or temporalised roles are available [17], and EXPTIME-complete if the language allows only temporalised concepts and global axioms [24, 2]. Finally, the 'undirected' temporal operators $\boxast$ and $\diamondast$ on concepts and roles together with global axioms result in a 2EXPTIME-complete extension of $\mathcal{ALC}$ [9].

2 Temporal DLs based on $\textit{DL-Lite}_{bool}^{\mathcal{N}}$

The TDL $T_{\mathcal{US}}\textit{DL-Lite}_{bool}^{\mathcal{N}}$ is based on $\textit{DL-Lite}_{bool}^{\mathcal{N}}$ [1, 5], which, in turn, extends $\textit{DL-Lite}_{\sqcap,\mathcal{F}}$ [12] with full Booleans over concepts and cardinality restrictions over roles. The language of $T_{\mathcal{US}}\textit{DL-Lite}_{bool}^{\mathcal{N}}$ contains *object names* $a_0, a_1, \ldots$, *concept names* $A_0, A_1, \ldots$, *local role names* $P_0, P_1, \ldots$ and *rigid role names* $G_0, G_1, \ldots$. *Roles* R, *basic concepts* B and *concepts* C are defined as follows:

$$
\begin{array}{lcl}
S & ::= & P_i \quad | \quad G_i, \qquad\qquad R \quad ::= \quad S \quad | \quad S^-,\\
B & ::= & \bot \quad | \quad A_i \quad | \quad \geq q\, R,\\
C & ::= & B \quad | \quad \neg C \quad | \quad C_1 \sqcap C_2 \quad | \quad C_1 \,\mathcal{U}\, C_2 \;|\; \quad C_1 \,\mathcal{S}\, C_2,
\end{array}
$$

where $q \geq 1$ is a natural number (the results obtained below do not depend on whether q is given in unary or binary). A $T_{\mathcal{US}}\textit{DL-Lite}_{bool}^{\mathcal{N}}$ *interpretation* is a function $\mathcal{I}$ on the integers $\mathbb{Z}$ (the intended flow of time):

$$
\mathcal{I}(n) \;=\; \big(\Delta^{\mathcal{I}}, a_0^{\mathcal{I}}, \ldots, A_0^{\mathcal{I}(n)}, \ldots, P_0^{\mathcal{I}(n)}, \ldots, G_0^{\mathcal{I}(n)}, \ldots\big),
$$

where $\Delta^{\mathcal{I}}$ is a nonempty set, the (constant) domain of $\mathcal{I}$, $a_i^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, $A_i^{\mathcal{I}(n)} \subseteq \Delta^{\mathcal{I}}$ and $P_i^{\mathcal{I}(n)}, G_i^{\mathcal{I}(n)} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ with $G_i^{\mathcal{I}(n)} = G_i^{\mathcal{I}(m)}$, for $i \in \mathbb{N}$ and $n, m \in \mathbb{Z}$. We adopt the unique name assumption according to which $a_i^{\mathcal{I}} \neq a_j^{\mathcal{I}}$, for $i \neq j$, although our complexity results would not change if we dropped it, apart from the NLogSpace bound of Theorem 4, which would increase to NP [5]. The role and concept constructs are interpreted in $\mathcal{I}$ as follows:

$$
\begin{array}{l}
(S^-)^{\mathcal{I}(n)} = \{(y,x) \mid (x,y) \in S^{\mathcal{I}(n)}\}, \quad \bot^{\mathcal{I}(n)} = \emptyset, \quad (\neg C)^{\mathcal{I}(n)} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}(n)},\\
(C_1 \sqcap C_2)^{\mathcal{I}(n)} = C_1^{\mathcal{I}(n)} \cap C_2^{\mathcal{I}(n)}, \quad (\geq q\, R)^{\mathcal{I}(n)} = \big\{x \mid \sharp\{y \mid (x,y) \in R^{\mathcal{I}(n)}\} \geq q\big\},\\
(C_1 \,\mathcal{U}\, C_2)^{\mathcal{I}(n)} = \bigcup_{k>n}\big(C_2^{\mathcal{I}(k)} \cap \bigcap_{n<m<k} C_1^{\mathcal{I}(m)}\big),\\
(C_1 \,\mathcal{S}\, C_2)^{\mathcal{I}(n)} = \bigcup_{k<n}\big(C_2^{\mathcal{I}(k)} \cap \bigcap_{n>m>k} C_1^{\mathcal{I}(m)}\big).
\end{array}
$$

Note that our *until* and *since* operators are 'strict' (i.e., do not include the current moment). We also use the temporal operators $\Diamond_F$ ('some time in the future'), $\Diamond_P$ ('some time in the past'), $\Diamond\!\!\!\!\ast$ ('some time'), their duals $\Box_F$, $\Box_P$ and $\Box\!\!\!\!\ast$, $\bigcirc_F$ ('next time') and $\bigcirc_P$ ('previous time'), which are all expressible by means of $\mathcal{U}$ and $\mathcal{S}$, e.g., $\Diamond_F C = \neg\bot\mathcal{U}C$, $\Box_F C = \neg\Diamond_F\neg C$, $\bigcirc_F C = \bot\mathcal{U}C$, $\Diamond\!\!\!\!\ast\, C = \Diamond_F\Diamond_P C$ and $\Box\!\!\!\!\ast\, C = \Box_F\Box_P C$. (Other standard abbreviations we use include $C_1 \sqcup C_2$, $\exists R$ and $\top = \neg\bot$.) Apart from full $T_{\mathcal{US}}\textit{DL-Lite}_{bool}^{\mathcal{N}}$, we consider a few of its sublanguages allowing only some of the (definable) temporal operators mentioned above:

- $T_{FP}\textit{DL-Lite}_{bool}^{\mathcal{N}}$, which allows only $\Diamond_F C$, $\Diamond_P C$ and their duals (but no $\bigcirc_F C$ or $C_1 \,\mathcal{U}\, C_2$), and its extension $T_{FPX}\textit{DL-Lite}_{bool}^{\mathcal{N}}$ with $\bigcirc_F C$ and $\bigcirc_P C$;
- $T_U\textit{DL-Lite}_{bool}^{\mathcal{N}}$, allowing only $\Diamond\!\!\!\!\ast\, C$ and $\Box\!\!\!\!\ast\, C$, and its extension $T_{UX}\textit{DL-Lite}_{bool}^{\mathcal{N}}$ with $\bigcirc_F C$ and $\bigcirc_P C$.

A *TBox*, $\mathcal{T}$, in any of our languages $\mathcal{L}$ is a finite set of *concept inclusions* (CIs) of the form $C_1 \sqsubseteq C_2$, where the C_i are $\mathcal{L}$-concepts. An *ABox*, $\mathcal{A}$, consists of assertions of the form $\bigcirc^n B(a)$ and $\bigcirc^n S(a,b)$, where B is a basic concept, S a (local or rigid) role name, a,b object names and $\bigcirc^n$, for $n \in \mathbb{Z}$, is a sequence of n operators $\bigcirc_F$ if $n \geq 0$ and $|n|$ operators $\bigcirc_P$ if $n < 0$. Taken together, the TBox $\mathcal{T}$ and ABox $\mathcal{A}$ form the *knowledge base* (KB) $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ in $\mathcal{L}$.

The *truth-relation* is defined as usual: $\mathcal{I} \models C_1 \sqsubseteq C_2$ iff $C_1^{\mathcal{I}(n)} \subseteq C_2^{\mathcal{I}(n)}$, for all $n \in \mathbb{Z}$, that is, we interpret concept inclusions globally, $\mathcal{I} \models \bigcirc^n B(a)$ iff $a^{\mathcal{I}} \in B^{\mathcal{I}(n)}$, and $\mathcal{I} \models \bigcirc^n S(a,b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in S^{\mathcal{I}(n)}$. We call $\mathcal{I}$ a *model* of a KB $\mathcal{K}$ and write $\mathcal{I} \models \mathcal{K}$ if $\mathcal{I} \models \alpha$ for all α in $\mathcal{K}$. If $\mathcal{K}$ has a model then it is said to be *satisfiable*. A concept C (role R) is *satisfiable* w.r.t. $\mathcal{K}$ if there are a model $\mathcal{I}$ of $\mathcal{K}$ and $n \in \mathbb{Z}$ such that $C^{\mathcal{I}(n)} \neq \emptyset$ (respectively, $R^{\mathcal{I}(n)} \neq \emptyset$). Clearly, the concept and role satisfiability problems are equivalent to KB satisfiability.

Our first result states that the satisfiability problem for $T_{\mathcal{US}}$*DL-Lite*$_{bool}^{\mathcal{N}}$ KBs is as complex as satisfiability in propositional temporal logic *LTL*.

Theorem 1. *Satisfiability of* $T_{\mathcal{US}}$*DL-Lite*$_{bool}^{\mathcal{N}}$ *KBs is* PSPACE*-complete.*

The proof is by a two-step (non-deterministic polynomial) reduction to *LTL*. First, we reduce satisfiability of a $T_{\mathcal{US}}$*DL-Lite*$_{bool}^{\mathcal{N}}$ KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ to satisfiability in the one-variable first-order temporal logic in a way similar to [8]. For each basic concept B $(\neq \bot)$, we take a fresh *unary* predicate $B^*(x)$ and encode $\mathcal{T}$ as

$$\mathcal{T}^\dagger \quad = \quad \bigwedge_{C_1 \sqsubseteq C_2 \in \mathcal{T}} \boxast \forall x \left(C_1^*(x) \to C_2^*(x)\right),$$

where the C_i^* are the results of replacing each B with $B^*(x)$ ($\sqcap$ with $\land$, etc.). We assume that $\mathcal{T}$ contains CIs of the form $\geq q\, R \sqsubseteq\, \geq q'\, R$, for $\geq q\, R$, $\geq q'\, R$ in $\mathcal{T}$ such that $q > q'$ and there is no q'' with $q > q'' > q'$ and $\geq q''\, R$ in $\mathcal{T}$. We also assume that $\mathcal{T}$ contains $\geq q\, R \equiv \boxast \geq q\, R$ if $\geq q\, R$ occurs in $\mathcal{T}$, for a rigid role R (i.e., for G_i or G_i^-). To take account of the fact that roles are *binary* relations, we add to $\mathcal{T}^\dagger$ the following formula, for each role name S:

$$\varepsilon_S \quad = \quad \boxast\big(\exists x\, (\exists S)^*(x) \leftrightarrow \exists x\, (\exists S^-)^*(x)\big)$$

(which says that at each moment of time the domain of S is nonempty iff its range is nonempty). The ABox $\mathcal{A}$ is encoded by a conjunction $\mathcal{A}^\dagger$ of ground atoms of the form $\bigcirc^m B^*(a)$ and $\bigcirc^n (\geq q\, R)^*(a)$ in the same way as in [8]. Thus, $\mathcal{K}$ is satisfiable iff the formula

$$\mathcal{K}^\dagger \quad = \quad \mathcal{T}^\dagger \quad \land \quad \bigwedge_S \varepsilon_S \quad \land \quad \mathcal{A}^\dagger$$

is satisfiable. The second step of our reduction is based on the observation that if $\mathcal{K}^\dagger$ is satisfiable then it can be satisfied in a model such that

(R) if $(\exists S)^*(x)$ is true at some moment (on some domain element) then it is true at all moments of time (perhaps on different domain elements).

Indeed, if $\mathcal{K}^\dagger$ is satisfied in $\mathcal{I}$ then it is satisfied in the disjoint union $\mathcal{I}^*$ of all $\mathcal{I}^n$, $n \in \mathbb{Z}$, obtained from $\mathcal{I}$ by shifting its time line n moments forward. It follows from **(R)** that $\mathcal{K}^\dagger$ is satisfiable iff there is a set Σ of role names such that

$$\mathcal{K}^{\dagger\Sigma} \;=\; \mathcal{T}^\dagger \;\wedge\; \bigwedge_{S\in\Sigma} \big((\exists S)^*(d_S) \wedge (\exists S^-)^*(d_{S^-})\big) \;\wedge\; \bigwedge_{S\notin\Sigma} \boxast \forall x\, \neg\big((\exists S)^*(x) \vee (\exists S^-)^*(x)\big) \;\wedge\; \mathcal{A}^\dagger$$

is satisfiable, where the d_S are fresh constants (informally, the roles in Σ are nonempty at some moment, whereas all other roles are always empty). Finally, as $\mathcal{K}^{\dagger\Sigma}$ contains no existential quantifiers, it can be regarded as an *LTL*-formula because all the universal quantifiers can be instantiated by all the constants in the formula, which results only in a polynomial blow-up of $\mathcal{K}^{\dagger\Sigma}$.

This reduction can also be used to obtain complexity results for the fragments of $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{bool}$ mentioned above. Using the well-known facts that satisfiability in the fragments of *LTL* with $\Diamond_F/\Diamond_P$ and with $\diamondast$ is NP-complete, and that the extension of any of these fragments with $\bigcirc_F/\bigcirc_P$ becomes PSPACE-complete again, we obtain:

Theorem 2. (i) *Satisfiability of* T_{FP}*DL-Lite*$^{\mathcal{N}}_{bool}$ *and* T_U*DL-Lite*$^{\mathcal{N}}_{bool}$ *KBs is* NP-*complete.* (ii) *For* T_{FPX}*DL-Lite*$^{\mathcal{N}}_{bool}$ *and* T_{UX}*DL-Lite*$^{\mathcal{N}}_{bool}$ *KBs, satisfiability is* PSPACE-*complete.*

3 Temporal DLs based on *DL-Lite*$^{\mathcal{N}}_{core}$

So far, to decrease complexity we have restricted the expressive power of the temporal component of $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{bool}$. But the underlying DL *DL-Lite*$^{\mathcal{N}}_{bool}$ also has some natural fragments of lower complexity [5]. In this section, we consider the simplest of them known as *DL-Lite*$^{\mathcal{N}}_{core}$ and containing only CIs of the form $B_1 \sqsubseteq B_2$ and $B_1 \sqcap B_2 \sqsubseteq \bot$, where the B_i are basic concepts. Satisfiability of *DL-Lite*$^{\mathcal{N}}_{core}$ KBs is NLOGSPACE-complete.

Let $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{core}$ be the fragment of $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{bool}$ with CIs of the form $D_1 \sqsubseteq D_2$ and $D_1 \sqcap D_2 \sqsubseteq \bot$, where the D_i are defined by the rule:

$$D \;::=\; B \;\mid\; B_1\,\mathcal{U}\,B_2 \mid\; B_1\,\mathcal{S}\,B_2.$$

By restricting D_1 and D_2 to concepts of the form

$$D \;::=\; B \;\mid\; \Diamond_F B \;\mid\; \Diamond_P B \;\mid\; \Box_F B \;\mid\; \Box_P B$$

we obtain T_{FP}*DL-Lite*$^{\mathcal{N}}_{core}$. These restrictions do not improve the complexity of reasoning: satisfiability of $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{core}$ KBs is PSPACE-complete, while for T_{FP}*DL-Lite*$^{\mathcal{N}}_{core}$ it is NP-complete [8].

What is really surprising and nontrivial is that extending T_{FP}*DL-Lite*$^{\mathcal{N}}_{core}$ with the next- and previous-time operators does not increase the complexity; cf. Theorem 2 (ii). More formally, define T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ by restricting D_1 and D_2 to concepts of the form:

$$D \;::=\; B \;\mid\; \Diamond_F B \;\mid\; \Diamond_P B \;\mid\; \Box_F B \;\mid\; \Box_P B \;\mid\; \bigcirc_F B \;\mid\; \bigcirc_P B,$$

and let T_{UX}*DL-Lite*$^{\mathcal{N}}_{core}$ be the logic with the D_i of the form:

$$D \;::=\; B \;\mid\; \diamondast B \;\mid\; \boxast B \mid\; \bigcirc_F B \;\mid\; \bigcirc_P B.$$

Theorem 3. *Satisfiability of T_{FPX} DL-Lite$^{\mathcal{N}}_{core}$ and T_{UX} DL-Lite$^{\mathcal{N}}_{core}$ KBs is* NP-*complete.*

We present only a sketch of the proof here; the full proof can be found at `http://www.dcs.bbk.ac.uk/~roman/papers/dl10-full.pdf`.

In a way similar to the proof of Theorem 1, one can (non-deterministically and polynomially) reduce satisfiability of a T_{FPX} *DL-Lite*$^{\mathcal{N}}_{core}$ KB to satisfiability of an *LTL*-formula $\varphi = \bigwedge_i \boxast(E_i \vee E'_i) \wedge \psi$, where the E_i and E'_i are of the form p, $\Diamond_F p$, $\Diamond_P p$, $\Box_F p$, $\Box_P p$, $\bigcirc_F p$, $\bigcirc_P p$ or a negation thereof, and ψ is a conjunction of formulas of the form $\bigcirc^n p$, p a propositional variable. Let Γ be the set of all subformulas of φ of the form $\Diamond_F p$, $\Diamond_P p$, $\Box_F p$ or $\Box_P p$. It should be clear that if φ is satisfied in an interpretation then the flow of time can be partitioned into $|\Gamma| + 1$ intervals $I_0, \dots, I_{|\Gamma|}$ such that, for each $\gamma \in \Gamma$ and each I_i, γ is true at *some* point in I_i iff γ is true at *every* point in I_i. The existence of such intervals can be expressed by certain syntactic conditions on their 'states,' the most crucial of which is satisfiability of a formula of the form

$$\chi \quad = \quad \Psi \wedge \Box^{\leq m}\Phi \wedge \bigcirc^m(\Psi' \wedge \bigcirc\Psi''),$$

for $\Phi = \bigwedge_i (D_i \vee D'_i)$, with each of the D_i and D'_i being a literal L (a propositional variable or its negation) or $\bigcirc L$, conjunctions Ψ, Ψ' and Ψ'' of literals, and $m \geq 0$, where $\bigcirc^n\Psi$ is the result of attaching n operators $\bigcirc$ to each literal in Ψ and $\Box^{\leq m}\Phi = \bigwedge_{0 \leq i \leq m} \bigcirc^i\Phi$. Intuitively, m is the number of distinct states in an interval I_i, Ψ and Ψ' are the first and the last states in I_i, Ψ'' is the first state of the next interval I_{i+1}, and Φ a set of binary clauses that describe possible transitions between the states. Let $cons^m_\Phi(\Psi)$ be the set of all literals L that are true at the moment $m \geq 0$ in every model of $\Psi \wedge \Box^{\leq m}\Phi$. As the formula $\Psi \wedge \Box^{\leq m}\Phi$ is essentially a 2CNF, one can compute $cons^m_\Phi(\Psi)$ inductively as follows:

$$cons^0_\Phi(\Psi) = \{L \mid \Phi \cup \Psi \models L\},$$
$$cons^m_\Phi(\Psi) = \{L \mid \Phi \models L' \rightarrow \bigcirc L, L' \in cons^{m-1}_\Phi(\Psi)\} \cup \{L \mid \Phi \models L\}.$$

For each L, construct a non-deterministic finite automaton $\mathfrak{A}_L = (Q, Q_0, \sigma, F_L)$ over the alphabet $\{0\}$ that accepts 0^m iff $L \in cons^m_\Phi(\Psi)$. Define the states in Q to be all the literals from χ, the set of initial states $Q_0 = cons^0_\Phi(\Psi)$, the accepting states $F_L = \{L\}$, and the transition relation

$$\sigma = \{(L'', L') \mid \Phi \models L'' \rightarrow \bigcirc L'\} \cup \{(L', L') \mid \Phi \models L'\}.$$

Then a state L is reachable in m σ-steps from a state in Q_0 iff $L \in cons^m_\Phi(\Psi)$, and so $\mathfrak{A}_L$ is as required. Every such $\mathfrak{A}_L$ can be converted into an equivalent automaton in the Chrobak normal form [14] using Martinez's algorithm [29], which gives rise to M_L-many arithmetic progressions $a^L_1 + b^L_1\mathbb{N}, \dots, a^L_{M_L} + b^L_{M_L}\mathbb{N}$, where $a + b\mathbb{N} = \{a + bn \mid n \in \mathbb{N}\}$, such that

(A$_1$) $M_L, a^L_i, b^L_i \leq |\Phi \cup \Psi|^2$, for $1 \leq i \leq M_L$, and
(A$_2$) $L \in cons^m_\Phi(\Psi)$ iff $m \in \bigcup_{i=1}^{M_L}(a^L_i + b^L_i\mathbb{N})$.

Satisfiability of χ can now be established by a polynomial-time algorithm which checks whether the following three conditions hold:

1. $p, \neg p \in cons^n_{\Phi}(\Psi)$, for no variable p and no $0 \leq n \leq m+1$;
2. $\neg L \notin cons^m_{\Phi}(\Psi)$, for all literals $L \in \Psi'$;
3. $\neg L \notin cons^{m+1}_{\Phi}(\Psi)$, for all literals $L \in \Psi''$.

To verify 1, we check, for each variable p, whether the linear Diophantine equations $a^p_i + b^p_i x = a^{\neg p}_j + b^{\neg p}_j y$, for $1 \leq i \leq M_p$ and $1 \leq j \leq M_{\neg p}$, have a solution (x_0, y_0) such that $0 \leq a^p_i + b^p_i x_0 \leq m+1$. Set $a = b^p_i$, $b = -b^{\neg p}_j$ and $c = a^{\neg p}_j - a^p_i$, which gives us the equation $ax + by = c$. If $a \neq 0$ and $b \neq 0$ then, by Bézout's lemma, it has a solution iff c is a multiple of the greatest common divisor d of a and b, which can be checked in polynomial time using the Euclidean algorithm (provided that the numbers are encoded in unary, which can be assumed in view of $(\mathbf{A}_1)$). Moreover, if the equation has a solution, then the Euclidean algorithm also gives us a pair (u_0, v_0) such that $d = au_0 + bv_0$, in which case all the solutions of the above equation form the set $\big\{\big((cu_0 + bk)/d, (cv_0 - ak)/d\big) \mid k \in \mathbb{Z}\big\}$. Thus, it remains to check whether a number between 0 and $m+1$ is contained in $a^p_i + b^p_i(a^{\neg p}_j - a^p_i)u_0/d + b^p_i b^{\neg p}_j/d\mathbb{N}$. The case $a = 0$ or $b = 0$ is left to the reader. To verify condition 2, we check, for each $L \in \Psi'$, whether m belongs to one of $a^{\neg L}_i + b^{\neg L}_i\mathbb{N}$, for $1 \leq i \leq M_L$, which can be done in polynomial time. Condition 3 is analogous. This gives us the NP upper bound for the logics mentioned in Theorem 3. The lower bound can be proved by reduction of the 3-colourability problem to satisfiability of $T_{U X}$*DL-Lite*$^{\mathcal{N}}_{core}$ KBs.

Theorem 3 shows that T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ can be regarded as a good candidate for representing temporal conceptual data models. Although not able to express covering constraints, T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ still appears to be a reasonable compromise compared to the full PSPACE-complete logic T_{FPX}*DL-Lite*$^{\mathcal{N}}_{bool}$.

By restricting the temporal constructs to the undirected universal modalities $\boxast$ and $\diamondast$, we obtain an even simpler logic:

Theorem 4. *Satisfiability of* T_U*DL-Lite*$^{\mathcal{N}}_{core}$ *KBs is* NLOGSPACE-*complete.*

The proof of the upper bound is by embedding into the universal Krom fragment of first-order logic.

4 Temporal DLs with Temporalised Roles

As we have seen before, in order to express lifespan cardinalities, temporal operators on roles are required. Modalised roles are known to be 'dangerous' and very difficult to deal with when temporalising expressive DLs such as $\mathcal{ALC}$ [17, Section 14.2]. To our surprise, even in the case of *DL-Lite*, temporal operators on roles may cause undecidability (while rigid roles are 'mostly harmless'). Denote by $T^{\mathcal{R}}_X$*DL-Lite*$^{\mathcal{N}}_{bool}$ the fragment of $T_{\mathcal{US}}$*DL-Lite*$^{\mathcal{N}}_{bool}$ with $\bigcirc_F$ as the only temporal operator over concepts and with roles R of the form

$$R \quad ::= \quad S \quad | \quad S^- \quad | \quad \diamondast R \quad | \quad \boxast R.$$

The extensions of $\diamondast R$ and $\boxast R$ in an interpretation $\mathcal{I}$ are defined as follows:

$$(\diamondast R)^{\mathcal{I}(n)} = \bigcup\nolimits_{k \in \mathbb{Z}} R^{\mathcal{I}(k)} \quad \text{and} \quad (\boxast R)^{\mathcal{I}(n)} = \bigcap\nolimits_{k \in \mathbb{Z}} R^{\mathcal{I}(k)}.$$

Theorem 5. *Satisfiability of $T_X^{\mathcal{R}}$DL-Lite$_{bool}^{\mathcal{N}}$ KBs is undecidable.*

The proof is by reduction of the $\mathbb{N} \times \mathbb{N}$-tiling problem: given a finite set T of tile types $t = (up(t), down(t), left(t), right(t))$, decide whether T can tile the $\mathbb{N} \times \mathbb{N}$-grid. We assume that the tiles use k colours numbered from 1 to k.

We construct a $T_X^{\mathcal{R}}$DL-Lite$_{bool}^{\mathcal{N}}$ KB $\mathcal{K}_T$ such that $\mathcal{K}_T$ is satisfiable iff T tiles $\mathbb{N} \times \mathbb{N}$. The temporal dimension clearly provides us with one of the two axes of the grid. The other axis is constructed from the domain elements: let R be a role such that $\geq 2 \diamondplus R \sqsubseteq \bot$ and $\geq 2 \diamondplus R^- \sqsubseteq \bot$. In other words, if xRy at some moment of time then there is no $y' \neq y$ with xRy' at any moment of time (and the same for R^-). We can generate an infinite sequence of the domain elements by saying that $\exists R^- \sqcap \bigcirc_F \exists R^-$ is nonempty and $\exists R^- \sqcap \bigcirc_F \exists R^- \sqsubseteq \exists R \sqcap \bigcirc_F \exists R$. (The reason for generating the R-arrows at two consecutive moments of time will become apparent below.) It should be also noted that the produced sequence may in fact be either a finite loop or an infinite sequence of distinct elements.

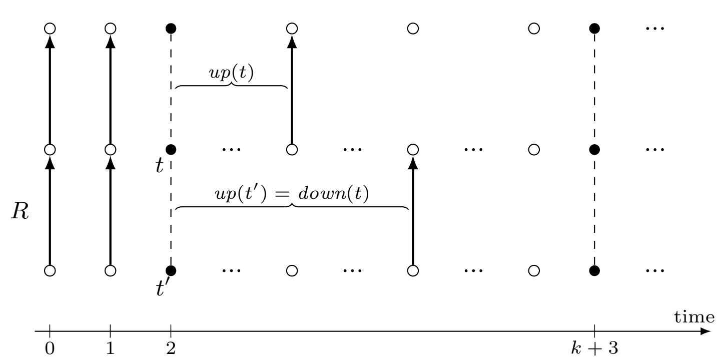
Now, let t be a fresh concept name, for each $t \in T$, and let tile types be disjoint, i.e., $t \sqcap t' \sqsubseteq \bot$ for $t \neq t'$. After the double R-arrows we place the first column of tiles, and every $k+1$ moments afterwards we place a column of tiles that matches the colours of the previous column:

$$\exists R^- \sqcap \bigcirc_F \exists R^- \sqsubseteq \bigsqcup_{t \in T} \bigcirc_F \bigcirc_F t, \qquad t \sqsubseteq \bigsqcup_{right(t)=left(t')} \bigcirc_F^{k+1} t', \quad \text{for each } t \in T.$$

It remains to ensure that the tiles are arranged in a proper grid and have matching top-bottom colours. It is for this purpose that we have (i) used the double R-arrows to generate the sequence of domain elements, and (ii) placed the columns of tiles every $k+1$ moments of time (not every moment). Consider the following CIs, for $t \in T$ and $1 \leq i \leq k$:

$$t \sqsubseteq \neg \exists R^-, \quad t \sqsubseteq \neg \bigcirc_F^i \exists R^- \quad (\text{if } i \neq down(t)) \quad \text{and} \quad t \sqsubseteq \bigcirc_F^{up(t)} \exists R.$$

The first two CIs ensure that between any two tiles $k+1$ moments apart there may be only one incoming R-arrow. This, in particular, means that after the double R-arrows no other two consecutive R-arrows are possible, and thus the proper $\mathbb{N} \times \mathbb{N}$-grid is ensured. Moreover, the exact position of the incoming R-arrow is uniquely determined by the *down*-colour of the tile, which in conjunction with the last CI guarantees that this colour matches the tile below. The following picture illustrates the construction:



Note that the next-time operator $\bigcirc_F$ is heavily used in the encoding above. If we replace it with $\Diamond^*$ and $\Box^*$ on concepts, then reasoning in the resulting logic $T^{\mathcal{R}}_{U}$*DL-Lite*$^{\mathcal{N}}_{bool}$ becomes much simpler:

Theorem 6. *Satisfiability of* $T^{\mathcal{R}}_{U}$*DL-Lite*$^{\mathcal{N}}_{bool}$ *KBs is* NP*-complete.*

This result is proved using a modification of the quasimodel construction from [7, 8]: we show that a KB is satisfiable iff there exists a *quasimodel* of polynomial size. In the *types* of our quasimodels, concepts $\geq q\,R$, $\geq q\,\Diamond^* R$ and $\geq q\,\Box^* R$ reflect the number of R-successors of the element required, respectively, in the current moment of time, 'sometime' ($\Diamond^* R$-successors) and 'always' ($\Box^* R$-successors). In order to deal with temporalised roles, we have to introduce the following conditions on quasimodels: (*i*) the numbers of $\Diamond^* R$-successors and $\Box^* R$-successors in types do not change along a *run* (in other words, temporalised roles are rigid roles); (*ii*) the number of R-successors in every type is sandwiched between the number of $\Box^* R$- and the number of $\Diamond^* R$-successors; (*iii*) if there is a run with more $\Diamond^* R$-successors than $\Box^* R$-successors, then there is a run with more $\Diamond^* R^-$-successors than $\Box^* R^-$-successors; (*iv*) in each run with more $\Diamond^* R$-successors than $\Box^* R$-successors, not all R-successors are $\Box^* R$-successors, and not all $\Diamond^* R$-successors are R-successors at all moments of time. Special conditions are also required for the runs on the objects in the ABox. Full details can be found at `http://www.dcs.bbk.ac.uk/~roman/papers/dl10-full.pdf`.

5 Conclusion

From the complexity-theoretic point of view, the best candidates for reasoning about TCMs appear to be T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ and T_{FPX}*DL-Lite*$^{\mathcal{N}}_{bool}$: the former is NP-complete and the latter PSpace-complete. Moreover, we believe that the reduction of T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ to *LTL* in the proof of Theorem 3 can be done deterministically, in which case one can use standard *LTL* provers for TCM reasoning. We also believe that T_{FPX}*DL-Lite*$^{\mathcal{N}}_{core}$ extended with temporalised roles can be decidable, which remains one of the most challenging open problems. But it seems to be next to impossible to reason in an effective way about all TCM constraints without any restrictions.

References

1. A. Artale, D. Calvanese, R. Kontchakov, and M. Zakharyaschev. *DL-Lite* in the light of first-order logic. In *Proc. of AAAI*, 2007.
2. A. Artale, E. Franconi, F. Wolter, and M. Zakharyaschev. A temporal description logic for reasoning about conceptual schemas and queries. In *Proc. of JELIA*, 2002.
3. A. Artale, C. Parent, and S. Spaccapietra. Evolving objects in temporal information systems. *Annals of Mathematics and Artificial Intelligence*, 50:5–38, 2007.
4. A. Artale, D. Calvanese, R. Kontchakov, V. Ryzhikov, and M. Zakharyaschev. Reasoning over extended ER models. In *Proc. of ER*, 2007.
5. A. Artale, D. Calvanese, R. Kontchakov, and M. Zakharyaschev. The *DL-Lite* family and relations. *J. of Artificial Intelligence Research (JAIR)*, 36:1–69, 2009.

6. A. Artale and E. Franconi. Foundations of temporal conceptual data models. In *Conceptual Modeling: Foundations and Applications*, vol. 5600 of *LNCS*. 2009.
7. A. Artale, R. Kontchakov, C. Lutz, F. Wolter, and M. Zakharyaschev. Temporalising tractable description logics. In *Proc. of TIME*, 2007.
8. A. Artale, R. Kontchakov, V. Ryzhikov, and M. Zakharyaschev. *DL-Lite* with temporalised concepts, rigid axioms and roles. In *Proc. of FroCoS*, 2009.
9. A. Artale, C. Lutz, and D. Toman. A description logic of change. In *Proc. of IJCAI*, 2007.
10. F. Baader, S. Ghilardi, and C. Lutz. LTL over description logic axioms. In *Proc. of KR*, 2008.
11. D. Berardi, D. Calvanese, and G. De Giacomo. Reasoning on UML class diagrams. *Artificial Intelligence*, 168:70–118, 2005.
12. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. of Automated Reasoning*, 39:385–429, 2007.
13. J. Chomicki and D. Toman. Temporal relational calculus. In *Encyclopedia of Database Systems*, pages 3015–3016. Springer, 2009.
14. M. Chrobak. Finite automata and unary languages. *Theor. Comput. Sci.*, 47:149–158, 1986.
15. C. Date, H. Darwen, and N. Lorentzos. *Temporal Data and the Relational Model*. Morgan Kaufmann, 2002.
16. M. Finger and P. McBrien. Temporal conceptual-level databases. In *Temporal Logics—Mathematical Foundations and Computational Aspects*. Oxford University Press, 2000.
17. D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-Dimensional Modal Logics: Theory and Applications*. Elsevier, 2003.
18. H. Gregersen and J. Jensen. Temporal Entity-Relationship models—a survey. *IEEE TKDE*, 11, 1999.
19. G. Hall and R. Gupta. Modeling transition. In *Proc. of ICDE*, 1991.
20. C. Jensen and R. Snodgrass. Temporally enhanced database design. In *Advances in Object-Oriented Data Modeling*. MIT Press, 2000.
21. C. Lutz, F. Wolter, and M. Zakharyaschev. Temporal description logics: A survey. In *Proc. of TIME*, 2008.
22. P. McBrien, A. Seltveit, and B. Wangler. An Entity-Relationship model extended to describe historical information. In *Proc. of CISMOD*, 1992.
23. A. Mendelzon, T. Milo, and E. Waller. Object migration. In *Proc. of PODS*, 1994.
24. K. Schild. Combining terminological logics with tense logic. *Proc. of EPIA*, 1993.
25. S. Spaccapietra, C. Parent, and E. Zimanyi. *Conceptual Modeling for Traditional and Spatio-Temporal Applications—The MADS Approach*. Springer, 2006.
26. J. Su. Dynamic constraints and object migration. *Theoretical Computer Science*, 184:195–236, 1997.
27. B. Tauzovich. Towards temporal extensions to the entity-relationship model. In *Proc. of ER*. Springer, 1991.
28. C. Theodoulidis, P. Loucopoulos, and B. Wangler. A conceptual modelling formalism for temporal database applications. *Information Systems*, 16:401–416, 1991.
29. A. To. Unary finite automata vs. arithmetic progressions. *Inf. Process. Lett.*, 109:1010–1014, 2009.