Identifying Objects Over Time with Description Logics

David Toman[†,‡], and Grant Weddell[†]

[†]D. R. Cheriton School of Computer Science, University of Waterloo, Canada

[‡]Faculty of Computer Science, Free University of Bozen-Bolzano, Italy

Abstract. A fundamental requirement for cooperating agents is to agree on a selection of component values of objects that can be used for reliably communicating references to the objects, to function as their keys. In a distributed environment, such as the web, it is more likely that a choice of such values may have time limits on the duration of their ability to serve as keys, e.g., values denoting permissions, authorizations, service codes, mobile addresses and so on. In this paper, we consider how a Boolean complete description logic can be embellished with a concept constructor for dynamic or temporal forms of equality generating constraints we call *temporal path functional dependencies*. In particular, we introduce the description logic $\mathcal{DLFD}_{\text{temp}}$, demonstrate how it can be used to capture and reason about temporal keys and functional dependencies for a hypothetical distributed hospital database, and prove that the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ is EXPTIME-complete. The latter is accomplished by exhibiting a reduction of the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ to the simpler dialect $\mathcal{DLF}$.

1 Introduction

Consider a situation where two agents a_1 and a_2 operating on behalf of two hospitals must exchange information about staff and departments over the web. Effective communication between a_1 and a_2 requires that they have a common understanding of this information in the form of a shared ontology. The current best practices for expressing this ontology, measured in terms of established reasoning technology, are the *description logic* (DL) based fragments of the OWL web ontology language, called OWL Lite and OWL DL [22]. Building on RDF Schema [14], they enable a_1 and a_2 to share, for example, a mutual understanding that:

– Each staff member has a staff number, a name, a phone number, an associated department and a chief who is also staff; and
– Each department has a name, a hospital name and a staff member who serves as the head.

Although OWL Lite and OWL DL are able to capture such knowledge, neither is sufficiently expressive to capture additional knowledge that would enable a_1 and

a_2 to reliably identify staff and departments *over time*. Such knowledge could include, for example, that:

- Any pair of staff members at either hospital at any time will not share the same (combination of) staff number and hospital name of their department;
- In any given year for a staff member, his or her staff number, telephone number and department are not changed;
- Neither department names nor hospital names ever change, and no two departments in the same hospital share the same names;
- A phone number cannot be assigned to two distinct employees during the first nine months of a year, during the last nine months of a year or during a workyear (i.e., any month excluding July and August); and
- phone numbers that are no longer in use can be reassigned to other staff, but only after a waiting period of ninety days.

The description logic $\mathcal{DLFD}_{\text{temp}}$ introduced in this paper manifests a first attempt to help remedy this situation. $\mathcal{DLFD}_{\text{temp}}$ is an extension of the description logic $\mathcal{DLFD}$, an earlier dialect that incorporated a concept constructor for capturing various forms of static keys and functional dependencies. Concepts using this constructor were called *path functional dependencies* (PFDs).

Like its predecessor, $\mathcal{DLFD}_{\text{temp}}$ is based on *attributes* (also called *features*) instead of the more common case of *roles*. $\mathcal{DLFD}_{\text{temp}}$ extends $\mathcal{DLFD}$ by augmenting PFDs with a temporal component in the same way that Wijsen's Temporal FDs generalize functional dependencies [24]. (The above scenario is an elaboration of sample cases introduced in [24] and used as running examples in the remainder of the paper.) This new more general form of PFDs are called *temporal path functional dependencies*(TPFDs). With $\mathcal{DLFD}_{\text{temp}}$, it becomes possible for agents a_1 and a_2 to now have an additional shared ontology based on TPFDs that captures all of the above.

Finally, we prove that the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ is EXPTIME-complete. This is accomplished by exhibiting a reduction of the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ to the simpler dialect $\mathcal{DLF}$. By using existing decision procedures for $\mathcal{DLF}$, it now becomes possible for agents a_1 and a_2 to know, for example, that a staff member can be

> *reliably identified within any given year by communicating either the combination of values for his or her staff numbers and department hospital names or by his or her telephone number.* (1)

The remainder of the paper is organized as follows. A review of related work completes our introductory comments. In Section 2, we define $\mathcal{DLFD}_{\text{temp}}$ and illustrate its use for the above hospital ontology. Our reduction to $\mathcal{DLF}$ is then presented in Section 3, followed by our summary comments and suggestions for future research in Section 4.

1.1 Related Work

In addition to OWL DL, description logics have been used extensively as a formal way of understanding a large variety of languages for specifying meta-data, including ER diagrams, UML class and object diagrams, relational database schema, and so on [15].

TPFDs introduced in this paper are a generalization of PFDs first introduced in [20]. Less expressive first order PFDs were introduced and studied in the context of object-oriented data models [9, 23]. An FD concept constructor was proposed and incorporated in Classic [4], an early DL with a PTIME reasoning procedure, without changing the complexity of its implication problem. The generalization of this constructor to PFDs alone leads to EXPTIME completeness of the implication problem [11]; this complexity remains unchanged in the presence of additional concept constructors common in rich DLs such as roles, qualified number restrictions, and so on [20, 21].

Recall from the above that TPFDs are also a generalization of *temporal functional dependencies* (TFDs) in [24], which also serves as a source for our example scenarios. TFDs are based on the same underlying data model in [9], and share the same origins in functional dependencies for the relational model.

In [6], the authors consider a DL with functional dependencies and a general form of keys added as additional varieties of dependencies, called a *key box*. They show that their dialect is undecidable for DLs with inverse roles, but becomes decidable when unary functional dependencies are disallowed. This line of investigation is continued in the context of PFDs and inverse features, with analogous results [18], and for this reason, inverse features are not included in $\mathcal{DLFD}_{\text{temp}}$ in order to avoid an already known cause for undecidability.

PFDs have also been used in a number of applications in object-oriented schema diagnosis and synthesis [2, 3], in query optimization [7, 10] and in the selection of indexing for a database [16].

A form of key dependency with left hand side feature paths has been considered for a DL coupled with various concrete domains [12, 13]. In this case, the authors explore how the complexity of satisfaction is influenced by the selection of a particular concrete domain together with various syntactic restrictions on the key dependencies themselves. Note that this earlier work strictly separates objects that serve as "domain values" from abstract objects such as tuples.

Temporal extensions of description logics, in particular extensions based on combining description logics with existing temporal or modal logics has been studied extensively (for a survey see [1], for more details on combining modal logics see [8]). However, *identification constraints*, such as functional dependencies have not been explored in this context beyond unary keys induced by number restrictions.

2 Definitions

$\mathcal{DLFD}_{\text{temp}}$ extends the atemporal logic $\mathcal{DLFD}$ with the ability of identifying objects over time. This is achieved by extending the $\mathcal{DLFD}$'s PFD constructor to allow expressing dependencies between pairs of objects *at different time points*. The extension is based on the notion of *time relations* to describe pertinent relations between time instants (such as a *year*) and by relativizing the

SYNTAX	SEMANTICS
	DEFN OF $(\cdot)^{\mathcal{I}(t)}$
$C ::= A$	(*an arbitrary subset of* Δ_t)
$\mid\ C_1 \sqcap C_2$	$(C_1)^{\mathcal{I}(t)} \cap (C_2)^{\mathcal{I}(t)}$
$\mid\ \neg C$	$\Delta_t \setminus (C)^{\mathcal{I}(t)}$
$\mid\ \forall f.C$	$\{x : (f)^{\mathcal{I}(t)}(x) \in (C)^{\mathcal{I}(t)}\}$
$D ::= C$	
$\mid\ D_1 \sqcap D_2$	$(D_1)^{\mathcal{I}(t)} \cap (D_2)^{\mathcal{I}(t)}$
$\mid\ C : \mathsf{Pf}_1, ..., \mathsf{Pf}_k \rightarrow_T \mathsf{Pf}$	$\{x : \forall (t, t') \in (T)^{\mathcal{I}}, \forall y \in (C)^{\mathcal{I}(t')}. \left(\bigwedge_{i=1}^{k} (\mathsf{Pf}_i)^{\mathcal{I}(t)}(x) = (\mathsf{Pf}_i)^{\mathcal{I}(t')}(y)\right) \Rightarrow (\mathsf{Pf})^{\mathcal{I}(t)}(x) = (\mathsf{Pf})^{\mathcal{I}(t')}(y)\}$
	DEFN OF $(\cdot)^{\mathcal{I}}$
$T ::= \mathsf{curr}$	$\{(t, t) : t \in \mathcal{W}\}$
$\mid\ \mathsf{forever}$	$\mathcal{W} \times \mathcal{W}$
$\mid\ W$	(*an arbitrary subset of* $(\mathsf{forever})^{\mathcal{I}}$ *containing* $(\mathsf{curr})^{\mathcal{I}}$)
$\mid\ T^-$	$\{(t_2, t_1) : (t_1, t_2) \in (T)^{\mathcal{I}}\}$
$\mid\ T_1 \sqcap T_2$	$(T_1)^{\mathcal{I}} \cap (T_2)^{\mathcal{I}}$
$\mid\ T_1 \sqcup T_2$	$(T_1)^{\mathcal{I}} \cup (T_2)^{\mathcal{I}}$

Fig. 1. SYNTAX AND SEMANTICS OF $\mathcal{DLFD}_{\text{temp}}$.

interpretation of the PFD (and the rest of $\mathcal{DLFD}$ as well) with respect to such relations.

A formal definition of $\mathcal{DLFD}_{\text{temp}}$ is given below. Regarding expressiveness, the logic $\mathcal{DLFD}$ is already able to simulate $\mathcal{ALCQI}$ [17] which can in turn simulate $\mathcal{DLR}$ [5]. With temporal PFDs, $\mathcal{DLFD}_{\text{temp}}$ gains an ability, among other things, to assert periods of time during which attributes remain unchanged and during which communicating agents can reliably identify objects in terms of the values for one or more of their attributes.

Definition 1 (Description Logic $\mathcal{DLFD}_{\text{temp}}$) *Let* F*,* A*, and* W *be disjoint sets of attribute names, concept names and time relation names, respectively. A* path expression *is defined by the grammar "*$\mathsf{Pf} ::= f.\,\mathsf{Pf} \mid Id$*" for* $f \in \mathsf{F}$*. We define derived* concept descriptions*,* C *and* D*, and derived* time relation descriptions*,* T*, by the grammar on the left-hand-side of Figure 1. A concept description obtained by using the sixth production is called a* temporal path functional dependency *(TPFD).*

An inclusion dependency $\mathcal{C}$ *is an expression of the form* $C \sqsubseteq D$ *and a* time relation axiom $\mathcal{R}$ *is an expression of the form* $T_1 \sqsubseteq T_2$*. A* terminology (TBox) $\mathcal{T}$ *consists of a finite set of inclusion dependencies and time relation axioms.*

$$
\begin{aligned}
\text{PERSON} &\sqsubseteq \forall Name.\text{STRING}\\
\text{STAFF} &\sqsubseteq \text{PERSON}\\
&\sqcap \forall Snum.\text{INTEGER}\\
&\sqcap \forall PhoneNum.\text{INTEGER}\\
&\sqcap \forall Dept.\text{DEPARTMENT}\\
&\sqcap \forall Chief.\text{STAFF}\\
\text{DEPARTMENT} &\sqsubseteq \neg\text{PERSON}\\
&\sqcap \forall Name.\text{STRING}\\
&\sqcap \forall Hospital.\text{STRING}\\
&\sqcap \forall Head.\text{STAFF}
\end{aligned}
$$

Fig. 2. Static structure for the HOSPITAL ontology in $\mathcal{DLFD}_{\text{temp}}$

The semantics *of expressions is defined with respect to a temporal structure*

$$\mathcal{I} = \left\langle \langle \Delta_t, (\cdot)^{\mathcal{I}(t)} \rangle \mid t \in \mathcal{W} \right\rangle,$$

where $\mathcal{W}$ denotes a non-empty domain of time points or chronons, and $\langle \Delta_t, (\cdot)^{\mathcal{I}(t)} \rangle$ a standard (atemporal) DL interpretation that, for each $t \in \mathcal{W}$, fixes the interpretation of attribute names f to be total functions $(f)^{\mathcal{I}(t)} : \Delta_t \to \Delta_t$. The interpretation is extended to path expressions, $(Id)^{\mathcal{I}(t)} = \lambda x.x$, $(f.\,\mathsf{Pf})^{\mathcal{I}(t)} = (\mathsf{Pf})^{\mathcal{I}(t)} \circ (f)^{\mathcal{I}(t)}$, and to concept descriptions, C and D, and time relation descriptions, T, as defined on the right-hand-side of Figure 1.

The equality symbol is interpreted as the diagonal relation on the set $\bigcup_{t\in\mathcal{W}} \Delta_t$.

An interpretation $\mathcal{I}$ satisfies *an inclusion dependency $C \sqsubseteq D$ if $(C)^{\mathcal{I}(t)} \subseteq (D)^{\mathcal{I}(t)}$ for every $t \in \mathcal{W}$. $\mathcal{I}$* satisfies *a time relation axiom $T_1 \sqsubseteq T_2$ if $(T_1)^{\mathcal{I}} \subseteq (T_2)^{\mathcal{I}}$.*

The $\mathcal{DLFD}_{\text{temp}}$ logical implication problem *asks if $\mathcal{T} \models \mathcal{C}$ holds; that is, for a* posed question *$\mathcal{C}$, if $\mathcal{C}$ is satisfied by any interpretation that satisfies all inclusion dependencies and time relation axioms in $\mathcal{T}$.*

To improve readability in the following, path expressions are written without trailing "*Id*"s when they consist of at least one attribute name. Also, in keeping with Wijsen's Temporal FDs [24], observe that our semantics allows the possibility that the underlying domains at different time points may not coincide. This is in contrast to the so-called *constant domain assumption* commonly utilized by temporal description logics.

Our introductory ontology can be captured as a HOSPITAL terminology in $\mathcal{DLFD}_{\text{temp}}$ as illustrated in Figure 2 for static aspects of information structure, and in Figure 3 for temporal aspects relating to keys and functional dependencies. Note in the latter case the inclusion of various time relation axioms asserting, e.g., that the time relation Year must be symmetric. Agents a_1 and a_2 are

$$
\begin{aligned}
\text{STAFF} \sqsubseteq\ & \text{STAFF} : \mathit{Snum}, \mathit{Dept.Hospital} \rightarrow_{\mathsf{forever}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{Id} \rightarrow_{\text{Year}} \mathit{Snum} \\
\sqcap\ & \text{STAFF} : \mathit{Id} \rightarrow_{\text{Year}} \mathit{Dept} \\
\sqcap\ & \text{STAFF} : \mathit{Id} \rightarrow_{\text{Year}} \mathit{PhoneNum}
\end{aligned}
$$

$$
\begin{aligned}
\text{DEPARTMENT} \sqsubseteq\ & \text{DEPARTMENT} : \mathit{Id} \rightarrow_{\mathsf{forever}} \mathit{Name} \\
\sqcap\ & \text{DEPARTMENT} : \mathit{Id} \rightarrow_{\mathsf{forever}} \mathit{Hospital} \\
\sqcap\ & \text{DEPARTMENT} : \mathit{Name}, \mathit{Hospital} \rightarrow_{\mathsf{forever}} \mathit{Id}
\end{aligned}
$$

$$
\begin{aligned}
\text{STAFF} \sqsubseteq\ & \text{STAFF} : \mathit{PhoneNum} \rightarrow_{\text{FirstNineMonths}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{PhoneNum} \rightarrow_{\text{LastNineMonths}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{PhoneNum} \rightarrow_{\text{Workyear}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{PhoneNum} \rightarrow_{\text{NinetyDays}} \mathit{Id}
\end{aligned}
$$

$$
\begin{aligned}
\text{Year} &\sqsubseteq (\text{FirstNineMonths} \sqcup \text{LastNineMonths} \sqcup \text{Workyear}) \sqcap \text{Year}^{-} \\
\text{FirstNineMonths} &\sqsubseteq \text{Year} \sqcap \text{FirstNineMonths}^{-} \\
\text{LastNineMonths} &\sqsubseteq \text{Year} \sqcap \text{LastNineMonths}^{-} \\
\text{Workyear} &\sqsubseteq \text{Year} \sqcap \text{Workyear}^{-}
\end{aligned}
$$

Fig. 3. Dynamic structure for the HOSPITAL ontology in $\mathcal{DLFD}_{\text{temp}}$

now able to formally express (1) above in terms of the inclusion dependency

$$
\begin{aligned}
\text{STAFF} \sqsubseteq\ & \text{STAFF} : \mathit{Snum}, \mathit{Dept.Hospital} \rightarrow_{\text{Year}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{Id} \rightarrow_{\text{Year}} \mathit{Dept.Hospital} \\
\sqcap\ & \text{STAFF} : \mathit{PhoneNum} \rightarrow_{\text{Year}} \mathit{Id} \\
\sqcap\ & \text{STAFF} : \mathit{Id} \rightarrow_{\text{Year}} \mathit{PhoneNum}.
\end{aligned}
\tag{2}
$$

Our decision procedure for the $\mathcal{DLFD}_{\text{temp}}$ implication problem can then verify that (2) is a logical consequence of the HOSPITAL terminology. This has the crucial consequence that a_1 and a_2 can know that they are able to unambiguously communicate a reference to a staff person within a calendar year by exchanging, e.g., a combination of his or her current staff number and current name of the hospital of their current department.

3 Decision Procedure

We now prove that the membership problem for $\mathcal{DLFD}_{\text{temp}}$ is complete for EXPTIME by exhibiting a reduction of the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ to the simpler dialect $\mathcal{DLF}$. The result then follows by appeal to existing decision procedures and complexity bounds for $\mathcal{DLF}$.

First, let $\mathsf{TR}(\mathcal{T})$ be the set of all time descriptions that appear in $\mathcal{T}$. We associate two auxiliary primitive concepts $\mathrm{C}_{\mathrm{T}}^{\mathrm{LR}}$ and $\mathrm{C}_{\mathrm{T}}^{\mathrm{RL}}$ with each time description $\mathrm{T} \in \mathsf{TR}(\mathcal{T})$. To mimic the behaviour of time relations we constrain the behaviour of these concepts a follows:

$$\mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_1\sqcap \mathrm{T}_2} = \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_1} \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_2} \qquad \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}^-} = \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}} \qquad \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_1\sqcap \mathrm{T}_2} = \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_1} \sqcap \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_2}$$
$$\mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_1\sqcup \mathrm{T}_2} = \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_1} \sqcup \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_2} \qquad \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}} = \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}^-} \qquad \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_1\sqcup \mathrm{T}_2} = \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_1} \sqcup \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_2}$$

where $T, T_1, T_2 \in \mathsf{TR}(\mathcal{T})$ and equality denotes subsumptions in both directions.

Now given a $\mathcal{DLFD}_{\text{temp}}$ terminology $\mathcal{T}$, we define a $\mathcal{DLF}$ terminology $\mathcal{T}^*$ as a set consisting of the above axioms together with the following:

1. $\mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_1} \sqsubseteq \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}_2}$ and $\mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_1} \sqsubseteq \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}_2}$ for each time relation axiom $T_1 \sqsubseteq T_2 \in \mathcal{T} \cup \{\mathsf{curr} \sqsubseteq \mathsf{curr}^-, \mathsf{forever} \sqsubseteq \mathsf{forever}^-, \mathsf{curr} \sqsubseteq \mathrm{T}, \mathrm{T} \sqsubseteq \mathsf{forever} \mid \mathrm{T} \in \mathsf{TR}(\mathcal{T})\}$,
2. $\mathrm{C}^{\mathrm{LR}}_{\mathrm{T}} \sqsubseteq \forall f.\mathrm{C}^{\mathrm{LR}}_{\mathrm{T}}$ and $\mathrm{C}^{\mathrm{RL}}_{\mathrm{T}} \sqsubseteq \forall f.\mathrm{C}^{\mathrm{RL}}_{\mathrm{T}}$ for all $\mathrm{T} \in \mathsf{TR}(\mathcal{T})$ and features f in $\mathcal{T}$,
3. $\mathrm{C}^L \sqsubseteq \mathrm{D}^L$ and $\mathrm{C}^R \sqsubseteq \mathrm{D}^R$ for each subsumption axiom $\mathrm{C} \sqsubseteq \mathrm{D} \in \mathcal{T}$ that is not a time relation axiom and such that D is not TPFD,
4. $\mathrm{C}^L \sqcap \mathrm{D}^R \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}} \sqcap (\sqcap_{i\leq k} \forall\, \mathsf{Pf}_i\,.\mathsf{Eq}) \sqsubseteq \forall\, \mathsf{Pf}\,.\mathsf{Eq}$ and $\mathrm{C}^R \sqcap \mathrm{D}^L \sqcap \mathrm{C}^{\mathrm{RL}}_{\mathrm{T}} \sqcap (\sqcap_{i\leq k} \forall\, \mathsf{Pf}_i\,.\mathsf{Eq}) \sqsubseteq \forall\, \mathsf{Pf}\,.\mathsf{Eq}$ for each subsumption axiom $\mathrm{C} \sqsubseteq \mathrm{D} : \mathsf{Pf}_1, \ldots, \mathsf{Pf}_k \rightarrow_T \mathsf{Pf} \in \mathcal{T}$,
5. $(\mathsf{Eq} \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathsf{curr}}) \sqsubseteq \forall f.\mathsf{Eq}$ for each primitive feature f in $\mathcal{T}$, and
6. $(\mathsf{Eq} \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathsf{curr}} \sqcap \mathrm{A}^L) \sqsubseteq \mathrm{A}^R$ and $(\mathsf{Eq} \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathsf{curr}} \sqcap \mathrm{A}^R) \sqsubseteq \mathrm{A}^L$ for each primitive concept A in $\mathcal{T}$.

where D^L (resp. D^R) denotes a $\mathcal{DLF}$ concept description D in which all occurrences of primitive concept description A has been replaced by A^L (resp. A^R). For a given posed question $\mathcal{Q} = \mathrm{C} \sqsubseteq \mathrm{D}$, we need to distinguish two cases:

1. $\mathcal{Q}$ is not a TPFD: define $\mathcal{Q}^*$ to be $\mathrm{C}^L \sqsubseteq \mathrm{D}^L$; and
2. $\mathcal{Q}$ is of the form $\mathrm{C} \sqsubseteq \mathrm{D} : \mathsf{Pf}_1, \ldots, \mathsf{Pf}_k \rightarrow_T \mathsf{Pf}$: define $\mathcal{Q}^*$ to be

$$\mathrm{C}^L \sqcap \mathrm{D}^R \sqcap \mathrm{C}^{\mathrm{LR}}_{\mathrm{T}} \sqcap (\sqcap_{i\leq k} \forall\, \mathsf{Pf}_i\,.\mathsf{Eq}) \sqsubseteq \forall\, \mathsf{Pf}\,.\mathsf{Eq}.$$

Theorem 2 *Let $\mathcal{T} \models \mathcal{Q}$ be a $\mathcal{DLFD}_{\text{temp}}$ implication problem. Then*

$$\mathcal{T} \models \mathcal{Q} \;\; \textit{if and only if} \;\; \mathcal{T}^* \models \mathcal{Q}^*.$$

<u>Proof:</u> (outline) Let $\mathcal{Q} = \mathrm{C} \sqsubseteq \mathrm{D}$. The case where D is not a TPFD is straightforward since we can test for logical implication in a single world using a tree model. Hence, this case reduces immediately to reasoning in $\mathcal{DLF}$ [17, 20].

We now consider both implications in the alternative case where $\mathcal{Q}$ has the form

$$\mathrm{C} \sqsubseteq \mathrm{D} : \mathsf{Pf}_1, \ldots, \mathsf{Pf}_k \rightarrow_T \mathsf{Pf}.$$

Possible conjunctions that are allowed in D reduce to one of the above cases simply by considering each conjunct separately.

($\Rightarrow$) Assume that $\mathcal{T}^* \not\models \mathcal{Q}^*$. Then there must exist a tree model $\mathcal{I}^*$ of $\mathcal{T}^*$ with a root, o, satisfying the concept

$$\mathrm{C}^L \sqcap \mathrm{D}^R \sqcap \mathrm{C}_\mathrm{T}^{\mathrm{LR}} \sqcap (\sqcap_{i\leq k} \forall\, \mathsf{Pf}_i\, .\mathsf{Eq}) \sqcap \neg\forall\, \mathsf{Pf}\, .\mathsf{Eq}.$$

We construct a model $\mathcal{I}$ for $\mathcal{T}$ that falsifies $\mathcal{Q}$ as follows: let o_1 and o_2 be objects in the domain of $\mathcal{I}$ and t_1 and t_2 time instants.

We distinguish two cases based on T. For $T = \mathsf{curr}$ we have $t_1 = t_2$. It therefore follows that $o_1 \neq o_2$. The construction then proceeds as in the atemporal case [17] with the final interpretation consisting of a single world, $\mathcal{W} = \{t_1\}$.

For $t_1 \neq t_2$ (but $(t_1, t_2) \in (T)^\mathcal{I}$) we define an interpretation $\mathcal{I}$ as follows: the interpretation consists of two worlds, $\mathcal{W} = \{t_1, t_2\}$, with each containing an interpretation in which objects are terms of the form $\mathsf{Pf}(o_1)$ and $\mathsf{Pf}(o_2)$, where Pf is a path expression. The temporal interpretation is then defined as

$$\mathcal{I} = \left\langle \langle \{\mathsf{Pf}(o_0)\}, (.)^{\mathcal{I}(t_0)}\rangle, \langle \{\mathsf{Pf}(o_1)\}, (.)^{\mathcal{I}(t_1)}\rangle \right\rangle.$$

We define the following relation on $\Delta_{t_1} \cup \Delta_{t_2}$

$$\begin{array}{l}\{(\mathsf{Pf}(o_1), \mathsf{Pf}(o_1)), (\mathsf{Pf}(o_2), \mathsf{Pf}(o_2)) \mid \mathsf{Pf} \text{ a path description}\}\cup\\ \{(\mathsf{Pf}(o_1), \mathsf{Pf}(o_2)), (\mathsf{Pf}(o_2), \mathsf{Pf}(o_1)) \mid (\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathsf{Eq})^{\mathcal{I}^*}\},\end{array}$$

to identify objects that are equal in the two worlds (note that technically we chose a representative for each equivalence class of the above relation for the equality to be truly a diagonal relation on $\Delta_{t_1} \cup \Delta_{t_2}$), and the interpretation functions $(.)^{\mathcal{I}(t_i)}$ of primitive concepts as follows:

- $(\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1) \in (\mathrm{A})^{\mathcal{I}(t_1)}$ if $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{A}^L)^{\mathcal{I}^*}$,
- $(\mathsf{Pf})^{\mathcal{I}(t_2)}(o_2) \in (\mathrm{A})^{\mathcal{I}(t_2)}$ if $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{A}^R)^{\mathcal{I}^*}$;

the interpretations are then extended to complex concepts in a standard way. Finally, the interpretation of attributes is defined as $(f)^{\mathcal{I}(t_i)}(x) = f.x$ (since objects are represented by terms).

By case analysis, it is straightforward to verify that $\mathcal{I} \models \mathcal{T}$. However, $\mathcal{I} \not\models \mathcal{Q}$, a fact witnessed by the two objects o_1 and o_2 since $o_1 \in (\mathrm{C})^{\mathcal{I}(t_1)}$, $o_2 \in (\mathrm{D})^{\mathcal{I}(t_2)}$, $(\mathsf{Pf}_i)^{\mathcal{I}(t_1)}(o_1) = \mathsf{Pf}_i(o_1) = \mathsf{Pf}_i(o_2) = (\mathsf{Pf}_i)^{\mathcal{I}(t_1)}(o_1)$ as $(\mathsf{Pf}_i)^{\mathcal{I}^*}(o) \in (\mathsf{Eq})^{\mathcal{I}^*}$ (or equivalently, $o \in (\forall\, \mathsf{Pf}_i\, .\mathsf{Eq})^{\mathcal{I}^*}$) for all $i \leq k$, but $(\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1) = \mathsf{Pf}(o_1) \neq \mathsf{Pf}(o_2) = (\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1)$ as $(\mathsf{Pf})^{\mathcal{I}^*}(o) \notin (\mathsf{Eq})^{\mathcal{I}^*}$ (or equivalently, $o \in \neg(\forall\, \mathsf{Pf}\, .\mathsf{Eq})^{\mathcal{I}^*}$).

($\Leftarrow$) Now assume $\mathcal{T} \not\models \mathcal{Q}$. Then there is an interpretation $\mathcal{I}$ that is a model for $\mathcal{T}$ but falsifies $\mathcal{Q}$. Hence, there must be $o_1 \in (C)^{\mathcal{I}(t_1)}$ and $o_2 \in (D)^{\mathcal{I}(t_2)}$ such that $(t_1, t_2) \in (T)^\mathcal{I}$, $(\mathsf{Pf}_i)^{\mathcal{I}(t_1)}(o_1) = (\mathsf{Pf}_i)^{\mathcal{I}(t_2)}(o_2)$, and $(\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1) \neq (\mathsf{Pf})^{\mathcal{I}(t_2)}(o_2)$. Note that we can allow interpretations in which $o_1 = o_2$ or $t_1 = t_2$ (but not both) as long as the above conditions are met.

Now define a $\mathcal{DLF}$ interpretation $\mathcal{I}^*$: let o be an arbitrary object in the domain of $\mathcal{I}^*$, and assign the interpretation of primitive concepts as follows:

1. If $(\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1) \in (\mathrm{A})^{\mathcal{I}(t_1)}$ then $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{A}^L)^{\mathcal{I}^*}$;
2. If $(\mathsf{Pf})^{\mathcal{I}(t_2)}(o_2) \in (\mathrm{A})^{\mathcal{I}(t_2)}$ then $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{A}^R)^{\mathcal{I}^*}$;
3. If $(\mathsf{Pf})^{\mathcal{I}(t_1)}(o_1) = (\mathsf{Pf})^{\mathcal{I}(t_2)}(o_2)$ then $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathsf{Eq})^{\mathcal{I}^*}$;
4. If $(t_1, t_2) \in (T)^{\mathcal{I}}$ then $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{C}_{\mathrm{T}}^{\mathrm{LR}})^{\mathcal{I}^*}$; and
5. If $(t_2, t_1) \in (T)^{\mathcal{I}}$ then $(\mathsf{Pf})^{\mathcal{I}^*}(o) \in (\mathrm{C}_{\mathrm{T}}^{\mathrm{RL}})^{\mathcal{I}^*}$.

for all primitive concepts A, all path functions Pf and all time relations T. It is easy to verify that $\mathcal{I}^*$ is a model of $\mathcal{T}^*$ and that o falsifies

$$\mathrm{C}^L \sqcap \mathrm{D}^R \sqcap \mathrm{C}_{\mathrm{T}}^{\mathrm{LR}} \sqcap (\sqcap_{i \le k} \forall\, \mathsf{Pf}_i\, .\mathsf{Eq}) \sqcap \neg\forall\, \mathsf{Pf}\, .\mathsf{Eq}.$$

□

And since $\mathcal{T}^* \models \mathcal{Q}^*$ is a $\mathcal{DLF}$ implication problem:

Corollary 3 *The logical implication problem for $\mathcal{DLFD}_{\text{temp}}$ is decidable and EXPTIME-complete.*

Proof: Follows immediately from Theorem 2 above and results in [17]. □

4 Summary and Discussion

We have introduced $\mathcal{DLFD}_{\text{temp}}$, a Boolean complete description logic with a concept constructor for expressing dynamic or temporal forms of equality generating constraints called *temporal path functional dependencies* (TPFDs), and have illustrated how TPFDs can be used to capture and reason about temporal keys and functional dependencies for a hypothetical distributed hospital database. We have also proven that the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ is EXPTIME-complete by exhibiting a reduction of the general membership problem for $\mathcal{DLFD}_{\text{temp}}$ to the simpler dialect $\mathcal{DLF}$ for which existing decisions procedures are known [17].

There are several worthwhile directions for future work. Possibilities that we suspect are in increasing order of difficulty are as follows.

- We wish to investigate how time relation descriptions can be generalized in various ways, e.g., to enable specifying time relations that are irreflexive, or transitive, that encode “next time”, are complements of other time relations, and so on.
- We have investigated relaxing the restrictions on the location of the PFD concept constructor in the dialect $\mathcal{DLFD}$, showing that it is possible to allow PFDs to occur in right-hand-sides of inclusion dependencies in the scope of monotonic concept constructors [19]. We wish to investigate a similar possibility with TPFDs.

– Our motivating application for the development of $\mathcal{DLFD}_{\text{temp}}$ is to enable agents to know how to communicate unambiguous references to objects over time. This suggests the likely efficacy of a new reasoning service for DL reasoners: one that responds to *requests* by agents for a combination of component path expressions that can reliably serve as object keys over a given time duration. For example, in our sample application, agent a_1 would supply the parameters HOSPITAL, Year and STAFF as such a request, possibly getting in return $\{Snum, Dept.Hospital\}$.
– Finally, we believe that a reduction of the $\mathcal{DLFD}_{\text{temp}}$ membership problem to the $\mathcal{DLFD}$ membership problem is still possible under the constant domain assumption. However, our initial investigations along this line suggest the the reduction will be considerably more involved.

References

1. Alessandro Artale and Enrico Franconi. Temporal description logics. In Michael Fisher, Dov Gabbay, and Lluis Vila, editors, *Handbook of Temporal Reasoning in Artificial Intelligence*, pages 375–388. Elsevier, 2005.
2. Joachim Biskup and Torsten Polle. Decomposition of Database Classes under Path Functional Dependencies and Onto Constraints. In *Foundations of Information and Knowledge Systems*, pages 31–49, 2000.
3. Joachim Biskup and Torsten Polle. Adding inclusion dependencies to an object-oriented data model with uniqueness constraints. *Acta Informatica*, 39:391–449, 2003.
4. Alexander Borgida and Grant E. Weddell. Adding Uniqueness Constraints to Description Logics (Preliminary Report). In *International Conference on Deductive and Object-Oriented Databases*, pages 85–102, 1997.
5. Diego Calvanese, Giuseppe De Giacomo, and Maurizio Lenzerini. On the Decidability of Query Containment under Constraints. In *ACM Symposium on Principles of Database Systems*, pages 149–158, 1998.
6. Diego Calvanese, Giuseppe De Giacomo, and Maurizio Lenzerini. Identification Constraints and Functional Dependencies in Description Logics. In *Proc. of the 17th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 155–160, 2001.
7. David DeHaan, David Toman, and Grant E. Weddell. Rewriting Aggregate Queries using Description Logics. In *Description Logics 2003*, pages 103–112. CEUR-WS vol.81, 2003.
8. D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-dimensional modal logics: theory and applications*. Studies in Logic. Elsevier, 2003.
9. Minoru Ito and Grant E. Weddell. Implication Problems for Functional Constraints on Databases Supporting Complex Objects. *Journal of Computer and System Sciences*, 49(3):726–768, 1994.
10. Vitaliy L. Khizder, David Toman, and Grant E. Weddell. Reasoning about Duplicate Elimination with Description Logic. In *Rules and Objects in Databases (DOOD, part of CL'00)*, pages 1017–1032, 2000.
11. Vitaliy L. Khizder, David Toman, and Grant E. Weddell. On Decidability and Complexity of Description Logics with Uniqueness Constraints. In *International Conference on Database Theory ICDT'01*, pages 54–67, 2001.

12. C. Lutz and M. Milicic. Description Logics with Concrete Domains and Functional Dependencies. In *Proc. European Conference on Artificial Intelligence (ECAI)*, pages 378–382, 2004.
13. Carsten Lutz, Carlos Areces, Ian Horrocks, and Ulrike Sattler. Keys, Nominals, and Concrete Domains. In *Proc. of the 18th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 349–354, 2003.
14. Resource Description Framework (RDF). http://www.w3.org/RDF/.
15. U. Sattler, D. Calvanese, and R. Molitor. Relationships with other formalisms. In *The Description Logic Handbook: Theory, Implementation, and Applications*, chapter 4, pages 137–177. Cambridge University Press, 2003.
16. Lubomir Stanchev and Grant E. Weddell. Index Selection for Embedded Control Applications using Description Logics. In *Description Logics 2003*, pages 9–18. CEUR-WS vol.81, 2003.
17. David Toman and Grant Weddell. On Reasoning about Structural Equality in XML: A Description Logic Approach. *Theoretical Computer Science*, 336(1):181–203, 2005.
18. David Toman and Grant Weddell. On the Interaction between Inverse Features and Path-functional Dependencies in Description Logics. In *Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 603–608, 2005.
19. David Toman and Grant Weddell. On keys and functional dependencies as first-class citizens in description logics. *Journal of Automated Reasoning*, 2008, to appear.
20. David Toman and Grant E. Weddell. On Attributes, Roles, and Dependencies in Description Logics and the Ackermann Case of the Decision Problem. In *Description Logics 2001*, pages 76–85. CEUR-WS vol.49, 2001.
21. David Toman and Grant E. Weddell. Attribute Inversion in Description Logics with Path Functional Dependencies. In *Description Logics 2004*, pages 178–187. CEUR-WS vol.104, 2004.
22. Web Ontology Language (OWL). http://www.w3.org/2004/OWL/.
23. Grant Weddell. A Theory of Functional Dependencies for Object Oriented Data Models. In *International Conference on Deductive and Object-Oriented Databases*, pages 165–184, 1989.
24. Jef Wijsen. Temporal FDs on Complex Objects. *ACM Transactions on Database Systems*, 24(1):127–176, 1999.