

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, that is, to function as their keys. In distributed environments 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 with a concept constructor for expressing “always” can also 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 temporal description logic $\mathcal{DLFD}_{\text{temp}}$, demonstrate how it can be used, among other things, 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} . We also show that the addition of very simple kinds of eventualities leads to a significant increase in the complexity of the membership problem.

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 (W3C 2004b). They build on RDF Schema (W3C 2004a) and enable a_1 and a_2 to share a common understanding of their hospitals:

- Each staff member is a person with a name who also has

Copyright © 2008, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

an assigned staff number, a phone number, an associated department and a chief who is also staff; and that

- 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, they are not able to capture additional knowledge that would enable a_1 and a_2 to reliably identify staff and departments, in particular *over time*. This kind of temporal knowledge might allow a_1 and a_2 to also know the following:

- Any pair of staff members at either hospital at any time do 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 temporal description logic $\mathcal{DLFD}_{\text{temp}}$ introduced in this paper is the first DL dialect that is able to capture this kind of knowledge. This logic 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 (Toman and Weddell 2008). Concepts using this constructor were called *path functional dependencies* (PFDs).

Similarly to its predecessor, $\mathcal{DLFD}_{\text{temp}}$ is based on *attributes* (also called *features*) instead of the more common *roles*, and extends \mathcal{DLFD} in two ways. The first augments

*A preliminary version of this work was presented at the Description Logics Workshop DL 2007.

PFDs with a *temporal component* in a similar fashion to how Wijsen's Temporal FDs generalize functional dependencies (Wijsen 1999). The above scenario is an elaboration of sample cases introduced by Wijsen and is used as running examples in the remainder of the paper. This new more general form of PFDs are called *temporal path functional dependencies* (TPFDs). Thus, 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.

The second extension to \mathcal{DLFD} adds an additional concept constructor for expressing "always". This enables $\mathcal{DLFD}_{\text{temp}}$ to capture additional temporal knowledge, in particular:

- A person is always a person; and
- Staff remain staff for a working year.

TPFDs and the added "always" concept constructor are based on an underlying notion of a *time relation*. Consequently, $\mathcal{DLFD}_{\text{temp}}$ is also able to capture a (sub)ontology of time relations over pairs of time points, for example that

the first nine months of the same year

is a subset of

same year,

and that both of these time relations are symmetric.

The primary contributions of this paper are as follows.

1. 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} .
2. We prove that this is no longer possible if one extends $\mathcal{DLFD}_{\text{temp}}$ to enable capturing very simple kinds of eventualities. In this case the complexity must increase at least to 2-EXPTIME.

Note that, in contrast to many other temporal description logics, $\mathcal{DLFD}_{\text{temp}}$ is carefully designed to preserve the complexity of the membership problem in the underlying atemporal logic while providing considerable modelling power as witnessed by our running example introduced above. Returning to the example, existing decision procedures for \mathcal{DLF} now make it possible for agents a_1 and a_2 to know (or be told) the following.

A staff member can be reliably identified within any given year by communicating the combination of values for his or her staff numbers and department hospital names, or, alternatively, 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. Section 4 complements this development by showing that extensions of the $\mathcal{DLFD}_{\text{temp}}$ lead to a significant increase in the complexity of the associated decision problem. We conclude with summary comments and suggestions for future research in Section 5.

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 (Sattler, Calvanese, and Molitor 2003).

TPFDs introduced in this paper are a generalization of PFDs first introduced in (Toman and Weddell 2001). Less expressive first order PFDs were introduced and studied in the context of object-oriented data models (Ito and Weddell 1994; Weddell 1989). An FD concept constructor was proposed and incorporated in Classic (Borgida and Weddell 1997), 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 (Khizder, Toman, and Weddell 2001); this complexity remains unchanged in the presence of additional concept constructors common in rich DLs such as roles, qualified number restrictions, and so on (Toman and Weddell 2001; 2004).

Recall from the above that TPFDs are also a generalization of *temporal functional dependencies* (TFDs) in (Wijsen 1999), which also serves as a source for our example scenarios. TFDs are based on the same underlying data model in (Ito and Weddell 1994), and share the same origins in functional dependencies for the relational model.

Calvanese *et al.* (Calvanese, De Giacomo, and Lenzerini 2001) 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 combined with inverse features, with analogous results (Toman and Weddell 2005b), 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 (Biskup and Polle 2000; 2003), in query optimization (DeHaan, Toman, and Weddell 2003; Khizder, Toman, and Weddell 2000) and in the selection of indexing for a database (Stanchev and Weddell 2003).

A form of key dependency with left hand side feature paths has been considered for a DL coupled with various concrete domains (Lutz and Milicic 2004; Lutz *et al.* 2003). In this case, the authors explore how the complexity of satisfaction is influenced by the selection of a particular concrete

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

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

domain together with various syntactic restrictions on the key dependencies themselves. Note that these approaches 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 (Artale and Franconi 2005), for more details on combining modal logics see (Gabbay et al. 2003). However, *identification constraints*, such as functional dependencies have not been explored in this context beyond unary keys induced by number restrictions. Also, in many of these approaches, allowing global roles (which can be easily captured in $\mathcal{DLFD}_{\text{temp}}$) leads to a significant increase in the computational hardness of the associated reasoning problems.

2 Definitions

$\mathcal{DLFD}_{\text{temp}}$ extends the atemporal logic \mathcal{DLFD} with the ability to identify 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 rela-*

tions to describe pertinent relations between time instants (such as a *year*) and by relativizing the interpretation of the PFD (and the rest of \mathcal{DLFD} as well) with respect to such relations. 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. A formal definition of $\mathcal{DLFD}_{\text{temp}}$ is given below.

Definition 1 (Description Logic $\mathcal{DLFD}_{\text{temp}}$) *Let F , A , and \mathcal{W} be disjoint sets of attribute names, concept names and time relation names, respectively. A path expression is defined by the grammar “ $\mathbf{Pf} ::= f.\mathbf{Pf} \mid \text{Id}$ ” for $f \in F$. 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 eighth production is called a temporal path functional dependency (TPFD). An inclusion dependency C is an expression of the form $C \sqsubseteq D$ and a time relation axiom R is an expression of the form $T_1 \sqsubseteq T_2$. A terminology (TBox) T consists of a finite set of inclusion dependencies and time relation axioms.*

PERSON	$\sqsubseteq \forall Name.STRING$	STAFF	$\sqsubseteq PERSON$
DEPARTMENT	$\sqsubseteq \neg PERSON$		$\sqcap \forall Snum.INTEGER$
	$\sqcap \forall Name.STRING$		$\sqcap \forall PhoneNum.INTEGER$
	$\sqcap \forall Hospital.STRING$		$\sqcap \forall Dept.DEPARTMENT$
	$\sqcap \forall Head.STAFF$		$\sqcap \forall Chief.STAFF$

Figure 2: STATIC STRUCTURE FOR THE HOSPITAL ONTOLOGY IN $\mathcal{DLFD}_{\text{temp}}$

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

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

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 $(f)^{\mathcal{I}(t)} : \Delta_t \rightarrow \Delta_t$. The interpretation is extended to path expressions, $(Id)^{\mathcal{I}(t)} = \lambda x.x$, $(f.Pf)^{\mathcal{I}(t)} = (Pf)^{\mathcal{I}(t)} \circ (f)^{\mathcal{I}(t)}$, 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 either $T \models C$ or $T \models R$ holds; that is, for a posed question C or R , if either is satisfied by any interpretation that satisfies all inclusion dependencies and time relation axioms in T .

To improve readability in the following, path expressions are written without trailing “ Id ”’s when they consist of at least one attribute name.

In keeping with Wijsen’s Temporal FDs (Wijsen 1999), 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 many temporal description logics. Also note that the semantics ensures that time relations always contain “curr” and therefore conforms to our intuition that each world should at least see itself. If subsets of “curr” are necessary for some application domain, however, it is an easy exercise to remove this condition.

It is straightforward to formally capture our introductory hospital ontology as a terminology in $\mathcal{DLFD}_{\text{temp}}$. In particular, inclusion dependencies that capture the static aspects of the information structure are given in Figure 2, while inclusion dependencies and time relation axioms that capture temporal aspects of this structure are given in Figure 3. For example, the first line of Figure 3 is a inclusion dependency that is only satisfied by a temporal structure for which no object can be a person at one point in time and not a person

at another, that is, ensures *once a person, always a person*. The second and third lines ensure the following:

1. Any pair of distinct staff at any two (possibly different) points in time will either have a distinct staff number or will be assigned to departments in different hospitals; and
2. The staff number of a staff member will not be changed within any given calendar year.

The last line of Figure 3 is a time relation axiom that distinguishes temporal structures for which a working year time relation is a more special case of a year time relation, and is also a time relation that is reflexive. An interesting consequence of the final four lines of Figure 3 is the time relation axiom

$$\text{Year} \sqsubseteq \text{Year}^{-},$$

asserting that year is also a symmetric time relation.

Agents a_1 and a_2 are now able to formally express (1) above in terms of the inclusion dependency (2) below. Our decision procedure for the $\mathcal{DLFD}_{\text{temp}}$ implication problem can then verify that (2) is indeed a logical consequence of the HOSPITAL terminology.

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

This has the desired 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 in either of two ways:

1. by exchanging current values for a combination of his or her staff number and department’s hospital name, or
2. by exchanging the current value of his or her phone number.

TPFDs vs. Global Features/Roles

It is easy to see that assertions of the form

$$C \sqsubseteq C : Id \rightarrow_{\text{forever}} g$$

essentially state that the feature g emanating from a particular C -object o in arbitrary two different worlds leads necessarily to the same object $g(o)$ in both the worlds; in other words, g is *global feature* (on the class C).

PERSON	\sqsubseteq	$\square_{\text{forever}} \text{PERSON}$
STAFF	\sqsubseteq	$\text{STAFF} : Snum, Dept. Hospital \rightarrow_{\text{forever}} Id$
	\sqcap	$\text{STAFF} : Id \rightarrow_{\text{Year}} Snum$
	\sqcap	$\text{STAFF} : Id \rightarrow_{\text{Year}} Dept$
	\sqcap	$\text{STAFF} : Id \rightarrow_{\text{Year}} PhoneNum$
DEPARTMENT	\sqsubseteq	$\text{DEPARTMENT} : Id \rightarrow_{\text{forever}} Name$
	\sqcap	$\text{DEPARTMENT} : Id \rightarrow_{\text{forever}} Hospital$
	\sqcap	$\text{DEPARTMENT} : Name, Hospital \rightarrow_{\text{forever}} Id$
STAFF	\sqsubseteq	$\text{STAFF} : PhoneNum \rightarrow_{\text{FirstNineMonths}} Id$
	\sqcap	$\text{STAFF} : PhoneNum \rightarrow_{\text{LastNineMonths}} Id$
	\sqcap	$\text{STAFF} : PhoneNum \rightarrow_{\text{Workyear}} Id$
	\sqcap	$\text{STAFF} : PhoneNum \rightarrow_{\text{NinetyDays}} Id$
	\sqcap	$\square_{\text{Workyear}} \text{STAFF}$
Year	\sqsubseteq	$(\text{FirstNineMonths} \sqcup \text{LastNineMonths} \sqcup \text{Workyear})$
FirstNineMonths	\sqsubseteq	$\text{Year} \sqcap \text{FirstNineMonths}^-$
LastNineMonths	\sqsubseteq	$\text{Year} \sqcap \text{LastNineMonths}^-$
Workyear	\sqsubseteq	$\text{Year} \sqcap \text{Workyear}^-$

Figure 3: DYNAMIC STRUCTURE FOR THE HOSPITAL ONTOLOGY IN $\mathcal{DLFD}_{\text{temp}}$

3 Decision Procedure for $\mathcal{DLFD}_{\text{temp}}$

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} .

Reasoning with Time Relations

First, let $\text{TR}(\mathcal{T})$ be the set of all time relation descriptions that appear in \mathcal{T} . We associate two new primitive concepts C_T^{LR} and C_T^{RL} with each $T \in \text{TR}(\mathcal{T})$. Intuitively, these concepts simulate the behaviour of the time relations for two chronons: a “left” one and a “right” one. The actual behaviour of time relations is thus captured by constraining the behaviour of these concepts as follows:

$$\begin{array}{ll} C_{T_1 \sqcap T_2}^{LR} = C_{T_1}^{LR} \sqcap C_{T_2}^{LR} & C_{T_1 \sqcap T_2}^{RL} = C_{T_1}^{RL} \sqcap C_{T_2}^{RL} \\ C_{T_1 \sqcup T_2}^{LR} = C_{T_1}^{LR} \sqcup C_{T_2}^{LR} & C_{T_1 \sqcup T_2}^{RL} = C_{T_1}^{RL} \sqcup C_{T_2}^{RL} \\ C_{\neg T}^{LR} = C_{\text{curr}}^{LR} \sqcup \neg C_T^{LR} & C_{\neg T}^{RL} = C_{\text{curr}}^{RL} \sqcup \neg C_T^{RL} \\ C_{T^-}^{LR} = C_T^{RL} & C_{T^-}^{RL} = C_T^{LR} \end{array}$$

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

The terminological constraints concerning time relations are then captured as follows:

$$C_{T_1}^{LR} \sqsubseteq C_{T_2}^{LR} \text{ and } C_{T_1}^{RL} \sqsubseteq C_{T_2}^{RL} \text{ for each } T_1 \sqsubseteq T_2 \in \mathcal{T}.$$

We add the following constraints to capture the behaviour of

the curr and forever time relations:

$$\begin{aligned} \text{curr} &\sqsubseteq \text{curr}^-, \text{forever} \sqsubseteq \text{forever}^-, \\ \text{curr} &\sqsubseteq T, T \sqsubseteq \text{forever} \text{ for } T \in \text{TR}(\mathcal{T}). \end{aligned}$$

We call the terminology obtained above \mathcal{T}^* .

Lemma 2 Let $\mathcal{T} \models T_1 \sqsubseteq T_2$ be a time relation implication problem. Then $\mathcal{T} \models T_1 \sqsubseteq T_2$ iff $\mathcal{T}^* \models C_{T_1}^{LR} \sqsubseteq C_{T_2}^{LR}$.

It is also easy to see that the problem is co-NP-complete by reduction to unsatisfiability in propositional logic.

Reasoning in full $\mathcal{DLFD}_{\text{temp}}$

Now given a $\mathcal{DLFD}_{\text{temp}}$ terminology \mathcal{T} , we extend \mathcal{T}^* to be a \mathcal{DLF} terminology consisting of the following additional subsumptions:

1. $C_T^{LR} \sqsubseteq \forall f.C_T^{LR}$ and $C_T^{RL} \sqsubseteq \forall f.C_T^{RL}$ for all $T \in \text{TR}(\mathcal{T})$ and features f in \mathcal{T} ,
2. $C^L \sqsubseteq D^L$ and $C^R \sqsubseteq D^R$ for each subsumption axiom $C \sqsubseteq D \in \mathcal{T}$ that is not a time relation axiom and such that D is an atemporal concept description;
3. $C^L \sqcap D^R \sqcap C_T^{LR} \sqcap (\bigwedge_{i \leq k} \forall Pf_i.Eq) \sqsubseteq \forall Pf.Eq$ and $C^R \sqcap D^L \sqcap C_T^{RL} \sqcap (\bigwedge_{i \leq k} \forall Pf_i.Eq) \sqsubseteq \forall Pf.Eq$ for each subsumption axiom $C \sqsubseteq D : Pf_1, \dots, Pf_k \rightarrow_T Pf \in \mathcal{T}$,
4. $C^L \sqcap C_T^{LR} \sqcap Eq \sqsubseteq D^R$ and $C^R \sqcap C_T^{RL} \sqcap Eq \sqsubseteq D^L$ for each $C \sqsubseteq \square_T D \in \mathcal{T}$,
5. $(Eq \sqcap C_{\text{curr}}^{LR}) \sqsubseteq \forall f.Eq$ for each primitive feature f in \mathcal{T} , and

6. $(\text{Eq} \sqcap C_{\text{curr}}^{LR} \sqcap A^L) \sqsubseteq A^R$ and $(\text{Eq} \sqcap C_{\text{curr}}^{LR} \sqcap A^R) \sqsubseteq A^L$ for each primitive concept A in \mathcal{T} ,

where the labelled descriptions C^L and D^L (resp., C^R and D^R) denote \mathcal{DLF} concept descriptions C and D in which all occurrences of primitive concept description A has been replaced by A^L (resp., A^R). The assertions state how a $\mathcal{DLFD}_{\text{temp}}$ terminology can be simulated by a \mathcal{DLF} terminology by simulating two interpretations (“left” and “right”) using the labelled concept descriptions, and how equalities induced by TPFDs can be simulated by the auxiliary primitive description Eq . Note also, that this extension of \mathcal{T}^* is conservative w.r.t. reasoning about time relations.

All that remains is to translate a posed question—an inclusion dependency—to an appropriate inclusion dependency in \mathcal{DLF} . However, note that the grammar rules prohibit a TPFD from occurring in the scope of negation. (Removing this restriction would lead immediately to undecidability (Toman and Weddell 2008).) Thus, $\mathcal{DLFD}_{\text{temp}}$ is not fully closed under negation and this final translation depends on the concept occurring on the right-hand-side of the posed question. In particular, given $\mathcal{C} = C \sqsubseteq D'$, there are three cases to consider:

1. if D' is an atemporal description D , then define \mathcal{C}^* to be $C^L \sqsubseteq D^L$;
2. if D' is of the form $D : \text{Pf}_1, \dots, \text{Pf}_k \rightarrow_T \text{Pf}$, then define \mathcal{C}^* to be

$$C^L \sqcap D^R \sqcap C_T^{LR} \sqcap (\sqcap_{i \leq k} \forall \text{Pf}_i. \text{Eq}) \sqsubseteq \forall \text{Pf}. \text{Eq};$$

otherwise,

3. otherwise, when D' is of the form $\square_T D$, define \mathcal{C}^* to be

$$C^L \sqcap C_T^{LR} \sqcap \text{Eq} \sqsubseteq D^R.$$

Note how the non-trivial cases—cases 2 and 3—use the concept labelling to simulate multiple worlds and objects in a single \mathcal{DLF} interpretation.

Theorem 3 *Let $\mathcal{T} \models \mathcal{C}$ be a $\mathcal{DLFD}_{\text{temp}}$ implication problem. Then*

$$\mathcal{T} \models \mathcal{C} \text{ if and only if } \mathcal{T}^* \models \mathcal{C}^*.$$

Proof Outline: Let $\mathcal{C} = C \sqsubseteq D$. The case where D is atemporal 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} (Toman and Weddell 2005a).

We now consider both implications in the remaining cases where \mathcal{C} has the form

$$C \sqsubseteq D : \text{Pf}_1, \dots, \text{Pf}_k \rightarrow_T \text{Pf} \text{ or } C \sqsubseteq \square_T D.$$

Possible conjunctions that are allowed on the right-hand side of \mathcal{C} reduce to one of the above cases simply by considering

each conjunct separately. We consider the $\mathcal{C} = C \sqsubseteq D : \text{Pf}_1, \dots, \text{Pf}_k \rightarrow_T \text{Pf}$ case first:

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

$$C^L \sqcap D^R \sqcap C_T^{LR} \sqcap (\sqcap_{i \leq k} \forall \text{Pf}_i. \text{Eq}) \sqcap \neg \forall \text{Pf}. \text{Eq}.$$

We construct a model \mathcal{I} for \mathcal{T} that falsifies \mathcal{C} 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 $T = \text{curr}$ we have $t_1 = t_2$. It therefore follows that $o_1 \neq o_2$. The construction then proceeds as in the atemporal case (Toman and Weddell 2005a) 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 $\text{Pf}(o_1)$ and $\text{Pf}(o_2)$, where Pf is a path expression. The temporal interpretation is then defined as

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

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

$$\{(\text{Pf}(o_1), \text{Pf}(o_1)), (\text{Pf}(o_2), \text{Pf}(o_2)) \mid \text{Pf} \text{ path description}\} \cup \{(\text{Pf}(o_1), \text{Pf}(o_2)), (\text{Pf}(o_2), \text{Pf}(o_1)) \mid (\text{Pf})^{\mathcal{I}^*}(o) \in (\text{Eq})^{\mathcal{I}^*}\},$$

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:

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

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 essentially represented by terms).

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

(\Leftarrow) Now assume $\mathcal{T} \not\models \mathcal{C}$. Then there is an interpretation \mathcal{I} that is a model for \mathcal{T} but falsifies \mathcal{C} . Hence, there must be $o_1 \in (C)^{\mathcal{I}(t_1)}$ and $o_2 \in (D)^{\mathcal{I}(t_2)}$ such that $(t_1, t_2) \in (T)^{\mathcal{I}}$, $(\text{Pf}_i)^{\mathcal{I}(t_1)}(o_1) = (\text{Pf}_i)^{\mathcal{I}(t_2)}(o_2)$, and

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

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

$$\text{C}^L \sqcap \text{D}^R \sqcap \text{C}_T^{\text{LR}} \sqcap (\bigwedge_{i \leq k} \forall \text{Pf}_i.\text{Eq}) \sqcap \neg \forall \text{Pf}.\text{Eq}.$$

The remaining cases for $\square_T D$ are as follows: The case $C = C \sqsubseteq \square_T D$ for $T = \text{curr}$ reduces to the plain \mathcal{DLF} case as $C \sqsubseteq \square_{\text{curr}} D$ is the same as $C \sqsubseteq D$.

For the case $C = C \sqsubseteq \square_T D$ where $T \neq \text{curr}$ we follow the development similar to the one in the TPFD case: in a counterexample interpretation \mathcal{I} to $\mathcal{T} \models C$ we can always find $t_1 \neq t_2$ and an object o that violates C , i.e., such that

$$o \in (C)^{\mathcal{I}(t_1)} - (D)^{\mathcal{I}(t_2)}.$$

We use this object to extract from \mathcal{I} a counterexample interpretation \mathcal{I}^* that falsifies

$$\text{C}^L \sqcap \text{C}_T^{\text{LR}} \sqcap \text{Eq} \sqsubseteq \text{D}^R,$$

and similarly, from an interpretation \mathcal{I}^* that shows $\mathcal{T}^* \not\models C^*$ we extract an interpretation \mathcal{I} that shows $\mathcal{T} \not\models C$. \square

Since $\mathcal{T}^* \models C^*$ is a \mathcal{DLF} implication problem we have:

Corollary 4 *The logical implication problem for $\mathcal{DLFD}_{\text{temp}}$ is decidable and EXPTIME-complete.*

Proof Outline: Follows immediately from Theorem 3 above and results in (Toman and Weddell 2005a). \square

4 Extensions and Lower Bounds

In this section, we consider extending $\mathcal{DLFD}_{\text{temp}}$ with a temporal concept constructor for expressing eventualities. Consider in particular a ‘‘diamond’’ constructor of the form $\diamond_T C$ with its semantics given by

$$(\diamond_T C)^{\mathcal{I}(t)} = \{x \in \Delta_t : \exists(t, t') \in (T)^\mathcal{I}. x \in (C)^{\mathcal{I}(t')}\}.$$

We shall refer to this extended logic as $\mathcal{DLFD}_{\text{temp}}^\diamond$.

A Lower Bound for Eventualities

We now show that adding a simple version of this eventuality with the form $\diamond_{\text{forever}} C$ already leads to a significant increase in the complexity of the decision problem. This is accomplished by virtue of a reduction of exponentially space bounded ATMs to our logic in a fashion similar to the lower bound construction for $S5_{ALCQT}$: the worlds of an intended model represent tape cells; computation is then modelled by traversing appropriate features similarly to the construction for $S5_{ALCQT}$ (Artale, Lutz, and Toman 2007).

However, unlike the case with $S5_{ALCQT}$, we are unable to rely on the constant domain assumption in the reduction which prevents us from using the same method to generate sufficiently many tape cells (by using a binary ALC tree of depth n). We therefore use an alternative construction that employs the eventuality $\diamond_{\text{forever}}$ to construct the needed tape cells. These cells are labelled, in the end, by the primitive concept description C and can be identified by the values of binary *counters* implemented by the primitive concept descriptions C_i , $0 < i \leq n$. In the construction, we use additional auxiliary primitive concept descriptions L_i and D_i ($0 \leq i \leq n$):

- For $0 \leq i \leq n$:

$$L_i \sqsubseteq (\square_{\text{forever}} L_i) \sqcap \forall f.(L_i \sqcap L_{i+1})$$

The concepts L_0, \dots, L_{n-1} mark n *levels* connected (eventually) by a feature f .

- For $0 < i \leq n$:

$$\begin{aligned} D_i \sqcap L_i &\sqsubseteq \square_{\text{forever}} D_i \sqcap \forall f.D_i \\ &\quad \sqcap \diamond_{\text{forever}} \forall f.D_{i+1} \sqcap \diamond_{\text{forever}} \forall f.\neg D_{i+1}, \\ \neg D_i \sqcap L_i &\sqsubseteq \square_{\text{forever}} \neg D_i \sqcap \forall f.\neg D_i \\ &\quad \sqcap \diamond_{\text{forever}} \forall f.D_{i+1} \sqcap \diamond_{\text{forever}} \forall f.\neg D_{i+1}. \end{aligned}$$

The goal of these assertions is to create 2^i objects at level i , each labelled by a unique combination of primitive concepts D_j and their negations for $j \leq i$. These combinations can be considered *counter values* in which the D_j concepts serve as bits. Note that the construction only works assuming non-constant domains.

- For $0 < i \leq n$:

$$D_i \sqsubseteq D_i : Id \rightarrow_{\text{forever}} f^{n-i}.g.$$

These constraints ensure that, from the n^{th} level, the feature g leads to a *single object*.

- We then *copy* the counters across the g feature:

$$D_i \sqsubseteq \forall g.C_i \quad \neg D_i \sqsubseteq \forall g.\neg C_i$$

Hence, in each of the worlds in which the above object exists it must be *labelled* by a *counter value*. Since each distinct counter value must be present, the construction yields exponentially many worlds (as no two counter values can coexist in a single world).

- We assign the concept C to the cells at this level:

$$L_0 \sqcap D_0 \sqsubseteq \forall f^n.g.C \quad C \sqsubseteq \square_{\text{forever}} C$$

These cells are the tape cells at the beginning of the computation; their successors will hold the contents of the tape corresponding to the moves of the ATM.

We then ensure that the contents of the new counters (based on the concepts C_i) cannot change by traversing additional features:

$$C_i \sqsubseteq \forall h_j.C_i \quad \neg C_i \sqsubseteq \forall h_j.\neg C_i$$

The features h_j are used to model transitions of the ATM and, together with the *tape*, to encode a computation of an ATM with an exponential bound on space. This follows the development used for $S5_{ALCQT}$ (Artale, Lutz, and Toman 2007) since starting from the C object we can, w.l.o.g., assume that all objects needed to encode the computation appear in all of the exponentially-many worlds. This follows from totality of features and the fact that the *counters* cannot change by traversing h_j . The above construction (with the additional assertions coding states and transitions of an ATM) yields the following:

Theorem 5 *The logical implication problem for $\mathcal{DLFD}_{\text{temp}}^\diamond$ is 2-EXPTIME-hard. This remains true even when all time relations correspond to forever.*

Upper bounds for $\mathcal{DLFD}_{\text{temp}}^\diamond$ are beyond the scope of this paper. However, techniques similar to those used for $S5_{ALCQT}$ (Artale, Lutz, and Toman 2007) are likely to work since, for upper bounds, a varying domain can easily be modelled by additional concepts in a constant domain setting.

5 Summary and Discussion

We have introduced $\mathcal{DLFD}_{\text{temp}}$, a Boolean complete description logic with a pair of concept constructors for capturing the “always” modality and dynamic or temporal forms of equality generating constraints called *temporal path functional dependencies* (TPFDs). We have illustrated how these constructors can be used to capture and reason about temporal keys and functional dependencies for a hypothetical distributed hospital database and 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 (Toman and Weddell 2005a). Finally, we have also proven that no such bound remains possible if one adds any ability to express eventualities.

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 transitive or that encode (an approximation to) “next time”. Note however, that adding standard LTL operators, such as *next time*, leads immediately to undecidability. This follows since $\mathcal{DLFD}_{\text{temp}}$ would then be capable of expressing global features (which in turn enables a tiling reduction similar to a reduction for \mathcal{ALC}_{LTL}).
- 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 (Toman and Weddell 2008). We wish to investigate a similar possibility for 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 $\{Snum, Dept.Hospital\}$.
- There is also the possibility of treating the domain of time points \mathcal{W} as a concrete domain in $\mathcal{DLFD}_{\text{temp}}$. This would accommodate so-called data time-stamping in which factual data correlates knowledge of, e.g., updates over time. However, the interaction of TPFDs with the additional structure of this concrete domain requires care: the combination of functional dependencies with discretely ordered concrete time domains, for example, can lead to undecidability.
- 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 reduction is considerably more involved.

References

- Artale, A., and Franconi, E. 2005. Temporal description logics. In Fisher, M.; Gabbay, D.; and Vila, L., eds., *Handbook of Temporal Reasoning in Artificial Intelligence*. Elsevier. 375–388.
- Artale, A.; Lutz, C.; and Toman, D. 2007. A description logic of change. In *Int. Joint Conference on Artificial Intelligence (IJCAI-07)*, 218–223.
- Biskup, J., and Polle, T. 2000. Decomposition of Database Classes under Path Functional Dependencies and Onto

- Constraints. In *Foundations of Information and Knowledge Systems*, 31–49.
- Biskup, J., and Polle, T. 2003. Adding inclusion dependencies to an object-oriented data model with uniqueness constraints. *Acta Informatica* 39:391–449.
- Borgida, A., and Weddell, G. E. 1997. Adding Uniqueness Constraints to Description Logics (Preliminary Report). In *International Conference on Deductive and Object-Oriented Databases*, 85–102.
- Calvanese, D.; De Giacomo, G.; and Lenzerini, M. 2001. Identification Constraints and Functional Dependencies in Description Logics. In *Proc. of the 17th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, 155–160.
- DeHaan, D.; Toman, D.; and Weddell, G. E. 2003. Rewriting Aggregate Queries using Description Logics. In *Description Logics 2003*, 103–112. CEUR-WS vol.81.
- Gabbay, D.; Kurucz, A.; Wolter, F.; and Zakharyashev, M. 2003. *Many-dimensional modal logics: theory and applications*. Studies in Logic. Elsevier.
- Ito, M., and Weddell, G. E. 1994. Implication Problems for Functional Constraints on Databases Supporting Complex Objects. *Journal of Computer and System Sciences* 49(3):726–768.
- Khizder, V. L.; Toman, D.; and Weddell, G. E. 2000. Reasoning about Duplicate Elimination with Description Logic. In *Rules and Objects in Databases (DOOD, part of CL'00)*, 1017–1032.
- Khizder, V. L.; Toman, D.; and Weddell, G. E. 2001. On Decidability and Complexity of Description Logics with Uniqueness Constraints. In *International Conference on Database Theory ICDT'01*, 54–67.
- Lutz, C., and Milicic, M. 2004. Description Logics with Concrete Domains and Functional Dependencies. In *Proc. European Conference on Artificial Intelligence (ECAI)*, 378–382.
- Lutz, C.; Areces, C.; Horrocks, I.; and Sattler, U. 2003. Keys, Nominals, and Concrete Domains. In *Proc. of the 18th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, 349–354.
- Sattler, U.; Calvanese, D.; and Molitor, R. 2003. Relationships with other formalisms. In *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press. chapter 4, 137–177.
- Stanchev, L., and Weddell, G. E. 2003. Index Selection for Embedded Control Applications using Description Logics. In *Description Logics 2003*, 9–18. CEUR-WS vol.81.
- Toman, D., and Weddell, G. E. 2001. On Attributes, Roles, and Dependencies in Description Logics and the Ackermann Case of the Decision Problem. In *Description Logics 2001*, 76–85. CEUR-WS vol.49.
- Toman, D., and Weddell, G. E. 2004. Attribute Inversion in Description Logics with Path Functional Dependencies. In *Description Logics 2004*, 178–187. CEUR-WS vol.104.
- Toman, D., and Weddell, G. 2005a. On Reasoning about Structural Equality in XML: A Description Logic Approach. *Theoretical Computer Science* 336(1):181–203.
- Toman, D., and Weddell, G. 2005b. 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)*, 603–608.
- Toman, D., and Weddell, G. 2008. On keys and functional dependencies as first-class citizens in description logics. *Journal of Automated Reasoning* 40(2–3):117–132.
- W3C. 2004a. Resource Description Framework (RDF), <http://www.w3.org/RDF/>.
- W3C. 2004b. Web Ontology Language (OWL) <http://www.w3.org/2004/owl/>.
- Weddell, G. 1989. A Theory of Functional Dependencies for Object Oriented Data Models. In *International Conference on Deductive and Object-Oriented Databases*, 165–184.
- Wijsen, J. 1999. Temporal FDs on Complex Objects. *ACM Transactions on Database Systems* 24(1):127–176.