Connecting abstract description systems

Oliver Kutz, Frank Wolter*
and Michael Zakharyaschev†

March 3, 2002

Abstract

Combining knowledge representation and reasoning formalisms like description logics (DLs), temporal logics, and logics of space, is worthwhile but difficult. It is worthwhile because usually realistic application domains comprise various aspects of the world, thus requiring suitable combinations of formalisms modeling each of these aspects. It is difficult because the computational behavior of the resulting hybrids is often much worth than the behavior of its components. In this paper we propose a combination method which is robust in the computational sense and still allows for certain interactions between the combined systems. The combination method, called *E-connection*, will be defined and investigated for so-called abstract description systems (ADS) which include all standard description logics, various logics of time and space, modal logics, and epistemic logics. The main theoretical result is that every E-connection of any finite number of decidable ADSs is decidable as well. Four instances of E-connections of ADSs will be discussed: (1) the E-connection of DLs with the logic $\mathcal{MS}$ intended for quantitative reasoning about space, (2) the E-connection of DLs with the logic $\mathsf{S4}_u$ (containing $\mathsf{RCC\text{-}8}$) that can be used for qualitative reasoning about space, (3) the E-connection of two DLs ($\mathcal{ALCO}$ and $\mathcal{SHIQ}$), and (4) the E-connection of DLs with propositional temporal logic PTL and $\mathsf{S4}_u$.

*Computer Science Institute, Intelligent Systems Department, University of Leipzig, Augustus-Platz 10–11, 04109 Leipzig, Germany.
Email: {kutz,wolter}@informatik.uni-leipzig.de

†Department of Computer Science, King's College London, Strand, London WC2R 2LS, U.K. Email: mz@dcs.kcl.ac.uk

1 Introduction

Combining knowledge representation and reasoning formalisms is worthwhile but difficult. It is worthwhile because usually realistic application domains comprise various aspects of the world around us, thus requiring suitable combinations of formalisms modeling each of these aspects. It is difficult because the computational behavior of the resulting hybrids is often much worse than the behavior of the combined components; see, e.g., [11, 14, 16]. To be more specific, consider three examples related to description logics.

(i) Classical description logics (DLs) represent knowledge at a rather abstract logical level. To cope with applications which require predefined predicates or temporal and spatial dimensions, combinations of DLs with *concrete domains* such as the natural numbers equipped with predicates like $<$, Allen's interval logic [1], or region connection calculus $\mathsf{RCC\text{-}8}$ [29] have been proposed [7, 19]. The addition of a concrete domain to a DL is a rather sensitive operation as far as the preservation of its nice computational properties is concerned: even 'weak' DLs combined with 'weak' concrete domains can become undecidable; see, e.g., [8, 18, 28]. In fact, to investigate DLs with concrete domains is rather hard and requires developing new techniques, cf. [27].

(ii) Standard DLs have been designed to represent *static* knowledge which is time- and agent-independent. To take into account the dynamic aspects of knowledge, DLs have been extended with temporal, dynamic, epistemic and other intentional operators [25, 9, 10, 13, 2, 39, 41]. The resulting formalisms become 'many-dimensional' and sometimes show rather nasty computational behavior: combinations of simple description logics (say, $\mathcal{ALC}$) with simple temporal logics (say, propositional temporal logic PTL) can be highly undecidable [10, 41, 40, 16]. These logics also require new approaches [41, 16], and it is still unclear whether practical reasoning systems can be developed for many-dimensional logics, cf. [26].

(iii) Often there is a need to combine two or more description logics: while one part of the application domain may require constructors of DL_1, another part can only be represented using constructors of DL_2. Putting the constructors of DL_1 and DL_2 together to form a new DL may result in an undecidable logic, even if both components are decidable. As an example, consider the DLs $\mathcal{ALCF}$ (extending $\mathcal{ALC}$ with functional roles (or features) and the same-as constructor (or agreement) on chains of functional roles) and $\mathcal{ALC}^{+,\circ,\sqcup}$ (extending $\mathcal{ALC}$ with the transitive closure, composition, and union of roles). For both DLs, the subsumption of concept descriptions is known to be decidable [21, 32, 6]. However, the subsumption problem for their *union* $\mathcal{ALCF}^{+,\circ,\sqcup}$ is undecidable [3]. Recently, *fusions* (or independent joints) have been proposed as a more robust way of combining DLs [5, 4, 36]. But even fusions behave badly if the class of models is not closed under disjoint unions, which is the case when nominals or the negation of roles are required [5, 4] (or if we combine logics of time and space—while linear orders are natural models of time, their disjoint unions are certainly not).

In this situation, a natural question arises as to whether there exist at all sufficiently general and useful ways of combining representation and reasoning systems preserving their good computational properties. *The main aim of this paper is to give a positive answer to this question.* We propose a combination method which is robust in the computational sense and still allows for certain interactions between the combined components. Given n 'description systems' $L_1, \ldots, L_n$ talking about domains $D_1, \ldots, D_n$, we form a new language L containing the languages L_i, $1 \leq i \leq n$, and talking about the disjoint union $\bigcup_{i=1}^n D_i$, in which the D_i are connected by a finite number of relations $E_j \subseteq D_1 \times \cdots \times D_n$, $1 \leq j \leq m$. The fragments L_i of L still talk about D_i; moreover, the super-language L contains $n \times m$ extra $(n-1)$-ary operators $\mathcal{E}_j^i$ which, given an input $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$, for $X_\ell \subseteq D_\ell$, return the set of all $x \in D_i$ such that

$$\forall \ell \neq i \; \exists x_\ell \in X_\ell \; (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \in E_j.$$

In other words, $\mathcal{E}_j^i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_n)$ is the value of the i-th factor of

$$(X_1 \times \cdots \times X_{i-1} \times D_i \times X_{i+1} \times \cdots \times X_n) \cap E_j.$$

We call L an *E-connection* of $L_1, \ldots, L_n$.

This is a rough idea. To make it more precise, we use the notion of *abstract description system* (ADS, for short) introduced in [5, 4]. Basically, all description, modal, temporal, epistemic and similar logics (in particular, modal logics of space) can be conceived of as ADSs. For this reason, ADSs form a right level of abstraction to study combinations of knowledge representation formalisms such as DLs, temporal logics (TLs, for short), spatial logics (SLs), etc. The main technical result of this paper is the following theorem: *every E-connection of any finite number of decidable ADSs is decidable as well.*

Here are three simple examples of E-connections; in more detail they will be discussed in Section 4.

DL–SL: A DL L_1 (say, $\mathcal{ALC}$ or $\mathcal{SHIQ}$ [22]) talks about a domain D_1 of abstract objects. A spatial logic L_2 (say, qualitative $\mathsf{S4}_u$ [35, 12, 33, 16] or quantitative $\mathcal{MS}$ of [34, 24]) talks about a spatial domain D_2. An obvious E-connection is given by the relation $E \subseteq D_1 \times D_2$ defined by taking $(x, y) \in E$ iff y belongs to the spatial extension of x whenever x occupies some space. Then, given an L_1-concept C, say, river, the operator $\mathcal{E}^2(C)$ provides us with the spatial extension of all rivers. Conversely, given a spatial region X of L_2, say, the Alps, $\mathcal{E}^1(X)$ provides the concept comprising all objects the spatial extensions of which have a non-empty intersection with X. The concept $\mathsf{country} \sqcap \mathcal{E}^1(X)$ would denote then the union of all alpine countries.

DL–TL: Now, let L_3 be a temporal logic (say, point-based PTL [15] or Halpern-Shoham's logic of intervals HS [20]) and let D_3 be a set of time points or, respectively, time intervals interpreting L_3. In this case, a natural relation $E \subseteq D_1 \times D_3$ is given by taking $(x, y) \in E$ iff y belongs to the life-span of x.

DL–SL–TL: Further, we can combine all L_1, L_2, L_3 above into a single formalism by defining a ternary relation $E \subseteq D_1 \times D_2 \times D_3$ such that $(x, y, z) \in E$ iff y belongs to the spatial extension of x at moment (interval) z.

2 Abstract description systems

An abstract description system consists of an abstract description language and a class of admissible models specifying the intended semantics.

Definition 1. An *abstract description language* (ADL) is determined by a countably infinite set $\mathcal{V}$ of *set variables*, a countably infinite set $\mathcal{X}$ of *object variables*, a (possibly infinite) sequence $(R_i)_{i \in \mathcal{R}}$ of *relation symbols* of arity m_i, $i \in \mathcal{R}$, and a (possibly infinite) sequence $(f_i)_{i\in\mathcal{I}}$ of *function symbols* of arity n_i, $i \in \mathcal{I}$. The *terms* t_j of the ADL are built as follows:

$$t ::= x \mid \neg t \mid t_1 \wedge t_2 \mid t_1 \vee t_2 \mid f_i(t_1, \dots, t_{n_i}),$$

where $x \in \mathcal{V}$ and the Boolean operators $\neg$, $\wedge$, $\vee$ are different from all the f_i. The *term assertions* of the ADL are of the form

- $t_1 \sqsubseteq t_2$, where t_1 and t_2 are terms,

and the *object assertions* are

- $R_i(a_1, \dots, a_{m_i})$, for $a_1, \dots, a_{m_i} \in \mathcal{X}$ and $i \in \mathcal{R}$;
- $a : t$, for $a \in \mathcal{X}$ and t a term.

The sets of term and object assertions together form the set of the *ADL-assertions*.

Examples

(1) We remind the reader that $\mathcal{ALC}$-concept expressions C are composed from concept names by means of the operators $\sqcap$, $\neg$, $\forall R$, and $\exists R$, where R is a role name. The concept expressions of $\mathcal{ALC}$ (or any other standard description logic extending $\mathcal{ALC}$) can be regarded as terms $C^\sharp$ of an ADL $\mathcal{ALC}^\sharp$. Namely, with each concept name A we associate a set-variable $A^\sharp$, and with each role R we associate two unary function symbols $f_{\forall R}$ and $f_{\exists R}$. And then we put inductively:

$$\begin{aligned} (C \sqcap D)^\sharp &= C^\sharp \wedge D^\sharp \\ (\neg C)^\sharp &= \neg C^\sharp \\ (\forall R.C)^\sharp &= f_{\forall R}(C^\sharp) \\ (\exists R.C)^\sharp &= f_{\exists R}(C^\sharp) \end{aligned}$$

The object names of $\mathcal{ALC}$ are treated as object variables of $\mathcal{ALC}^\sharp$ and the role names as its binary relations. Thus, term assertions of $\mathcal{ALC}^\sharp$ correspond to general TBoxes, while object assertions correspond to ABoxes. (The connection between roles R and the function symbols $f_{\forall R}$ and $f_{\exists R}$ is fixed by choosing a proper class of admissible models; see below.) For transformations of more expressive DLs into ADLs consult [5, 4].

(2) The language of the logic $\mathsf{S4}_u$ (i.e., Lewis's modal system $\mathsf{S4}$ with the universal modality), with topological spaces as the intended interpretations, consists of set variables $X_1, \dots$ (in the modal context, propositional variables), the interior operator $\boldsymbol{I}$ (the necessity operator), the closure operator $\boldsymbol{C}$ (the possibility operator), the universal quantifier $\boxed{\forall}$ (the universal box), and the Booleans [35, 12, 33]. The corresponding ADL $\mathsf{S4}_u^\sharp$ would contain then the set variables $X_1^\sharp, \dots$, the unary function symbols $f_{\boldsymbol{I}}$, $f_{\boldsymbol{C}}$, $f_{\boxed{\forall}}$, and the Booleans (but no relation symbols). Besides, according to the definition, $\mathsf{S4}_u^\sharp$ must contain a countably infinite set of object variables a_i. The translation $^\sharp$ of $\mathsf{S4}_u$-formulas into terms of $\mathsf{S4}_u^\sharp$ is obvious; for example,

$$(\Box\varphi)^\sharp = f_\Box(\varphi^\sharp), \text{ for } \Box \in \{\boldsymbol{I}, \boldsymbol{C}, \boxed{\forall}\}.$$

Note, however, that $\mathsf{S4}_u^\sharp$ allows for object assertions of the form $a_i : t^\sharp$ which have no analogs in $\mathsf{S4}_u$.

(3) The logic of metric spaces $\mathcal{MS}$ of [34, 24][1] consists of terms constructed from set variables X_i and nominals N_i using the Booleans and the operators $\mathsf{A}_{\leq r}$, $\mathsf{E}_{\leq r}$, $\mathsf{A}_{>r}$ and $\mathsf{E}_{>r}$, for $r \in \mathbb{Q}^+$. Intuitively, given a set X in a metric space, $\mathsf{E}_{\leq r}X$ is the set of all points in the space located at distance $\leq r$ from (at least one point in) X. A point is in $\mathsf{A}_{>r}X$ iff the whole complement of its r-neighborhood is in X. Again, terms t of $\mathcal{MS}$ can be translated into terms $t^\sharp$ of an ADL $\mathcal{MS}^\sharp$ by associating with every nominal N_i a 0-ary function symbol f_{N_i} and with the operators $\mathsf{A}_{\leq r}$, $\mathsf{E}_{\leq r}$, $\mathsf{A}_{>r}$, $\mathsf{E}_{>r}$ corresponding unary function symbols $f_{\mathsf{A}_{\leq r}}$, $f_{\mathsf{E}_{\leq r}}$, $f_{\mathsf{A}_{>r}}$, $f_{\mathsf{E}_{\leq r}}$, and then proceeding as above. The location variables of $\mathcal{MS}$ correspond to the object variables of $\mathcal{MS}^\sharp$.

(4) In the same way the propositional temporal language PTL, in which formulas are composed from propositional variables by means of the Booleans and the binary operators $\mathcal{U}$ ('until') and $\mathcal{S}$ ('since'), can be represented as an ADL $\mathsf{PTL}^\sharp$. In this case we associate with $\mathcal{U}$ and $\mathcal{S}$ *binary* function symbols $f_\mathcal{U}$ and $f_\mathcal{S}$. Note again that $\mathsf{PTL}^\sharp$ allows for object assertions $a_i : t^\sharp$ which have no analogs in PTL. However, since our intended flow of time is $\langle \mathbb{N}, < \rangle$, object variables can be represented as $p \wedge \neg(\top \mathcal{U} p) \wedge \neg(\top \mathcal{S} p)$. As in (2) and (3), $\mathsf{PTL}^\sharp$ contains no relation symbols.

The semantics of ADLs is defined via abstract description models.

Definition 2. An *abstract description model* (ADM) for an ADL $\mathcal{L}$ is a structure of the form

$$\mathfrak{A} = \left\langle W, \mathcal{V}^\mathfrak{A}, \mathcal{X}^\mathfrak{A}, F^\mathfrak{A}, R^\mathfrak{A} \right\rangle,$$

where $\mathcal{V}^\mathfrak{A} = (x^\mathfrak{A})_{x \in \mathcal{V}}$, $\mathcal{X}^\mathfrak{A} = (a^\mathfrak{A})_{a \in \mathcal{X}}$, $F^\mathfrak{A} = (f_i^\mathfrak{A})_{i \in \mathcal{I}}$, $R^\mathfrak{A} = (R_i^\mathfrak{A})_{i \in \mathcal{R}}$, W is a non-empty set, $x^\mathfrak{A} \subseteq W$, $a^\mathfrak{A} \in W$, each $f_i^\mathfrak{A}$ is a function mapping n_i-tuples $\langle X_1, \ldots, X_{n_i} \rangle$ of subsets of W to a subset of W, and the $R_i^\mathfrak{A}$ are m_i-ary relations on W. The *value* $t^\mathfrak{A} \subseteq W$ of an $\mathcal{L}$-term t in $\mathfrak{A}$ is defined inductively by taking

$$(t_1 \wedge t_2)^\mathfrak{A} = t_1^\mathfrak{A} \cap t_2^\mathfrak{A}, \ (t_1 \vee t_2)^\mathfrak{A} = t_1^\mathfrak{A} \cup t_2^\mathfrak{A},$$
$$(\neg t)^\mathfrak{A} = W \setminus (t)^\mathfrak{A}, \ (f_i(t_1, \ldots, t_k))^\mathfrak{A} = f_i^\mathfrak{A}(t_1^\mathfrak{A}, \ldots, t_k^\mathfrak{A}).$$

The *truth-relation* $\mathfrak{A} \models \varphi$ for an $\mathcal{L}$-assertion φ is defined in the obvious way:

- $\mathfrak{A} \models R_i(a_1, \ldots, a_{m_i})$ iff $R_i^\mathfrak{A}(a_1^\mathfrak{A}, \ldots, a_{m_i}^\mathfrak{A})$,
- $\mathfrak{A} \models a : t$ iff $a^\mathfrak{A} \in t^\mathfrak{A}$,
- $\mathfrak{A} \models t_1 \sqsubseteq t_2$ iff $t_1^\mathfrak{A} \subseteq t_2^\mathfrak{A}$.

If $\mathfrak{A} \models \varphi$ holds, we say that φ is *satisfied* in $\mathfrak{A}$.

Definition 3. An *abstract description system* (ADS) is a pair $(\mathcal{L}, \mathcal{M})$, where $\mathcal{L}$ is an ADL and $\mathcal{M}$ is a class of ADMs for $\mathcal{L}$ that is closed under the following operation: if $\mathfrak{A} = \langle W, \mathcal{V}^\mathfrak{A}, \mathcal{X}^\mathfrak{A}, F^\mathfrak{A}, R^\mathfrak{A} \rangle$ is in $\mathcal{M}$, and $\mathcal{V}^{\mathfrak{A}'} = (x^{\mathfrak{A}'})_{x \in \mathcal{V}}$, $\mathcal{X}^{\mathfrak{A}'} = (a^{\mathfrak{A}'})_{a \in \mathcal{X}}$ are new assignments to set and object variables in W then $\mathfrak{A}' = \left\langle W, \mathcal{V}^{\mathfrak{A}'}, \mathcal{X}^{\mathfrak{A}'}, F^\mathfrak{A}, R^\mathfrak{A} \right\rangle \in \mathcal{M}$.

Let us now return to examples (1)–(4) above and supply the ADLs $\mathcal{ALC}^\sharp$, $\mathsf{S4}_u^\sharp$, $\mathcal{MS}^\sharp$ and $\mathsf{PTL}^\sharp$ with their intended ADMs.

[1] The logic we consider here is called $\mathcal{MS}_2$ in [34] and $\mathcal{MS}^\sharp$ in [24].

Examples (cont.)

(1) For $\mathcal{ALC}^\sharp$, the class $\mathcal{M}$ of ADMs is defined as follows. For any $\mathcal{ALC}$-model

$$\mathcal{I} = \left\langle \Delta, A_1^{\mathcal{I}}, \ldots, R_1^{\mathcal{I}}, \ldots, a_1^{\mathcal{I}}, \ldots \right\rangle,$$

$\mathcal{M}$ contains the model $\mathfrak{M} = \left\langle \Delta, \mathcal{V}^{\mathfrak{M}}, \mathcal{X}^{\mathfrak{M}}, F^{\mathfrak{M}}, R^{\mathfrak{M}} \right\rangle$, where F consists of the function symbols $f_{\forall R_i}$ and $f_{\exists R_i}$, and R is the set of all role names of $\mathcal{ALC}$:

- $(A^\sharp)^{\mathfrak{M}} = A^{\mathcal{I}}$, for all concept names A;
- $a^{\mathfrak{M}} = a^{\mathcal{I}}$, for all object names a;
- $R_i^{\mathfrak{M}} = R_i^{\mathcal{I}}$, for all roles R_i;
- $f_{\forall R_i} X = \{d \in \Delta \mid \forall d' \in \Delta \ (d R_i^{\mathcal{I}} d' \to d' \in X)\}$, for all role names R_i;
- $f_{\exists R_i} X = \{d \in \Delta \mid \exists d' \in \Delta \ (d R_i^{\mathcal{I}} d' \wedge d' \in X)\}$, for all role names R_i.

(2) An $\mathsf{S4}_u$-model $\mathcal{I} = \left\langle T, \mathbb{I}, \mathbb{C}, X_1^{\mathcal{I}}, \ldots \right\rangle$ consists of a topological space $\langle T, \mathbb{I} \rangle$, where $\mathbb{I}$ is an interior operator mapping subsets X of T to their interior $\mathbb{I}(X)$ and satisfying the equations $\mathbb{I}(X \cap Y) = \mathbb{I}(X) \cap \mathbb{I}(Y)$, $\mathbb{I}\mathbb{I}(X) = \mathbb{I}(X)$, $\mathbb{I}(X) \subseteq X$ and $\mathbb{I}(T) = T$ for all $X, Y \subseteq T$, $\mathbb{C}$ is the closure operator defined by $\mathbb{C}(X) = T - \mathbb{I}(T - X)$, and the $X_i^{\mathcal{I}}$ are subsets of T (interpreting the set variables of $\mathsf{S4}_u$). Of course, the operators $\boldsymbol{I}$ and $\boldsymbol{C}$ of $\mathsf{S4}_u$ are interpreted by $\mathbb{I}$ and $\mathbb{C}$, respectively. We define the class $\mathcal{M}$ of ADMs for $\mathsf{S4}_u^\sharp$ by taking, for every such $\mathsf{S4}_u$-model $\mathcal{I}$, the ADMs

$$\mathfrak{M} = \left\langle T, \mathcal{V}^{\mathfrak{M}}, \mathcal{X}^{\mathfrak{M}}, f_{\boldsymbol{I}}^{\mathfrak{M}}, f_{\boldsymbol{C}}^{\mathfrak{M}}, f_{\boxed{\forall}}^{\mathfrak{M}} \right\rangle,$$

where $(X_i^\sharp)^{\mathfrak{M}} = X_i^{\mathcal{I}}$, $a^{\mathfrak{M}} \in \Delta$, for $a \in \mathcal{X}$, $f_{\boldsymbol{I}}^{\mathfrak{M}} = \mathbb{I}$, $f_{\boldsymbol{C}}^{\mathfrak{M}} = \mathbb{C}$ and, for every $Y \subseteq T$,

$$f_{\boxed{\forall}}^{\mathfrak{M}} Y = \begin{cases} \emptyset & \text{if } Y \neq T \\ T & \text{if } Y = T. \end{cases}$$

(3) An $\mathcal{MS}$-model $\mathcal{I} = \left\langle W, \delta, X_1^{\mathcal{I}}, \ldots, N_1^{\mathcal{I}}, \ldots a_1^{\mathcal{I}}, \ldots \right\rangle$ consists of a metric space $\langle W, \delta \rangle$ together with interpretations of set variables X_i as subsets $X_i^{\mathcal{I}}$ of W, location variables a_i as elements $a_i^{\mathcal{I}}$ of W, and nominals N_i as singleton subsets $N_i^{\mathcal{I}}$ of W. Every such model $\mathcal{I}$ gives rise to the ADM

$$\mathfrak{M} = \left\langle W, \mathcal{V}^{\mathfrak{M}}, \mathcal{X}^{\mathfrak{M}}, F^{\mathfrak{M}} \right\rangle$$

for $\mathcal{MS}^\sharp$, where $(X_i^\sharp)^{\mathfrak{M}} = X_i^{\mathcal{I}}$, $a_i^{\mathfrak{M}} = a_i^{\mathcal{I}}$, and $F^{\mathfrak{M}}$ consists of all functions f_{N_i}, $f_{\mathsf{A}_{\leq r}}$, $f_{\mathsf{E}_{\leq r}}$, $f_{\mathsf{A}_{>r}}$, $f_{\mathsf{E}_{>r}}$, for $r \in \mathbb{Q}^+$, defined by

$$\begin{aligned}
f_{N_i}^{\mathfrak{M}} &= N_i^{\mathcal{I}} \\
f_{\mathsf{A}_{\leq r}}^{\mathfrak{M}}(Y) &= \{w \in W \mid \forall x \in W \ (\delta(w,x) \leq r \to x \in Y)\} \\
f_{\mathsf{E}_{\leq r}}^{\mathfrak{M}}(Y) &= \{w \in W \mid \exists x \in W \ (\delta(w,x) \leq r \wedge x \in Y)\} \\
f_{\mathsf{A}_{>r}}^{\mathfrak{M}}(Y) &= \{w \in W \mid \forall x \in W \ (\delta(w,x) > r \to x \in Y)\} \\
f_{\mathsf{E}_{>r}}^{\mathfrak{M}}(Y) &= \{w \in W \mid \exists x \in W \ (\delta(w,x) > r \wedge x \in Y)\}
\end{aligned}$$

(4) The class of ADMs for $\mathsf{PTL}^\sharp$ consists of structures

$$\mathfrak{M} = \left\langle \mathbb{N}, \mathcal{V}^{\mathfrak{M}}, \mathcal{X}^{\mathfrak{M}}, F^{\mathfrak{M}} \right\rangle,$$

in which $F^{\mathfrak{M}}$ contains two binary functions defined by taking, for all $Y, Z \subseteq \mathbb{N}$,

$$f_{\mathcal{U}}^{\mathfrak{M}}(Y,Z) = \{u \in \mathbb{N} \mid \exists z > u(z \in Z \wedge \forall y \in (u,z) y \in Y)\}$$
$$f_{\mathcal{S}}^{\mathfrak{M}}(Y,Z) = \{u \in \mathbb{N} \mid \exists z < u(z \in Z \wedge \forall y \in (z,u) y \in Y)\},$$

where $(u,v) = \{w \in \mathbb{N} \mid u < w < v\}$.

The main reasoning task for ADSs we are concerned with is the *satisfiability problem* for finite sets of assertions. In DL, this corresponds to the satisfiability of ABoxes with respect to general TBoxes.

Definition 4. Let $\mathcal{S} = (\mathcal{L}, \mathcal{M})$ be an ADS. A finite set Γ of $\mathcal{L}$-assertions is called *satisfiable* in $\mathcal{S}$ if there is an ADM $\mathfrak{A} \in \mathcal{M}$ which satisfies all assertions in Γ.

The satisfiability problem for an ADS $\mathcal{S}$ restricted to sets Γ of *object assertions* will be called the *A-satisfiability problem* for $\mathcal{S}$. The following theorem is an almost immediate consequence of the decidability of the corresponding logics, consult [17] for $\mathsf{S4}_u$, [34] for $\mathcal{MS}$, and [15] for PTL.

Theorem 5. *The satisfiability problem is decidable for the ADSs $\mathsf{S4}_u^\sharp$, $\mathcal{MS}^\sharp$, and $\mathsf{PTL}^\sharp$. It is also decidable for $\mathcal{L}^\sharp$ whenever $\mathcal{L}$ is a DL with a decidable satisfiability problem for ABoxes with respect to general TBoxes.*

3 Connections of ADSs

We are in a position now to define E-connections formally. Suppose that we have ADSs $\mathcal{S}_i = (\mathcal{L}_i, \mathcal{M}_i)$, for $1 \leq i \leq n$, with disjoint vocabularies apart from the Boolean operators.[2] Let E be a set of $n \times m$ function symbols $\mathcal{E}_j^i$ of arity $n-1$, for $1 \leq j \leq m$, $1 \leq i \leq n$. Define by induction the notions of *i-term* and *i-assertion* of the *E-connection* $\mathcal{C} = \mathcal{C}(\mathcal{S}_1, \ldots, \mathcal{S}_n)$, for $1 \leq i \leq n$:

- every set variable of $\mathcal{L}_i$ is an i-term;
- the set of i-terms is closed under $\neg$, $\wedge$, $\vee$, and the non-Boolean function symbols of $\mathcal{L}_i$;
- if $(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$ is a sequence of j-terms, for $j \neq i$, then $\mathcal{E}_k^i(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$, for $1 \leq k \leq m$, are i-terms;
- if $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ is a sequence of object variables a_j from $\mathcal{L}_j$, for $j \neq i$, then the $\mathcal{E}_k^i(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$, $1 \leq k \leq m$, are i-terms.

The *i-term assertions* of $\mathcal{C}$ are of the form $t_1 \sqsubseteq t_2$, where t_1 and t_2 are i-terms. The *i-object assertions* are all expressions of the form $R_\ell(a_1, \ldots, a_{m_\ell})$ and $a : t$, where R_ℓ is a relation symbol of $\mathcal{L}_i$, a and the a_k are object variables of $\mathcal{L}_i$, and t an i-term. The sets of all i-term and i-object assertions, $1 \leq i \leq n$, together form the set of *assertions* of the E-connection $\mathcal{C}$.

A structure of the form $\mathfrak{M} = \langle (\mathfrak{W}_i)_{i \leq n}, (E_j)_{j \leq m} \rangle$, where $E_j \subseteq W_1 \times \cdots \times W_n$ and $\mathfrak{W}_i = \langle W_i, \mathcal{V}_i^{\mathfrak{W}}, \mathcal{X}_i^{\mathfrak{W}}, F_i^{\mathfrak{W}}, R_i^{\mathfrak{W}} \rangle \in \mathcal{M}_i$, is called a *model* for $\mathcal{C}$. The *extension* $t^{\mathfrak{M}} \subseteq W_i$

[2] It is to be noted that this condition is different from the one required for fusions of ADSs in [5, 4]: when forming fusions, we assume that the set of set variables and object variables of the ADSs to be combined coincide. In the case of E-connections these sets of symbols should be disjoint, since they are used in the combined system to represent knowledge about disjoint domains.

of an i-term t is defined by simultaneous induction. For set and object variables X and a of $\mathcal{L}_i$, we put $X^{\mathfrak{M}} = X^{\mathfrak{W}_i}$ and $a^{\mathfrak{M}} = a^{\mathfrak{W}_i}$. The inductive steps for the Booleans and function symbols of $\mathcal{L}_i$ are as before. The new clauses are:

$$(\mathcal{E}_j^i(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n))^{\mathfrak{M}} = \{x \in W_i \mid \underset{\ell \neq i}{\exists} x_\ell \in t_\ell^{\mathfrak{M}} \ (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \in E_j\},$$

and

$$(\mathcal{E}_j^i(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n))^{\mathfrak{M}} = \{x \in W_i \mid (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) \in E_j\}.$$

The *truth-relation* $\models$ between models $\mathfrak{M}$ for $\mathcal{C}$ and assertions of $\mathcal{C}$ is defined in the obvious manner:

- $\mathfrak{M} \models t_i \sqsubseteq t_2$ iff $t_1^{\mathfrak{M}} \subseteq t_2^{\mathfrak{M}}$;
- $\mathfrak{M} \models a : t$ iff $a^{\mathfrak{M}} \in t^{\mathfrak{M}}$;
- $\mathfrak{M} \models R_i(a_1, \ldots, a_m)$ iff $R_i^{\mathfrak{M}}(a_1^{\mathfrak{M}}, \ldots, a_{m_i}^{\mathfrak{M}})$.

A set Γ of assertions of $\mathcal{C}$ is called *satisfiable* if there is a model for $\mathcal{C}$ in which all assertions in Γ are true. And an assertion φ *follows* from a set of assertions Γ in $\mathcal{C}$ if $\mathfrak{M} \models \varphi$ whenever $\mathfrak{M} \models \Gamma$, for every model $\mathfrak{M}$ for $\mathcal{C}$. Note that the problem whether an assertion follows from a set of assertions can be reduced to the satisfiability problem. For example, $t_1 \sqsubseteq t_2$ follows from Γ iff $\Gamma \cup \{a : t_1 \wedge \neg t_2\}$, a a fresh object variable, is not satisfiable. Note that, for any n object variables $a_i \in \mathcal{X}_i$ we have

$$\mathfrak{M} \models a_i : \mathcal{E}_j^i(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \text{ iff } E_j(a_1^{\mathfrak{M}}, \ldots, a_n^{\mathfrak{M}})$$

Hence, we can state that the tuple $\langle a_1^{\mathfrak{M}}, \ldots, a_n^{\mathfrak{M}} \rangle$ is an instance of the relation E_j.

The main results of this paper are as follows:

Theorem 6. (i) *Suppose that the satisfiability problem for each of the ADSs $\mathcal{S}_i$, $1 \leq i \leq n$, is decidable. Then the satisfiability problem for any E-connection of the $\mathcal{S}_i$ is decidable as well.*

(ii) *If the $\mathcal{S}_i$ are decidable in EXPTIME, then the E-connection is decidable in 2EXPTIME.*

Corollary 7. *The satisfiability problem for any E-connection of DLs with decidable satisfiability problems for ABoxes with respect to general TBoxes and logics like* PTL, *$\mathcal{MS}$, and* $\mathsf{S4}_u$ *is decidable.*

We know of no example where the $\mathcal{S}_i$ are decidable in EXPTIME but the E-connection is 2EXPTIME-hard. It is an open problem whether such ADSs exist.

It is of interest that A-satisfiability is not preserved under E-connections. More precisely, the following 'negative' result holds, where $\mathcal{ALCF}$ is the extension of $\mathcal{ALC}$ with functional roles and the same-as constructor (it is known that A-satisfiability for $\mathcal{ALCF}$ is decidable [21, 27] while satisfiability is not [3]):

Theorem 8. *Let $\mathcal{S}$ be an arbitrary ADS. Then the A-satisfiability problem for any E-connection $\mathcal{C}(\mathcal{ALCF}, \mathcal{S})$ with a non-empty E is undecidable.*

Before presenting the proofs, we illustrate the notion of E-connection with illuminative examples.

4 Examples of E-connections

In this section we give four examples of E-connections using the representation formalisms introduced above. Our aim is to demonstrate the versatility of the combination technique and to outline its limits. The first three examples are 'two-dimensional,' while the fourth one connects three ADSs. To simplify notation, we will not distinguish between description, spatial, metric, or temporal logics and the corresponding ADSs.

$\mathcal{C}(\mathcal{ALC}, \mathcal{MS})$: Suppose that you are developing a KR&R system for an estate agency. You imagine yourself to be a customer hunting for a house in London. What kind of requirements (constraints) could you have? Perhaps something like this:

(A) The house should not be too far from King's College, not more than 5 miles.

(B) The house should be close to shops, say, within 1 mile.

(C) There should be a 'green zone' around the house, at least within 2 miles in each direction.

(D) There must be a sports center around, and moreover, all sports centers of the district should be reachable on foot, i.e., they should be within, say, 3 miles.

(E) Public transport should easily be accessible: whenever you are not more than 8 miles away from home, the nearest bus stop or tube station should be reachable within 1 mile.

(F) The house should have a telephone.

(G) The neighbors shouldn't have children.

The terminology may require some clarification, so perhaps you may also need statements like

(H) All supermarkets are shops.

(Oh, you forgot about the most important constraint, the price, but let us deal with it a bit later.)

The resulting constraints (A)–(H) contain two kinds of knowledge. (F)–(H) can be classified as conceptual knowledge which is captured by almost any description logic, say, $\mathcal{ALC}$:

(F) $\mathit{house} : \exists \mathsf{has}.\mathsf{Telephone}$

(G) $\mathit{house} : \forall \mathsf{neighbor}.\forall \mathsf{child}.\bot$

(H) $\mathsf{Supermarket} \sqsubseteq \mathsf{Shop}$

(A)–(E) speak about distances and can be represented in the logic $\mathcal{MS}$ of metric spaces:

(A) $\mathit{house} \sqsubseteq \mathsf{E}_{\leq 5}\mathit{King's\_college}$

(B) $\mathit{house} \sqsubseteq \mathsf{E}_{\leq 1}\mathsf{Shop}$

(C) $\mathit{house} \sqsubseteq \mathsf{A}_{\leq 2}\mathsf{Green\_zone}$

(D) $\mathit{house} \sqsubseteq (\mathsf{E}_{\leq 3}\mathsf{Sports\_center}) \sqcap (\mathsf{A}_{>3}\neg\mathsf{Sports\_center})$

(E) $\mathit{house} \sqsubseteq \mathsf{A}_{\leq 8}\mathsf{E}_{\leq 1}\mathsf{Public\_transport}$

(Here, *house* and *King's_college* are nominals of $\mathcal{MS}$, while Shop, Green_zone, etc. are set variables.)

However, we can't just join these two knowledge bases together without connecting them. They speak about the same things, but from different points of view. For instance, in (H) 'shop' is used as a *concept*, while (B) deals with the *space* occupied by shops. Without connecting these different aspects we can't deduce from the knowledge base that a supermarket within 1 mile is sufficient to satisfy constraint (B). The required interaction can be easily captured by an E-connection between $\mathcal{ALC}$ and $\mathcal{MS}$. Indeed, take roles has, $\mathsf{neighbor}$, child, concepts $\mathsf{Telephone}$, $\mathsf{Supermarket}$, Shop, $\mathsf{Green\_zone}$ etc., and a nominal *King's_college* of $\mathcal{MS}$. Now, using the constructors $\mathcal{E}^2$ and $\mathcal{E}^1$ connecting $\mathcal{ALC}$- and $\mathcal{MS}$-models, we can represent constraints (A)–(H) as the concept $\mathsf{Good\_house}$ defined by the following knowledge base in $\mathcal{C}(\mathcal{ALC}, \mathcal{MS})$:

$$
\begin{aligned}
\mathsf{Good\_house} &= \mathsf{House} \sqcap \mathsf{Well\_located} \sqcap \exists \mathsf{has}.\mathsf{Telephone} \sqcap \\
&\quad\ \forall \mathsf{neighbor}.\forall \mathsf{child}.\bot \\
\mathsf{Well\_located} &= \mathcal{E}^1 \Big(\mathsf{E}_{\leq 5} \mathit{King's\_college} \sqcap \mathsf{A}_{\leq 2} \mathcal{E}^2(\mathsf{Green\_zone}) \\
&\quad \sqcap \mathsf{E}_{\leq 1} \mathcal{E}^2(\mathsf{Shop}) \sqcap \mathsf{E}_{\leq 3} \mathcal{E}^2(\mathsf{Sports\_center}) \\
&\quad \sqcap \mathsf{A}_{>3} \mathcal{E}^2(\neg \mathsf{Sports\_center}) \\
&\quad \sqcap \mathsf{A}_{\leq 8} \mathsf{E}_{\leq 1} \mathcal{E}^2(\mathsf{Public\_transport}) \Big) \\
\mathsf{Supermarket} &\sqsubseteq \mathsf{Shop}
\end{aligned}
$$

If we want to specify also that the house should be of reasonable price, the ADS $\mathcal{ALC}^\sharp$ can be extended with a suitable concrete domain dealing with natural numbers so that the resulting ADS is still decidable [5, 4]. The E-connection will be decidable as well.

By using a satisfiability checking algorithm for satisfiability in arbitrary metric spaces we can verify only that the requirements are consistent. To answer the query whether such a house really exists in London, we need a suitable map of London as our metric space and the agency's knowledge base about properties.

$\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$: Now imagine that you are employed by the EU parliament to develop a geographical information system about Europe. One part of the task is easy. You take the description logic $\mathcal{ALCO}$ (extending $\mathcal{ALC}$ with nominals, i.e., concept names which have to be interpreted as singleton sets [31, 23]) and, using concepts $\mathsf{Country}$, Treaty, etc., nominals EU, $\mathsf{Schengen\_treaty}$, object names *Spain*, *Luxembourg*, *UK*, etc., and a role member, write

$$
\begin{aligned}
&\mathit{Luxembourg} : \exists \mathsf{member}.\mathsf{EU} \sqcap \exists \mathsf{member}.\mathsf{Schengen\_treaty} \\
&\mathit{Iceland} : \exists \mathsf{member}.\mathsf{Schengen\_treaty} \sqcap \neg \exists \mathsf{member}.\mathsf{EU} \\
&\mathit{France} : \mathsf{Country} \\
&\mathsf{Schengen\_treaty} \sqsubseteq \mathsf{Treaty} \\
&\exists \mathsf{member}.\mathsf{Schengen\_treaty} \sqsubseteq \mathsf{Country}, \quad \text{etc.}
\end{aligned}
$$

After that you have to say something about the geography of Europe. To this end you can use the spatial logic RCC-8 [29, 12, 30, 37] or the more expressive formalism of $\mathsf{S4}_u$ in which one can encode the topological meaning of the RCC-8 predicates—DC (disconnected), EQ (equal), EC (externally connected), NTPP (non-tangential proper part), etc.—as term assertions of $\mathsf{S4}_u$ (see e.g., [12, 37]), say,

$$\begin{aligned}
&\mathsf{DC}(X,Y) \text{ as } \neg\Diamond(X\wedge Y),\\
&\mathsf{EQ}(X,Y) \text{ as } \Box(X\leftrightarrow Y),\\
&\mathsf{EC}(X,Y) \text{ as } \Diamond(X\wedge Y)\wedge\neg\Diamond(\boldsymbol{I}X\wedge\boldsymbol{I}Y),\\
&\mathsf{NTPP}(X,Y) \text{ as } \Box(\neg X\vee\boldsymbol{I}Y)\wedge\Diamond(\neg X\wedge Y).
\end{aligned}$$

Using an E-connection between $\mathcal{ALCO}$ and $\mathsf{S4}_u$ you can then continue:

$$\begin{aligned}
&\mathsf{EQ}(\mathcal{E}^2(\mathsf{EU}),\mathcal{E}^2(\mathit{Portugal})\sqcup\dots)\\
&\mathsf{EC}(\mathcal{E}^2(\mathit{France}),\mathcal{E}^2(\mathit{Luxembourg}))\\
&\mathsf{NTPP}(\mathcal{E}^2(\mathit{Luxembourg}),\mathcal{E}^2(\exists\mathsf{member.Schengen_Treaty}))\\
&\mathit{Austria}:\mathcal{E}^1(\mathit{alps})
\end{aligned}$$

i.e., 'the space occupied by the EU is the space occupied by its members,' 'France and Luxembourg have a common border,' 'if you cross the border of Luxembourg, then you enter a member of the Schengen Treaty,' 'Austria is an alpine country' (*alps* is a set variable of $\mathsf{S4}_u$). Of course, to ensure that the spatial extensions of the EU, *France*, etc. are not degenerate and to comply with requirements of RCC-8, you should guarantee that all mentioned spatial regions are interpreted by regular closed sets, i.e., $\mathcal{E}^2(\mathsf{EU})=\boldsymbol{CI}\mathcal{E}^2(\mathsf{EU})$, $\mathcal{E}^2(\mathit{France})=\boldsymbol{CI}\mathcal{E}^2(\mathit{France})$, etc.

Suppose now that you want to test your system and ask whether France is a member of the Schengen treaty, i.e., $\mathit{France}\sqsubseteq\exists\mathsf{member.Schengen\_treaty}$. The answer will be "Don't know!" because you did not tell your system that the spatial extensions of any two countries do not overlap. If you add, for example,

$$\boldsymbol{I}\mathcal{E}^2(\mathsf{Country}\sqcap\neg\exists\mathsf{member.Schengen\_treaty})\ \sqsubseteq\ \neg\boldsymbol{I}(\mathcal{E}^2(\exists\mathsf{member.Schengen\_treaty}))$$

('the members of the Schengen treaty do not overlap with the non-Schengen countries') to the knowledge base, then the answer to the query will be "Yes!"

$\mathcal{C}(\mathcal{SHIQ},\mathcal{ALCO})$: Having satisfied your boss in the EU parliament with the constructed GIS, you get a new task: develop a knowledge base regulating relations between people in the EU (citizenship, jobs, etc.). On the one hand, you already have the $\mathcal{ALCO}$ knowledge base describing countries in the EU from the previous example. But on the other hand, you must also be able to express laws like (i) 'no citizen of the EU may have more than one spouse,' (ii) 'all children of UK citizens are UK citizens,' or (iii) 'a person living in the UK is either child of somebody who is a UK citizen or has a work permit in the UK, or the person is a UK citizen or has a work permit in the UK herself.' This means, in particular, that you need more constructors than are provided by $\mathcal{ALCO}$, say, qualified number restrictions and inverse roles. It is known, however, that inverse roles, number restrictions, and nominals are difficult to handle algorithmically in one system [23]. The fusion of $\mathcal{ALCO}$ with, say, $\mathcal{SHIQ}$ of [22], having the required constructors, doesn't help either because transfer results for fusions are available so far only for DLs whose models are closed under disjoint unions [5, 4]—which is not the case if nominals are allowed as concept constructors. It seems that a perspective way to attack this problem is to *connect* $\mathcal{SHIQ}$ with $\mathcal{ALCO}$.

Let E contain three binary relations between the domains of $\mathcal{SHIQ}$ (people, companies, etc.) and $\mathcal{ALCO}$ (countries): $x\mathcal{S}y$ means that x is a citizen of y, $x\mathcal{L}y$ means that x lives in y, and $x\mathcal{W}y$ that x has a work permit in y. For example, $\mathcal{L}^1(UK)$ denotes all people living in the UK, while $\mathcal{S}^1(UK)$ all UK citizens. The subsumptions below represent the regulations (i)–(iii):

$$\begin{aligned}\mathcal{S}^1(\mathsf{EU}) &\sqsubseteq \neg\exists_{\geq 2}\mathsf{married}.\top\\ \exists\mathsf{child\_of}.\mathcal{S}^1(UK) &\sqsubseteq \mathcal{S}^1(UK)\\ \mathcal{L}^1(UK) &\sqsubseteq \exists\mathsf{child\_of}^{-1}.(\mathcal{S}^1(UK)\sqcup\mathcal{W}^1(UK))\\ &\qquad\sqcup\,\mathcal{S}^1(UK)\sqcup\mathcal{W}^1(UK)\end{aligned}$$

$\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u, \mathsf{PTL})$ "The EU is developing," said your boss, "we are going to have new members by 2005." So you extend the connection $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$ with one more ADS—propositional temporal logic PTL. Now, besides nominals EU, etc. and object names *Germany*, etc. of $\mathcal{ALCO}$ and set variables *alps*, *Basel*, etc. of $\mathsf{S4}_u$, we have *nominals* $0, 1, \dots$ interpreting time points (n can be regarded as an abbreviation for $\neg\bigcirc_P^n\top\wedge\bigcirc_P^{n-1}\top$, where $\bigcirc_P\varphi$ stands for $\bot\mathcal{S}\varphi$). The ternary relation $\mathcal{E}(x, y, z)$ means now that at moment z (from the domain of PTL) point y (in the domain of $\mathsf{S4}_u$) belongs to the spatial region occupied by object x (in the domain of $\mathcal{ALCO}$). Then we can say, for example:

$$\begin{aligned}&\mathcal{E}^2(Poland, 2005)\sqsubseteq\mathcal{E}^2(\mathsf{EU}, 2005)\\ &\mathsf{PO}(\mathcal{E}^2(Austria, 1914), \mathcal{E}^2(Italy, 1950))\\ &\Box_F\neg\mathcal{E}^3(Basel, \mathsf{EU}),\end{aligned}$$

i.e., 'in 2005, the territory of Poland will belong to the territory occupied by the EU,' 'the territory of Austria in 1914 partially overlaps the territory of Italy in 1950,' 'no part of Basel will ever belong to the EU.'

5 Proof

We prove Theorem 6 for the connection $\mathcal{C} = \mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$ between *two* ADSs $\mathcal{S}_i = (\mathcal{L}_i, \mathcal{M}_i)$, for $i = 1, 2$, by *one* relation E. The extra function symbols $\mathcal{E}^1$ and $\mathcal{E}^2$ of the connection are then unary. The general case of connections between n ADSs is treated analogously.

Let Γ be a finite set of assertion of $\mathcal{C}$. Denote by $ob_i(\Gamma)$ the set of object names from $\mathcal{L}_i$ which occur in Γ, $i = 1, 2$. Let $\overline{1} = 0$, $\overline{0} = 1$, and

$$o_i(\Gamma) = \{\mathcal{E}^i\neg\mathcal{E}^{\overline{i}}a \mid a \in ob_i(\Gamma)\}.$$

Note that $a^{\mathfrak{M}} \notin (\mathcal{E}^i\neg\mathcal{E}^{\overline{i}}a)^{\mathfrak{M}}$, for any model $\mathfrak{M}$ for $\mathcal{C}$ and object name a of $\mathcal{L}_i$. For every i-term t of the form $\mathcal{E}^i s$, where s is a term or an object name of $\mathcal{L}_{\overline{i}}$ occurring in $\Gamma' = \Gamma \cup o_1(\Gamma) \cup o_2(\Gamma)$, we introduce a fresh set variable x_t of $\mathcal{L}_i$. Given an i-term t, denote by $sur_i(t)$—the *surrogate of* t—the term which results from t by replacing all subterms $t' = \mathcal{E}^i s$ of t that are not within the scope of an $\mathcal{E}^i$ by $x_{t'}$. Clearly, $sur_i(t)$ belongs to the language $\mathcal{L}_i$.

Denote by Θ_i, $i = 1, 2$, the closure under negation of the set of i-terms which occur in Γ'. Without loss of generality we can identify $\neg\neg t$ with t. Thus, Θ_i is finite. The *i-consistency-set* $\mathcal{C}(\Theta_i)$ is defined as the set $\{t_c \mid c \subseteq \Theta_i\}$, where

$$t_c = \bigwedge\{\chi \mid \chi \in c\} \wedge \bigwedge\{\neg\chi \mid \chi \in \Theta_i \setminus c\}.$$

Sometimes we will identify $t \in \mathcal{C}(\Theta_i)$ with the set of its conjuncts; then $s \in t$ means that s is a conjunct of t. By $\top_i$ we denote $x^i \vee \neg x^i$, where x^i is a set variable from $\mathcal{L}_i$.

We are now ready to reduce the satisfiability in the connection $\mathcal{C}$ to the satisfiability problem in the components $(\mathcal{L}_1, \mathcal{M}_1)$ and $(\mathcal{L}_2, \mathcal{M}_2)$:

Theorem 9. *Γ is satisfiable iff there exist sets $\Delta_1 \subseteq \mathcal{C}(\Theta_1)$ and $\Delta_2 \subseteq \mathcal{C}(\Theta_2)$, a relation $e \subseteq \Delta_1 \times \Delta_2$, and for each $t \in \Delta_i$, $i = 1, 2$, a fresh object name a_t from $\mathcal{L}_i$ for which the following hold: there are functions σ_i from $ob_i(\Gamma)$ into Δ_i such that $\mathcal{E}^i\neg\mathcal{E}^{\bar{i}}a \notin \sigma_i(a)$, for any $a \in ob_i(\Gamma)$, the union Γ_i of the sets*

$$\{sur_i(\bigvee \Delta_i) = \top_i\};$$
$$\{a_t : sur_i(t) \mid t \in \Delta_i\};$$
$$\{a : sur_i(\sigma_i(a)) \mid a \in ob_i(\Gamma)\};$$
$$\{(a : sur_i(t)) \mid (a : t) \in \Gamma \text{ an } i\text{-term assertion}\};$$
$$\{sur_i(t_1) \sqsubseteq sur_i(t_2) \mid t_1 \sqsubseteq t_2 \in \Gamma \text{ an } i\text{-term assertion}\};$$
$$\{R_j(a_1, \ldots, a_{m_j}) \in \Gamma \mid R_j(a_1, \ldots, a_{m_j}) \text{ an } i\text{-object assertion}\};$$

is $(\mathcal{L}_i, \mathcal{M}_i)$-satisfiable, and

1. *for all $\mathcal{E}^1 s \in \Theta_1$, s a 2-term, and $t \in \Delta_1$, we have $\mathcal{E}^1 s \in t$ iff there is $t' \in \Delta_2$ with $(t, t') \in e$ and $s \in t'$,*
2. *for all $\mathcal{E}^1 a \in \Theta_1$, $a \in ob_2(\Gamma)$, and $t \in \Delta_1$, we have $\mathcal{E}^1 a \in t$ iff $(t, \sigma_2(a)) \in e$,*
3. *for all $\mathcal{E}^2 s \in \Theta_2$, s a 1-term, and $t \in \Delta_2$, we have $\mathcal{E}^2 s \in t$ iff there is $t' \in \Delta_1$ with $(t', t) \in e$ and $s \in t'$,*
4. *for all $\mathcal{E}^2 a \in \Theta_2$, $a \in ob_2(\Gamma)$, and $t \in \Delta_2$, we have $\mathcal{E}^2 a \in t$ iff $(\sigma_1(a), t) \in e$.*

Proof. ($\Rightarrow$) Suppose Γ is satisfiable in $\mathcal{C}$. Take a model $\mathfrak{M} = ((\mathfrak{W}_1, \mathfrak{W}_2), E)$ which satisfies Γ. Let, for $d \in W_i$,

$$t(d) = \bigwedge\{t \in \Theta_i \mid d \in t^{\mathfrak{M}}\}$$

and let $\Delta_i = \{t(d) | d \in W_i\}$. Take a fresh object name a_t from $\mathcal{L}_i$ for every $t \in \Delta_i$. Define, for $a \in ob_i(\Gamma)$,

$$\sigma_i(a) = \bigwedge\{t \in \Theta_i \mid a^{\mathfrak{M}} \in t^{\mathfrak{M}}\} \in \Delta_i$$

and define $e \subseteq \Delta_1 \times \Delta_2$ by taking $(t, t') \in e$ iff there are $d_1 \in W_1$ and $d_2 \in W_2$ such that $t = t(d_1)$, $t' = t(d_2)$, and $d_1 E d_2$. It remains to check that Δ_i, σ_i and e are as required.

First, we show that Γ_i is $(\mathcal{L}_i, \mathcal{M}_i)$-satisfiable. Take the model

$$\mathfrak{W}_i = \left\langle W_i, \mathcal{V}_i^{\mathfrak{W}_i}, \mathcal{X}_i^{\mathfrak{W}_i}, R_i^{\mathfrak{W}_i}, F_i^{\mathfrak{W}_i} \right\rangle.$$

$\mathfrak{W}_i$ is almost as required. We just have to give the appropriate values to the fresh set variables x_t and the fresh object names a_t. To this end put $x_t^{\mathfrak{W}'_i} = t^{\mathfrak{M}}$ for every $t = \mathcal{E}_i s$, $x^{\mathfrak{W}'_i} = x^{\mathfrak{W}_i}$ for the remaining variables, and $a_t^{\mathfrak{W}'_i} \in t^{\mathfrak{M}}$ for every $t \in \Delta_i$ and $a^{\mathfrak{W}'_i} = a^{\mathfrak{W}_i}$ for the remaining object names. Note that

$$\mathfrak{W}' = \left\langle W_i, \mathcal{V}_i^{\mathfrak{W}'_i}, \mathcal{X}_i^{\mathfrak{W}'_i}, R_i^{\mathfrak{W}_i}, F_i^{\mathfrak{W}_i} \right\rangle \in \mathcal{M}_i.$$

This follows from the closure conditions for the class $\mathcal{M}_i$. To prove that $\mathfrak{W}'_i \models \Gamma_i$, it is enough to show by induction that for all $d \in W_i$ and $s \in \Theta_i$, we have $d \in (sur_i(s))^{\mathfrak{W}'_i}$ iff $d \in s^{\mathfrak{M}}$. We leave this to the reader. Thus, the sets Γ_i are $(\mathcal{L}_i, \mathcal{M}_i)$-satisfiable.

Now we check that e satisfies conditions (1)–(4).

(1) Suppose $\mathcal{E}^1 s \in \Theta_1$, s is a 2-term and $t \in \Delta_1$. Assume first that $\mathcal{E}^1 s \in t$. Let $t = t(d)$ for some $d \in W_1$. Then there is $d' \in W_2$ with dEd' and $d' \in s^{\mathfrak{M}}$. But then $s \in t(d')$ and $(t, t(d')) \in e$. Assume now that $(t, t') \in e$ and $s \in t'$. Then there are $d \in W_1$ and $d' \in W_2$ with $t = t(d)$ and $t' = t(d')$ and dEd'. We have $d' \in s^{\mathfrak{W}}$ and so $d \in (\mathcal{E}^1 s)^{\mathfrak{M}}$. Hence $\mathcal{E}^1 s \in t$.

(3) is proved in the same manner.

(2) Suppose $\mathcal{E}^1 a \in \Theta_1$ and $t \in \Delta_1$. Assume first that $\mathcal{E}^1 a \in t$. Then, for any $d \in W_1$ with $t(d) = t$ we have $dEa^{\mathfrak{M}}$. Hence $(t, t(a^{\mathfrak{M}})) \in e$ and so $(t, \sigma_2(a)) \in e$. Conversely, let $(t, \sigma_2(a)) \in e$. Then $\mathcal{E}^2 \neg \mathcal{E}^1 a \notin \sigma_2(a)$. By (3), $\neg \mathcal{E}^1 a \notin t$. Hence $\mathcal{E}^1 a \in t$.

(4) is proved in the same manner as (2).

($\Leftarrow$) Suppose that Δ_i, σ_i and e satisfying the conditions of the theorem are given. We construct a model satisfying Γ. To this end take $\mathfrak{W}_i \in \mathcal{M}_i$ satisfying Γ_i, $i = 1, 2$. Let, for $d \in W_i$,

$$t(d) = \bigwedge \{ t \in \Theta_i \mid d \in (sur_i(t))^{\mathfrak{W}_i} \}.$$

Now define $E \subseteq W_1 \times W_2$ by taking dEd' iff $(t(d), t(d')) \in e$. We show that $\mathfrak{M} = (\mathfrak{W}_1, \mathfrak{W}_2, E)$ satisfies Γ. To this end we show by simultaneous induction for $i = 1, 2$ and all $d \in W_i$ and $s \in \Theta_i$, that $d \in (sur_i(s))^{\mathfrak{W}_i}$ iff $d \in s^{\mathfrak{M}}$. For set-variables the claim follows from the definition. Also the steps for the Boolean operators and for the function symbols of $\mathcal{L}_i$, $i = 1, 2$, are clear. It remains to consider the cases $t = \mathcal{E}^i s$, $i = 1, 2$. Let us assume $i = 1$, the case $i = 2$ is similar.

Suppose first that $t = \mathcal{E}^1 s$ for a 2-term s and that $d \in (sur_1(\mathcal{E}^1 s))^{\mathfrak{W}_1}$. Then $\mathcal{E}^1 s \in t(d)$. We know that

$$\mathfrak{W}_1 \models sur_1(\bigvee \Delta_1) = \top_1$$

and so $t(d) \in \Delta_1$. By (1), we find $t' \in \Delta_2$ with $(t(d), t') \in e$ and $s \in t'$. We know that

$$\mathfrak{W}_2 \models a_{t'} : sur_2(t'),$$

and so we find $d' \in W_2$ with $t' = t(d')$. Hence dEd'. From $s \in t'$ we obtain $d' \in (sur_2(s))^{\mathfrak{M}_2}$ and by the induction hypothesis $d' \in s^{\mathfrak{M}}$. Now $d \in (\mathcal{E}^1 s)^{\mathfrak{M}}$ follows. Conversely, suppose $d \in (\mathcal{E}^1 s)^{\mathfrak{M}}$. We find $d' \in W_2$ with dEd' and $d' \in s^{\mathfrak{M}}$. By the induction hypothesis, $d' \in (sur_2(s))^{\mathfrak{W}_2}$ and so $s \in t(d')$. By definition $(t(d), t(d')) \in e$ and so, by (1), $\mathcal{E}^1 s \in t(d)$ which implies $d \in (sur_1(\mathcal{E}^1 s))^{\mathfrak{W}_1}$.

Suppose now that $t = \mathcal{E}^1 a$ for an object name a of $\mathcal{L}_2$. Since $d \in (sur_1(\mathcal{E}^1 a))^{\mathfrak{W}_1}$, we have $\mathcal{E}^1 a \in t(d)$. As above we know that $t(d) \in \Delta_1$. By (2) $(t, \sigma_2(a)) \in e$. We also know that

$$\mathfrak{W}_2 \models a : sur_2(\sigma_2(a)).$$

Hence $dEa^{\mathfrak{W}_2}$, which implies $d \in (\mathcal{E}^1 a)^{\mathfrak{M}}$. Conversely, suppose that $d \in (\mathcal{E}^1 a)^{\mathfrak{M}}$. Then $dEa^{\mathfrak{M}}$, and so $(t(d), t(a^{\mathfrak{M}})) \in e$. We have $t(a^{\mathfrak{M}}) = \sigma_2(a)$ and therefore $(t(d), \sigma_2(a)) \in e$, which implies, by (2), that $\mathcal{E}^1 a \in t(d)$. This means $d \in (sur_1(\mathcal{E}^1 a))^{\mathfrak{W}_1}$. ❑

Theorem 6 follows from Theorem 9. Indeed, since the sets $\mathcal{C}(\Theta_i)$ are finite, Theorem 9 provides us with a decision procedure for $\mathcal{C}$ if decision procedures for $(\mathcal{L}_i, \mathcal{M}_i)$, $i = 1, 2$, are known. To decide whether a set Γ is satisfiable, 'guess' sets $\Delta_1 \subseteq \mathcal{C}(\Theta_1)$ and $\Delta_2 \subseteq \mathcal{C}(\Theta_2)$, functions $\sigma_i : ob_i(\Gamma) \to \Delta_i$, $i = 1, 2$, and a relation $e \subseteq \Delta_1 \times \Delta_2$ and check whether they satisfy the conditions listed in the formulation of the theorem. Regarding the complexity of the obtained decision procedure, the costly step is guessing the right sets Δ_i. The cardinality of the sets $\mathcal{C}(\Theta_i)$ is exponential in the size of Γ. Thus, there are double exponentially many different subsets to be chosen from. Since the cardinality of the chosen sets Δ_i may be exponential in the size of Γ, also the size of Γ_1 and Γ_2 may be exponential in Γ (because of the big disjunction over the Δ_i).

6 A-Satisfiability

Here we show that A-satisfiability is not preserved under connections. To this end, recall that $\mathcal{ALCF}$ is the extension of $\mathcal{ALC}$ by functional roles and the same-as constructor (see the introduction). A-satisfiability for $\mathcal{ALCF}$ is decidable [27], while satisfiability itself is not [3].

Proof. of Theorem 8. Let $(\mathcal{L}, \mathcal{M})$ be an arbitrary ADS. We show that for any finite set Γ of general TBox-axioms and any assertion $a : t$ of $\mathcal{ALCF}$, there exists a set Γ^* of object assertions of $\mathcal{C} = \mathcal{C}(\mathcal{ALCF}, (\mathcal{L}, \mathcal{M}))$ such that $\Gamma \cup \{a : t\}$ is satisfiable iff Γ^* is satisfiable in the connection. Suppose Γ and $a : t$ are given. We may assume that Γ consists of axioms of the form $u = \top$. Denote by $\mathcal{R}$ the set of all roles which occur in $\Gamma \cup \{t\}$. We put $\mathcal{A}^i s = \neg\mathcal{E}^i \neg s$. Let b be an object name of $(\mathcal{L}, \mathcal{M})$.

Define Γ^* as the union of the sets

$$\{a : \mathcal{E}^1(b)\} \cup \{b : \mathcal{A}^2 f_{\forall R}\mathcal{E}^1(b) \mid R \in \mathcal{R}\},$$
$$\{a : t\} \cup \{b : \mathcal{A}^2 u \mid (u = \top) \in \Gamma\}.$$

Suppose $\Gamma \cup \{a : t\}$ is satisfied in an $\mathcal{ALCF}$-model $\mathfrak{W}_1$ with domain Δ. Define a model $\mathfrak{M}$ for $\mathcal{C}$ by taking an arbitrary model $\mathfrak{W}_2$ with domain W for $(\mathcal{L}, \mathcal{M})$ and putting $E = \Delta \times W$. It is easily checked that $\mathfrak{M} \models \Gamma^*$.

Conversely, let us suppose that $\mathfrak{M} \models \Gamma^*$ for a $\mathcal{C}$-model $\mathfrak{M} = (\mathfrak{W}_1, \mathfrak{W}_2, E)$. Let Δ be the domain of $\mathfrak{W}_1$. Denote by Δ' the minimal subset of Δ containing $a^{\mathfrak{M}}$ and satisfying the following closure condition for all $d, d' \in \Delta$:

$$\text{if } d \in \Delta' \text{ and } \exists S \in \mathcal{R}\ dS^{\mathfrak{M}}d' \text{ then } d' \in \Delta'.$$

The model $\mathfrak{W}_1'$, defined as the substructure of $\mathfrak{W}_1$ induced by Δ', satisfies $\Gamma \cup \{a : t\}$. To see this, it is sufficient to show that $u^{\mathfrak{M}} \supseteq \Delta'$ for every term u with $u = \top \in \Gamma$. Note that $d \in u^{\mathfrak{M}}$ whenever $(d, b^{\mathfrak{M}}) \in E$, because $b : \mathcal{A}^2 u \in \Gamma^*$. Hence it is enough to prove that for all $d \in \Delta'$, $(d, b^{\mathfrak{M}}) \in E$.

Since $a : \mathcal{E}^1(b) \in \Gamma^*$, $(a^{\mathfrak{M}}, b^{\mathfrak{M}}) \in E$. Suppose $d \in \Delta'$, dSd' for some $S \in \mathcal{R}$, and $(d, b^{\mathfrak{M}}) \in E$. We have $b \in \mathcal{A}^2 f_{\forall R}\mathcal{E}^1(b)$, and so $d \in f_{\forall R}\mathcal{E}^1(b)$. This implies $d' \in \mathcal{E}^1(b)$, from which $(d', b^{\mathfrak{M}}) \in E$. ❑

Note that in the proof we only used the following properties of $\mathcal{ALCF}$: (1) ABox-reasoning (without general TBox-axioms) is decidable, (2) reasoning with general TBoxes is undecidable, (3) an $\mathcal{ALCF}$-concept applies to an object d iff it applies to d in the generated substructure based on Δ' as defined above. Condition (3) applies to all standard description logics. This means that, roughly speaking, the A-satisfiability problem for an E-connection of two ADS is undecidable, whenever at least one ADS has an undecidable satisfiability problem.

7 Undefinability

Given that the E-connection of any finite number of decidable ADSs is decidable as well, it is clear that the interaction between the components has to be rather limited. Yet, it is not obvious what exactly can and what can't be expressed in the combined language. We have gone into great depth to provide examples of potentially useful E-connections. This section is devoted to sheding some light on the question of expressivity.

As is well known, undefinability results in modal logic—such as the undefinability of the irreflexivity of a Kripke frame—are usually gained by the concept of *bisimulation*. In what is to follow, we will lift the concept of bisimulations to the case of E-connections and will then give some simple examples of undefinable properties of E-connections.

We will work with the following definition of *definability*:

Definition 10. Let $\mathcal{C}$ be an E-connection. A property $\mathcal{P}$ of models of $\mathcal{C}$ is definable in $\mathcal{C}$ iff there exists a finite set Γ of assertions of $\mathcal{C}$ such that for all models $\mathfrak{M}$ of $\mathcal{C}$ the following holds: $\mathfrak{M}$ has $\mathcal{P}$ iff $\mathfrak{M} \models \Gamma$.

Bisimulations for ADSs

As ADSs abstract from the concrete definition of a given logic, it is difficult to come up with a notion of bisimulation that is non-trivial in the sense that it reflects certain properties of the logics under investigation (as is the case with bisimulations for modal logics where the semantic definition of modal operators is reflected in the definition of bisimulations).

So let us use a rather straightforward definition of bisimulation that simply pins down exactly what is needed to ensure that two models are indistinguishable.[3]

Definition 11. Let $(\mathcal{L}, \mathcal{M})$ be an ADS, and let

$$\mathfrak{W}_k = \left\langle W_k, \mathcal{V}^{\mathfrak{W}_k}, \mathcal{X}^{\mathfrak{W}_k}, F^{\mathfrak{W}_k}, R^{\mathfrak{W}_k} \right\rangle,$$

$k = 1, 2$, be two abstract description models from the class $\mathcal{M}$.

We say that $\mathfrak{W}_1$ and $\mathfrak{W}_2$ are *locally bisimilar*, in symbols $\mathfrak{W}_1 \rightleftharpoons \mathfrak{W}_2$, if there exists a non-empty binary relation $\rightleftharpoons \subseteq W_1 \times W_2$, such that the following holds:

(a) For all object variables $a \in \mathcal{L}$, $a^{\mathfrak{W}_1} \rightleftharpoons a^{\mathfrak{W}_2}$ and $R(a_1^{\mathfrak{W}_1}, \ldots, a_n^{\mathfrak{W}_1})$ iff $R(a_1^{\mathfrak{W}_2}, \ldots, a_n^{\mathfrak{W}_2})$;

(b) If $u \rightleftharpoons v$, then $u \in x^{\mathfrak{W}_1}$ iff $v \in x^{\mathfrak{W}_2}$;

[3]If one is interested in particular classes of logics, this definition of bisimulation can be accordingly strengthened.

(c) If $u \rightleftharpoons v$, then $u \in f_i^{\mathfrak{W}_1}(t^{\mathfrak{W}_1})$ iff $v \in f_i^{\mathfrak{W}_2}(t^{\mathfrak{W}_2})$.

Further, we say that $\mathfrak{W}_1$ and $\mathfrak{W}_2$ are *globally bisimilar*, in symbols $\mathfrak{W}_1 \rightleftharpoons_G \mathfrak{W}_2$, if they are locally bisimilar and the relation $\rightleftharpoons \subseteq W_1 \times W_2$ is global in the sense that for all $u \in W_1$ there is some $v \in W_2$ such that $u \rightleftharpoons v$, and, conversely, for all $v \in W_2$ there is some $u \in W_1$ such that $u \rightleftharpoons v$.

Proposition 12. *Let $\mathcal{S} = (\mathcal{L}, \mathcal{M})$ be an ADS and $\mathfrak{W}_k$, $k = 1, 2$, any two abstract description models from the class $\mathcal{M}$ that are locally bisimilar, $\mathfrak{W}_1 \rightleftharpoons \mathfrak{W}_2$. Then:*

(i) For all terms t of $\mathcal{L}$ and all points $u \in W_1$, $v \in W_2$ such that $u \rightleftharpoons v$ it holds that $u \in t^{\mathfrak{W}_1}$ iff $v \in t^{\mathfrak{W}_2}$;

(ii) $\mathfrak{W}_1 \models \phi$ if and only if $\mathfrak{W}_2 \models \phi$, for all object assertions ϕ of $\mathcal{L}$;

(iii) If $\mathfrak{W}_1 \rightleftharpoons_G \mathfrak{W}_2$, then $\mathfrak{W}_1 \models t_1 \sqsubseteq t_2$ if and only if $\mathfrak{W}_2 \models t_1 \sqsubseteq t_2$, for all term assertions $t_1 \sqsubseteq t_2$.

Proof. Claim (i) follows directly from items (b) and (c) of Definition 11 (with the Boolean connectives being a trivial inductive step) and (ii) follows from (i) and item (a).

For the proof of the third claim, suppose that $\rightleftharpoons$ is global. Assume that $\mathfrak{W}_1 \models t_1 \sqsubseteq t_2$. If $t_1^{\mathfrak{W}_2} = \emptyset$, then $\mathfrak{W}_2 \models t_1 \sqsubseteq t_2$ follows. So assume otherwise. Take any $v \in t_1^{\mathfrak{W}_2}$. By globality, there is a $u \in W_1$ such that $u \rightleftharpoons v$. Hence $u \in t_1^{\mathfrak{W}_1}$ by (i). By assumption, $u \in t_2^{\mathfrak{W}_1}$ as well, whence $v \in t_2^{\mathfrak{W}_2}$ by (i). ❑

Bisimulations for E-connections

We will now extend the notion of bisimulation for ADSs to E-connections.

Definition 13. Let $\mathcal{S}_i = (\mathcal{L}_i, \mathcal{M}_i)$, $i = 1, \dots, n$, be n ADSs and let $\mathcal{C} = \mathcal{C}(\mathcal{S}_1, \dots, \mathcal{S}_n)$ be the corresponding E-connection. Two models $\mathfrak{M}_1 = \left\langle (\mathfrak{W}^1{}_i)_{i \le n}, (E_j^1)_{j \le m} \right\rangle$ and $\mathfrak{M}^2 = \left\langle (\mathfrak{W}^2{}_i)_{i \le n}, (E_j^2)_{j \le m} \right\rangle$ for $\mathcal{C}$ are called *E-bisimilar*, symbolically $\mathfrak{M}_1 \rightleftharpoons_E \mathfrak{M}_2$, if there are global relations $\rightleftharpoons_i$ $(i = 1, \dots, n)$ satisfying conditions (a)–(c) from definition 11 such that the following holds for any j:

(d) If $\langle u_1, \dots, u_n \rangle \in E_j^1$ and $u_i \rightleftharpoons_i v_i$ for some i, then there are $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ such that $\langle v_1, \dots, v_n \rangle \in E_j^2$ and $u_k \rightleftharpoons_k v_k$ for all k.

(e) If $\langle v_1, \dots, v_n \rangle \in E_j^2$ and $u_i \rightleftharpoons_i v_i$ for some i, then there are $u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n$ such that $\langle u_1, \dots, u_n \rangle \in E_j^1$ and $u_k \rightleftharpoons_k v_k$ for all k.

Proposition 14. *Let $\mathcal{C} = \mathcal{C}(\mathcal{S}_1, \dots, \mathcal{S}_n)$ be an E-connection and suppose that $\mathfrak{M}^1 \rightleftharpoons_E \mathfrak{M}^2$. Then for all assertions ϕ of $\mathcal{C}$ it holds that:*

$$\mathfrak{M}^1 \models \phi \text{ iff } \mathfrak{M}^2 \models \phi,$$

i.e., E-bisimilar models are indistinguishable by means of assertions.

Proof. For simplicity, let us assume that the E-connection is 2-dimensional, i.e. $\mathcal{C} = \mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$, and contains only two new function symbols, $\mathcal{E}^1$ and $\mathcal{E}^2$. Suppose $\phi = \mathcal{E}^1(t)$ with t a 2-term. We prove that for all $u_1 \in W_1^1$ and $v_1 \in W_1^2$ with $u_1 \rightleftharpoons_1 v_1$

$$u_1 \in (\mathcal{E}^1(t))^{\mathfrak{M}^1} \text{ iff } v_1 \in (\mathcal{E}^1(t))^{\mathfrak{M}^2}$$

The case $\phi = \mathcal{E}^2(s)$ can be treated similarly. From this and the fact that the $\rightleftharpoons_i$ are global bisimulations, the claim follows.

So suppose that $u_1 \in (\mathcal{E}^1(t))^{\mathfrak{M}^1}$, i.e. that there is a $u_2 \in t^{\mathfrak{M}^1}$ such that $\langle u_1, u_2 \rangle \in E^1$. Since $u_1 \rightleftharpoons_1 v_1$, it follows that there is a v_2 such that $\langle v_1, v_2 \rangle \in E^2$ and $u_2 \rightleftharpoons_2 v_2$. Hence $v_2 \in t^{\mathfrak{M}^2}$, whence $v_1 \in (\mathcal{E}^1(t))^{\mathfrak{M}^2}$. ❑

Examples

Let us now give a few examples of undefinable properties of E-connections. For brevity, we will restrict the examples to the case of 2-dimensional E-connections.

Theorem 15. *Let R and S be role names.*
(i) The property

$$(\dagger) \quad \forall x \forall y \forall z (xRy \rightarrow (xEz \rightarrow yEz))$$

is not definable in the E-connections $\mathcal{C}(\mathcal{ALC}, \mathcal{MS})$, $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$ and $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$.
(ii) The property

$$(\ddagger) \quad \forall x \forall y (xRy \wedge xEx' \wedge yEy' \rightarrow x'Sy')$$

is not definable in the E-connection $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$.

Proof. Our strategy will be to give appropriate pairs of models for the respective E-connections, one model satisfying the given property, the other not, and provide a bisimulation between them. This shows that the properties (†) and (‡) are not definable.

Let us prove (i) and consider first the case of the E-connection $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$. We treat this case rather detailed and will be briefer in the remaining cases. To visualize the models we define below, compare Figure 1. Let $\mathfrak{M}_1 = \langle \mathfrak{W}_1, \mathfrak{W}_2, E \rangle$ be a model for $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$ with $\mathfrak{W}_i = \left\langle W_i, \mathcal{V}^{\mathfrak{W}_i}, \mathcal{X}^{\mathfrak{W}_i}, F^{\mathfrak{W}_i}, R^{\mathfrak{W}_i} \right\rangle$ where $W_1 = \{a_1, a_2, b_1, b_2, o_1\}$, $x^{\mathfrak{W}_1} = \emptyset$ for all $x \in \mathcal{V}$ and $a^{\mathfrak{W}_1} = o_1$ for all $a \in \mathcal{X}$. There is one (non-trivial) role name R with $R^{\mathfrak{W}_1} = \{\langle a_1, b_1 \rangle, \langle a_2, b_2 \rangle\}$[4] and function symbols $f_{\forall R}$, $f_{\exists R}$, $f_{\exists R^{-1}}$, $f_{\forall R^{-1}}$, $f_{\exists_{\leq n}}$, $f_{\forall_{\leq n}}$, $f_{\exists_{\geq n}}$ and $f_{\forall_{\geq n}}$ for $n \in \mathbb{N}$ which are interpreted as defined on page 5. Further, let $W_2 = \{c_1, c_2, o_2\}$, $x^{\mathfrak{W}_2} = \emptyset$ for all $x \in \mathcal{V}$ and $a^{\mathfrak{W}_2} = o_2$ for all $a \in \mathcal{X}$. We assume that in $\mathfrak{W}_2$ all roles are interpreted by the empty set. For every nominal o of $\mathcal{ALCO}$ we have a 0-ary function symbol f_o which we interpret as $f_o^{\mathfrak{W}_2} = \{o_2\}$. Finally let

$$E = \{\langle a_1, c_1 \rangle, \langle a_2, c_2 \rangle, \langle b_1, c_1 \rangle, \langle b_2, c_2 \rangle\}.$$

Next, let the model $\mathfrak{M}_2 = \langle \mathfrak{W}_1', \mathfrak{W}_2', E' \rangle$ with $W_1' = \{a_1', a_2', b_1', b_2', o_1'\}$ and $W_2' = \{c_1', c_2', o_2'\}$ be defined just like $\mathfrak{M}_1$ except for E' which is given by

$$E' = \{\langle a_1', c_2' \rangle, \langle a_2', c_1' \rangle, \langle b_1', c_1' \rangle, \langle b_2', c_2' \rangle\}.$$

[4]For other roles S assume $S^{\mathfrak{W}_1} = \emptyset$.

It should be obvious that $\mathfrak{M}_1$ satisfies (†), while $\mathfrak{M}_2$ doesn't. We claim that the relation $\rightleftharpoons$ defined by

$$
\begin{aligned}
\rightleftharpoons \; := \; & \{\left\langle a_1, a_1' \right\rangle, \left\langle a_1, a_2' \right\rangle, \left\langle a_2, a_1' \right\rangle, \left\langle a_2, a_2' \right\rangle, \\
& \left\langle b_1, b_1' \right\rangle, \left\langle b_1, b_2' \right\rangle, \left\langle b_2, b_1' \right\rangle, \left\langle b_2, b_2' \right\rangle, \\
& \left\langle c_1, c_1' \right\rangle, \left\langle c_1, c_2' \right\rangle, \left\langle c_2, c_1' \right\rangle, \left\langle c_2, c_2' \right\rangle, \\
& \left\langle o_1, o_1' \right\rangle, \left\langle o_2, o_2' \right\rangle\}
\end{aligned}
$$

is a global bisimulation between $\mathfrak{M}_1$ and $\mathfrak{M}_2$. First, it is obvious that $\rightleftharpoons$ is global and that conditions (a) and (b) from Definition 11 are satisfied. Let us check condition (c). Suppose that $u \rightleftharpoons v$ with $u \in W_1$ and $v \in W_2$. By (b) it is clear that $u \in t^{\mathfrak{W}_1}$ iff $v \in t^{\mathfrak{W}_2}$ for all set terms t without occurences of function symbols. Then, $u \in f_{\exists R}^{\mathfrak{W}_1}(t^{\mathfrak{W}_1})$ iff u has a R successor u' in W_1 such that $u' \in t^{\mathfrak{W}_1}$ iff v has a R successor v' in W_2 such that $v' \in t^{\mathfrak{W}_2}$ iff $v \in f_{\exists R}^{\mathfrak{W}_2}(t^{\mathfrak{W}_2})$, according to the definition of $\rightleftharpoons$ and the induction hypotheses. We leave it to the reader to check the cases of the other function symbols in a similar manner.

It remains to establish that conditions (d) and (e) of the definition of E-bisimulation hold. We show only (d). There is a total of 16 cases to be considered. We go through some of them and leave the rest to the reader.

(1) We have $\langle a_1, c_1 \rangle \in E$ and four possibilities to instantiate the antecedent of (d). If the cases $a_1 \rightleftharpoons a_1'$ and $c_1 \rightleftharpoons c_2'$, choose $a_1' \rightleftharpoons c_2'$. In the cases $a_1 \rightleftharpoons a_2'$ and $c_1 \rightleftharpoons c_1'$, choose $a_2' \rightleftharpoons c_1'$. In all cases (d) is satisfied.

(2) We have $\langle b_2, c_2 \rangle \in E$ and again four possibilities to instantiate the antecedent of (d). If the cases $b_2 \rightleftharpoons b_1'$ and $c_2 \rightleftharpoons c_1'$, choose $b_1' \rightleftharpoons c_1'$. In the cases $b_2 \rightleftharpoons b_2'$ and $c_2 \rightleftharpoons c_2'$, choose $b_2' \rightleftharpoons c_2'$. Again, in all cases (d) is satisfied.

It should be clear how to check the remaining cases. We have thus shown that $\rightleftharpoons$ defines an E-bisimulation between the models $\mathfrak{M}_1$ and $\mathfrak{M}_2$.

Let us now briefly discuss which modifications are needed in the cases of the E-connections $\mathcal{C}(\mathcal{ALC}, \mathcal{MS})$ and $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$. Roughly, we can use the same models as

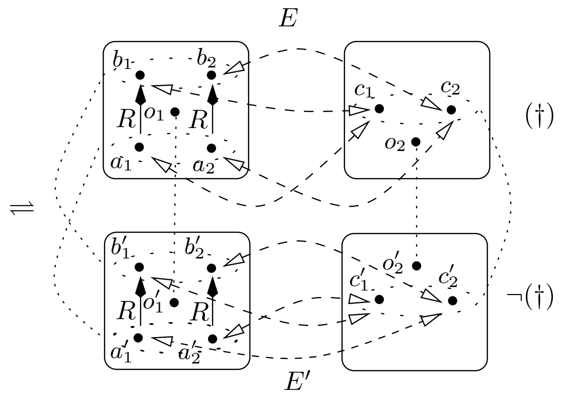

Figure 1: E-bisimilar models for $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$

before. But this time, to interpret the distance operators of $\mathcal{MS}$ and the interior/closure operators as well as the universal modalities of $\mathsf{S4}_u$, we need to specify metrics d/d' (in the case of $\mathcal{MS}$) and topologies $\mathfrak{T}/\mathfrak{T}'$ (in the case of $\mathsf{S4}_u$) for $\mathfrak{M}/\mathfrak{M}'$, respectively. A straightforward solution to this problem is to choose the discret metric in the case of $\mathcal{MS}$—i.e., the metric d such that $d(x, y) = 0$ iff $x = y$ and $d(x, y) = 1$, otherwise—and to choose the topology induced by the discret metric in the case of $\mathsf{S4}_u$. The above proof can then be mimicked without further major modifications.

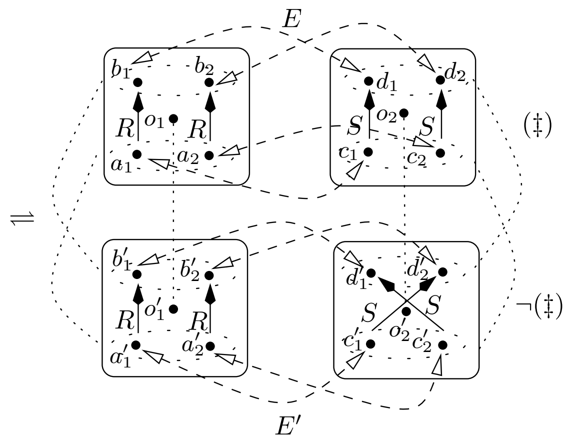

Figure 2: E-bisimilar models for $\mathcal{C}(\mathcal{SHIQ}, \mathcal{ALCO})$

For the proof of the second claim, we ask the reader to follow the lines of the proof of (i) and restore the details from Figure 2. ❑

For example, since we cannot express (†), we cannot say that the spatial extension of the capital of any country is included in the spatial extension of that country without enumerating the countries). Similarly, since (‡) is non-expressible, we cannot say that any child of any person who is a citizen of some country is a citizen of that same country.

8 Discussion

The investigation of combination methods for KR&R-formalisms consists, to a large extent, of the analysis of the trade-off between possible interactions of the components in the combined system and its computational properties. In this paper we studied a combination method which was proved to be extremely robust in the computational sense. Of course, the price for this is that the interaction between the components is limited. For example, in $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u, \mathsf{PTL})$ it is not possible to say that Germany and France were always externally connected or that the territory of the EU will never contract. Assertions of this type require more interaction between the components. We hope that starting from E-connections as a 'harmless' way of combining formalisms, it is possible to develop and study a hierarchy of more and more interactive combinations in a systematic manner.

Finally, note that the term 'harmless' used above is a bit misleading. Theorem 6 shows that decidability is inherited by the E-connection from its components. However, even if we have 'practical' algorithms for the components we do not obtain from our rather abstract model-theoretic proof a 'practical' algorithm for the E-connection. Moreover, it is unlikely that a 'practical' version of this general transfer result exists at all. Nevertheless, the proof of Theorem 6 indicates that in many cases existing practical decision procedures for the components can be combined so as to obtain practical decision procedures for the E-connection. We are currently working on such 'practical' combination techniques.

Acknowledgements

We should like to thank F. Baader and U. Sattler for helpful discussion. The work of M. Zakharyaschev was partially supported by U.K. EPSRC grants no. GR/R42474/01 and GR/R45369/01. The work of O. Kutz and F. Wolter was supported by DFG grant no. Wo 583/3-1.

References

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843, 1983.

[2] A. Artale and E. Franconi. A temporal description logic for reasoning about actions and plans. *Journal of Artificial Intelligence Research*, 9:463–506, 1998.

[3] F. Baader, H.-J. Bürckert, B. Nebel, W. Nutt, and G. Smolka. On the expressivity of feature logics with negation, functional uncertainty, and sort equations. *Journal of Logic and Computation*, (2):1–18, 1993.

[4] F. Baader, C. Lutz, H. Sturm, and F. Wolter. Fusions of description logics and abstract description systems. To appear in Journal of Artificial Intelligence Research (JAIR) 2002. (Visit `www.informatik.uni-leipzig.de/~wolter`.)

[5] F. Baader, C. Lutz, H. Sturm, and F. Wolter. Fusions of description logics. In F. Baader and U. Sattler, editors, *2000 International Workshop on Description Logics*, Aachener Beiträge zur Informatik, pages 21–30. Wissenschaftsverlag Mainz in Aachen, 2000.

[6] F. Baader. Augmenting concept languages by transitive closure of roles: An alternative to terminological cycles. In *Proceedings of the 13th IJCAI*, pages 446–451, Sidney, Australia, 1991.

[7] F. Baader and P. Hanschke. A scheme for integrating concrete domains into concept languages. In *Proceedings of the 13th IJCAI*, pages 452–457, Sidney, Australia, 1991.

[8] F. Baader and P. Hanschke. Extensions of concept languages for a mechanical engineering application. In *Proceedings of the 16th German Annual Conference on Artificial Intelligence (GWAI'92)*, volume 671 of *Lecture Notes in Computer Science*, pages 132–143, Bonn (Germany), 1992. Springer.

[9] F. Baader and A. Laux. Terminological logics with modal operators. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, pages 808–814, Montreal, Canada, 1995. Morgan Kaufmann, Los Altos.

[10] F. Baader and H.-J. Ohlbach. A multi-dimensional terminological knowledge representation language. *Journal of Applied Non-Classical Logics*, 5:153–197, 1995.

[11] F. Baader and W. Schulz, editors *Frontiers of Combinining Systems.* Kluwer Academic Publishers, 1996.

[12] B. Bennett. Modal logics for qualitative spatial reasoning. *Bulletin of the Interest Group in Pure and Applied Logic (IGPL)*, 4(1):23–45, 1996.

[13] C. Bettini. Time-dependent concepts: Representation and reasoning using temporal description logics. *Data and Knowledge Engineering*, 22:1–38, 1997.

[14] M. de Rijke and D. Gabbay, editors *Frontiers of Combining Systems II.* Research Studies Press LTD, England, 2000.

[15] D. Gabbay, I. Hodkinson, and M. Reynolds. *Temporal Logic: Mathematical Foundations and Computational Aspects*, volume 1. Oxford University Press, 1994.

[16] D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-Dimensional Modal Logics: Theory and Applications.* Studies in Logic. Elsevier, 2002. (To appear; avaliable at `http://www. dcs.kcl.ac.uk/staff/mz`.)

[17] V. Goranko and D. Passy. Using the universal modality: gains and questions. *Journal of Logic and Computation*, 2:5–30, 1992.

[18] V. Haarslev, C. Lutz, and R. Möller. Defined topological relations in description logics. In *Proceedings of the 6th International Conference on the Principles of Knowledge Representation and Reasoning (KR'98)*, pages 112–124, 1998.

[19] V. Haarslev, C. Lutz, and R. Möller. A description logic with concrete domains and role-forming predicates. *Journal of Logic and Computation*, 9(3):351–384, 1999.

[20] J. Halpern and Y. Shoham. A propositional modal logic of time intervals. *Journal of the ACM*, 38:935–962, 1991.

[21] B. Hollunder and W. Nutt. Subsumption algorithms for concept languages. 1990. DFKI Research Report RR-90-04, Germany Research Center for Artificial Intelligence, Kaiserslautern.

[22] I. Horrocks, U. Sattler, and S. Tobies. Practical reasoning for expressive description logics. In *Proceedings of the Sixth International Conference on Logic for Programming and Automated Reasoning (LPAR'99)*, pages 161–180, 1999.

[23] I. Horrocks and U. Sattler. Ontology reasoning in the $\mathsf{SHIQ}(d)$ description logic. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 199–204. Morgan Kaufmann, Los Altos, 2001.

[24] O. Kutz, H. Sturm, N.-Y. Suzuki, F. Wolter, and M. Zakharyaschev. Logics of metric spaces. Submitted 2001. (Visit `http://www.informatik.uni-leipzig.de/~wolter`.)

[25] A. Laux. Beliefs in multi-agent worlds: A terminological approach. In *Proceedings of the 11th European Conference on Artificial Intelligence (ECAI'94)*, pages 299–303, Amsterdam, The Netherlands, 1994.

[26] C. Lutz, H. Sturm, F. Wolter, and M. Zakharyaschev. Temporalizing tableaux. Manuscript, 2001.

[27] C. Lutz. Reasoning with concrete domains. In T. Dean, editor, *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI'99)*, pages 90–95, Stockholm, Sweden, 1999. Morgan Kaufmann, Los Altos.

[28] C. Lutz. NEXPTIME-Complete Description Logics with Concrete Domains. In R. Gore, A. Leitsch and T. Nipkov, editors, *Automated Reasoning, Proceedings of the First International Joint Conference (IJCAR'2001)*, pages 45–60, Lecture Notes on Artificial Intelligence 2083, Springer, 2001.

[29] D. Randell, Z. Cui, and A. Cohn. A spatial logic based on regions and connection. In *Proceedings of the 3rd International Conference on the Principles of Knowledge Representation and Reasoning (KR'92)*, pages 165–176. Morgan Kaufmann, Los Altos, 1992.

[30] J. Renz. A canonical model of the Region Connection Calculus. In A. Cohn, L. Schubert, and S. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference (KR-98)*, pages 330–341. Morgan Kaufman, 1998.

[31] A. Schaerf. Reasoning with individuals in concept languages. *Data and Knowledge Engineering*, 13(2):141–176, 1994.

[32] K. Schild. A correspondence theory for terminological logics: Preliminary report. In *Proceedings of the 13th IJCAI*, pages 466–471, Sidney, Australia, 1991.

[33] V. Shehtman. "Everywhere" and "Here". *Journal of Applied Non-Classical Logics*, 9, 1999.

[34] H. Sturm, N.-Y. Suzuki, F. Wolter, and M. Zakharyaschev. Semi-qualitative reasoning about distances: a preliminary report. In *Logics in Artificial Intelligence. Proceedings of JELIA 2000, Malaga, Spain*, pages 37–56, Berlin, 2000. Springer.

[35] A. Tarski. Der Aussagenkalkül und die Topologie. *Fundamenta Mathematicae*, 31:103–134, 1938.

[36] F. Wolter. Fusions of modal logics revisited. In *Advances in Modal Logic*, volume 1, pages 361–379, CSLI Publications, Stanford, 1998.

[37] F. Wolter and M. Zakharyaschev. Spatial reasoning in RCC-8 with Boolean region terms. In W. Horn, editor, *Proceedings of the 14th European Conference on Artificial Intelligence, ECAI 2000, Berlin, Germany*, pages 244–248. IOS Press, 2000.

[38] F. Wolter and M. Zakharyaschev. Spatio-temporal representation and reasoning based on RCC-8. In *Proceedings of the seventh Conference on Principles of Knowledge Representation and Reasoning, KR2000, Breckenridge, USA*, pages 3–14, Montreal, Canada, 2000. Morgan Kaufmann.

[39] F. Wolter and M. Zakharyaschev. Satisfiability problem in description logics with modal operators. In *Proceedings of the 6th International Conference on the Principles of Knowledge Representation and Reasoning (KR'98)*, pages 512–523. Morgan Kaufmann, Los Altos, 1998.

[40] F. Wolter and M. Zakharyaschev. Modal description logics: modalizing roles. *Fundamenta Informaticae*, 39:411–438, 1999.

[41] F. Wolter and M. Zakharyaschev. Multi-dimensional description logics. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI'99)*, pages 104–109. Morgan Kaufmann, Los Altos, 1999.