E-Connections

Oliver Kutz

LAC 2018 - Melbourne

joint work with
Carsten Lutz, Frank Wolter,
and Michael Zakharyaschev

Overview

- Why combine logics?
- Why are existing accounts not sufficient?
- Abstract Description Systems.
- Basic ideas of E-connections.
- Examples of various E-connections.
- E-connections and Ontologies.

Why combine Logics?

- Many application domains require different **perspectives** to be adequately captured.
- **`Universal languages'** like first-order logic are computationally difficult (undecidable) and/or incomplete (second-order);
- **Special purpose formalism** are often more natural, both semantically and syntactically;
- **Highly expressive** special purpose formalism are difficult to extend.

How to combine logics?

- Most combination techniques are algorithmically very difficult (e.g. **products**).
- Transfer results for **fusions** require their models be closed under disjoint unions.
- General **meta-theorems** (transfer results) presuppose an abstract framework for KR-formalisms.

Three Goals:

Find a combination technique that is **widely applicable**, **useful**, and **computationally robust**:

- use a suitable (common) representation of various KR formalisms;
- allow for (regimented) interaction between the components;
- ensure that the combination of decidable components remains decidable.

A Solution

- The languages of Description Logics, Modal Logics, Temporal and Spatial Logics etc. can be translated into the common language of ADLs. (**Abstract Description Languages**)
- The intended semantics of each logic is converted into classes of ADMs. (**Abstract Description Models**)
- **E-Connections of ADSs** allow for interaction between the components, while decidability is inherited from the component logics.

Abstract Description Systems

Definition 1. 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 $\mathcal{L}$.

Basic ideas of ADS

The languages of ADSs are **quantifier-free** and solely build from **terms** (denoting sets), **objects variables** (denoting individuals), **Booleans**, and **function** and **relation symbols.**

- prop. variables/concept expressions → terms;
- roles of DLs → binary relations;
- modalities/DL constructors etc. → functions;
- The intended interpretation of the symbols is fixed by choosing an appropriate class of **admissible models.**

Abstract Description Language

Definition 2. The **terms** of an **abstract description language** (ADL) are of the following form:

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

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

- $R(a_1, \ldots, a_m)$, for a_i object variables;
- $a : t$, for a an object variable and t a term.

Example: $\mathcal{ALC}$ as ADL

- The concept expressions C of $\mathcal{ALC}$ can be regarded as terms $C^\sharp$ of an ADS $\mathcal{ALC}^\sharp$.
- $A \leadsto A^\sharp$, A a concept name, $A^\sharp$ a set-variable;
- $R \leadsto f_{\forall R}, f_{\exists R}$, R a role, $f_{\forall R}$, $f_{\exists R}$ unary function symbols;

Then, put inductively

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

$\mathcal{ALC}$ as ADL

- Object names of $\mathcal{ALC}$ are treated as obj. variables;
- Role names are treated as binary relations;
- Term assertions correspond to general TBoxes;
- Object assertions correspond to ABoxes;
- The connection between roles R and the associated function symbols is fixed by choosing a suitable class of admissible models.

Semantics for ADLs

Definition 3. An **abstract description model** (ADM) for an ADL $\mathcal{L}$ is a structure of the form

$$\mathfrak{W} = \left\langle W, (x^{\mathfrak{W}})_{x \in \mathcal{V}}, (a^{\mathfrak{W}})_{a \in \mathcal{X}}, (f_i^{\mathfrak{W}})_{i \in \mathcal{I}}, (R_i^{\mathfrak{W}})_{i \in \mathcal{R}} \right\rangle,$$

where

- W is a non-empty set;
- $x^{\mathfrak{W}} \subseteq W$;
- $a^{\mathfrak{W}} \in W$;
- $f_i^{\mathfrak{W}}$ is a function mapping n_i-tuples $\langle X_1, \ldots, X_{n_i} \rangle$ of subsets of W to a subset of W;
- $R_i^{\mathfrak{W}}$ are m_i-ary relations on W.

$\mathcal{ALC}$ as ADS: Models

Given an $\mathcal{ALC}$-model

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

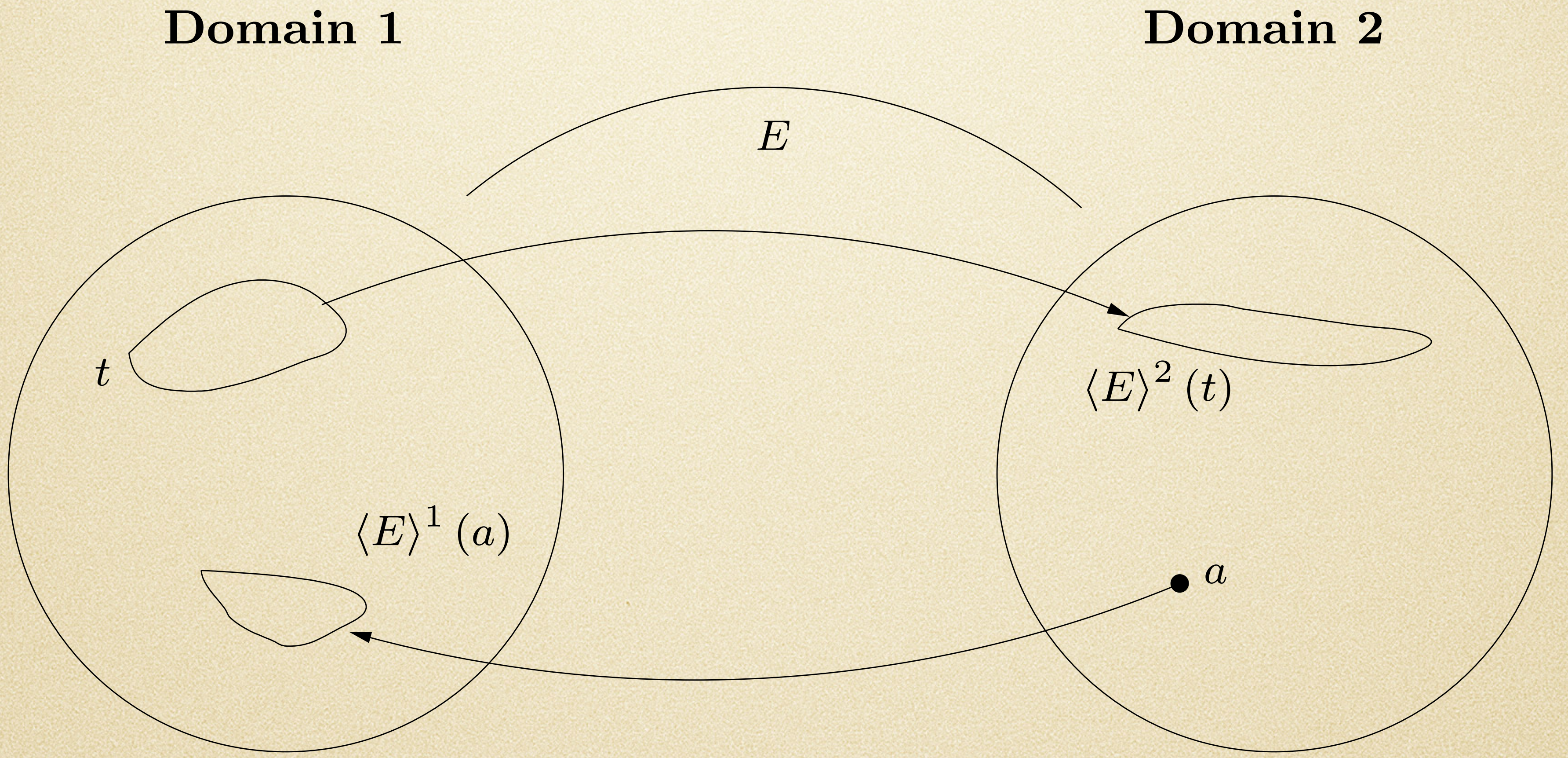
$\mathcal{M}$ contains the model $\mathfrak{M} = \left\langle \Delta, \mathcal{V}^{\mathfrak{M}}, \mathcal{X}^{\mathfrak{M}}, F^{\mathfrak{M}}, R^{\mathfrak{M}} \right\rangle$, where F consists of the function symbols $f_{\forall R_i}$ and $f_{\exists R_i}$, and R is the set of all role names of $\mathcal{ALC}$,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

- $(A^{\sharp})^{\mathfrak{M}} = A^{\mathcal{I}}$, 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;
- $f_{\forall R_i} X = \{d \in \Delta \mid \forall d' \in \Delta \; (d R_i^{\mathcal{I}} d' \rightarrow d' \in X)\}$, for all roles 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 roles R_i.

E-Connections: Basics

- Take a finite number of logical formalism, representable as ADSs;
- Assume their languages (signatures) to be disjoint (might share logical symbols);
- Assume an arbitrary number of link relations, establishing links between foreign domains;
- Add operators to the disjoint union of the languages which interpret these links;

2-dim E-connection



2-dim E-connection

A structure

$$\mathfrak{M} = \langle \mathfrak{W}_1, \mathfrak{W}_2, E \rangle \,,$$

where $\mathfrak{W}_i \in \mathcal{M}_i$ and $E \subseteq W_1 \times W_2$ is called a **model** for $\mathcal{C}$.

To define e.g. the **extension** $t^{\mathfrak{M}} \subseteq W_2$ of a 2-term t we add:

- $(\langle E \rangle^2 (s))^{\mathfrak{M}} = \{y \in W_2 \mid \exists x \in s^{\mathfrak{M}} : xEy\}$

The **truth-relation** $\models$ between models $\mathfrak{M}$ for $\mathcal{C}$ and assertions of $\mathcal{C}$ is standard.

General Case

- Allow n ADSs, n finite, allow arbitrarily many n-ary link relations, and add (n-1)-ary operators:

$$\left\langle E_j \right\rangle^i (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$$

$$E_j^{\mathfrak{M}} \subseteq W_1 \times \cdots \times W_n$$

$$(\left\langle E_j \right\rangle^i (\bar{t}_i))^{\mathfrak{M}} = \{x \in W_i \mid \underset{\ell \neq i}{\exists} x_\ell \in t_\ell^{\mathfrak{M}} \ (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \in E_j^{\mathfrak{M}}\}.$$

Important: Operators can be iterated

2-dim Example

Take the description logic $\mathcal{ALCO}$ (extending $\mathcal{ALC}$ with nominals) and, using concepts $\mathsf{Country}$, Treaty, etc., nominals EU, $\mathsf{Schengen\_treaty}$, object names France etc., and a role member, etc.

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

After that you want to say something about borders in Europe.

S4u as Spatial Logic

- The modal logic S4u, i.e., Lewis's modal system S4 enriched with the universal modality, is an important formalism for reasoning about spatial knowledge.
- Tarski interpreted the basic S4 (without the universal modality) in topological spaces as early as 1938.
- Later, the universal box was added in order to allow the representation of and reasoning about the well-known RCC-8 set of relations between two regions in a topological space.

2-dim Example

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

$$\mathsf{EQ}(\langle E\rangle^2(\mathsf{EU}), \langle E\rangle^2(\mathsf{Portugal}) \sqcup \ldots)$$

$$\mathsf{EC}(\langle E\rangle^2(\mathsf{France}), \langle E\rangle^2(\mathsf{Luxembourg}))$$

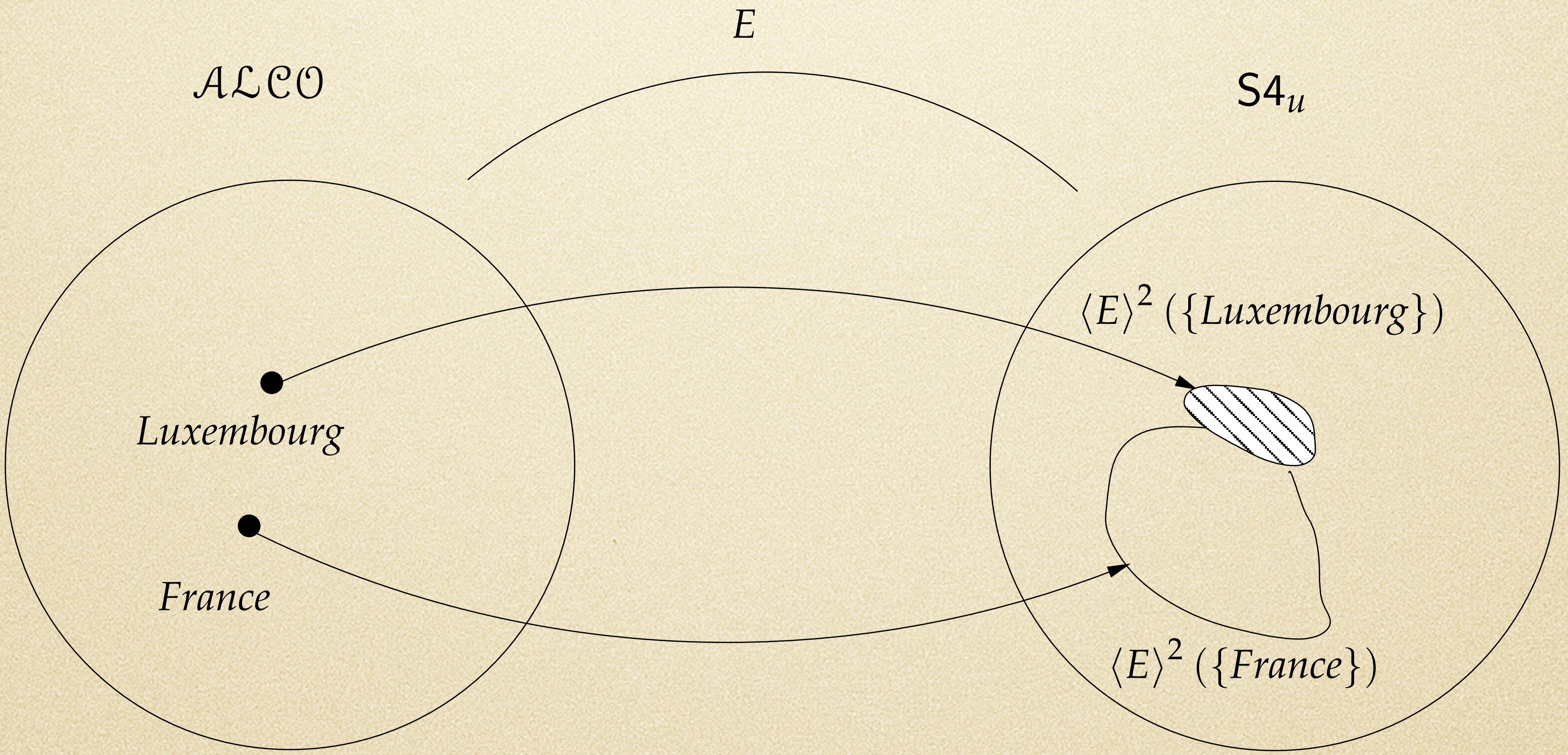
$$\mathsf{NTPP}(\langle E\rangle^2(\mathsf{Luxembourg}), \langle E\rangle^2(\exists\mathsf{member.Schengen\_Treaty}))$$

$$\langle E\rangle^2(\mathsf{France}) = \mathbb{CI}\,\langle E\rangle^2(\mathsf{France})$$

i.e., 'the space occupied by the EU is the space occupied by its members', etc.

Interaction

France : $\exists$member.Schengen_Treaty ?



Going 3-dim:

$$\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u, \mathsf{PTL})$$

- Add a **temporal dimension** to the connection $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$:
  - Extend the connection $\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u)$ with one more ADS—e.g. propositional temporal logic PTL, which uses the constructors `Since` and `Until` and is interpreted in **flows of time** like $\mathbb{N}$.
  - The ternary relation $E(x, y, z)$ 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 the $\mathcal{ALCO}$).

3-dim E-Connection

$$\mathcal{C}(\mathcal{ALCO}, \mathsf{S4}_u, \mathsf{PTL})$$

Then we can say, for example:

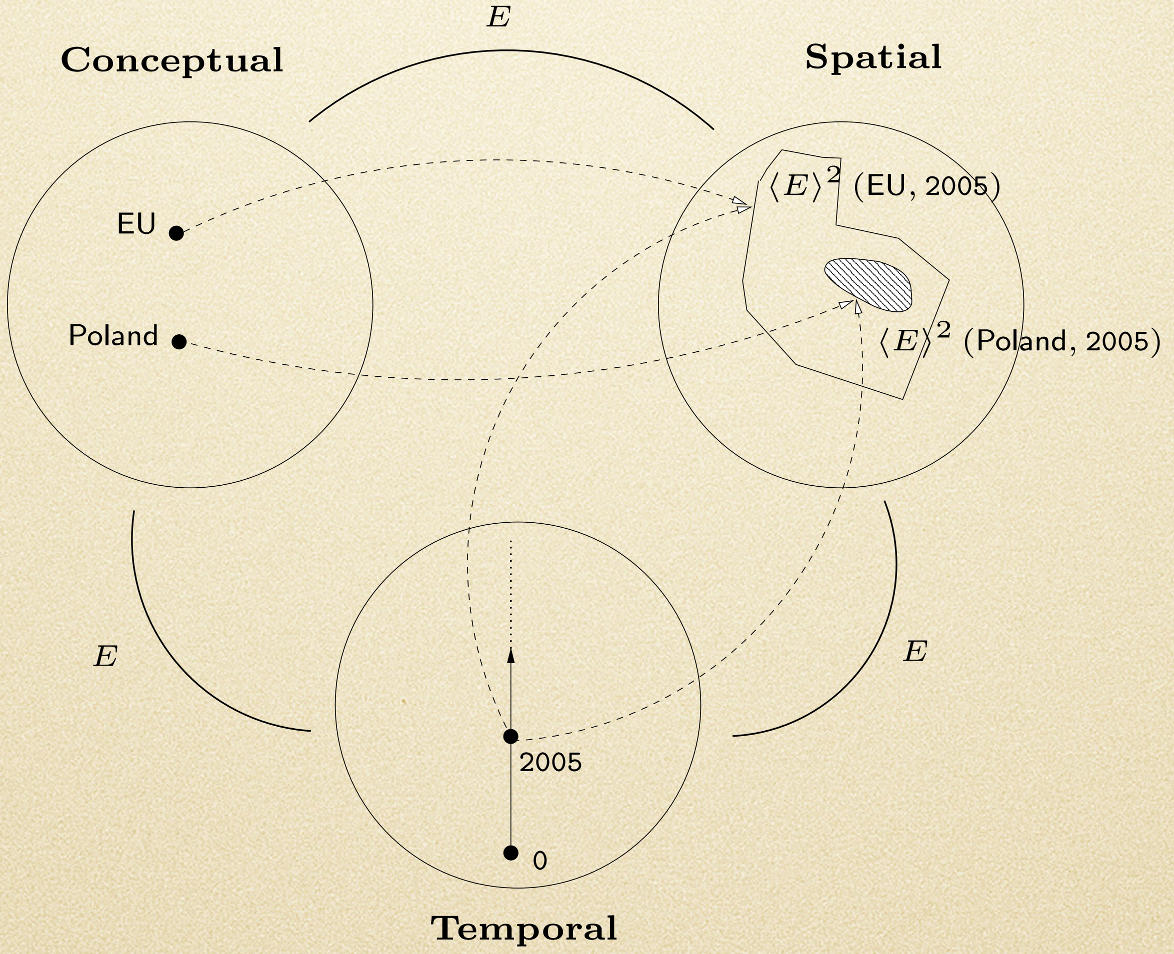
$$\langle E \rangle^2 (\mathsf{Poland}, 2005) \sqsubseteq \langle E \rangle^2 (\mathsf{EU}, 2005)$$

$$\mathsf{PO}(\langle E \rangle^2 (\mathsf{Austria}, 1914), \langle E \rangle^2 (\mathsf{Italy}, 1950))$$

$$\Box_F \neg \langle E \rangle^3 (Basel, \mathsf{EU}),$$

- '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.'.

A 3-dim connection



Results for Basic E-Connections:

Theorem 4. *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.*

Corollary 5. *The satisfiability problem for any E-connection of DLs with a decidable satisfiability problem for ABoxes with respect to TBoxes as well as logics like* PTL, $\mathcal{MS}$, *and* $\mathsf{S4}_u$ *is decidable.*

- The time complexity of the decision problem is at most one non-deterministic exponential higher than the worst complexity of one of the components.

E-Connections: Extensions

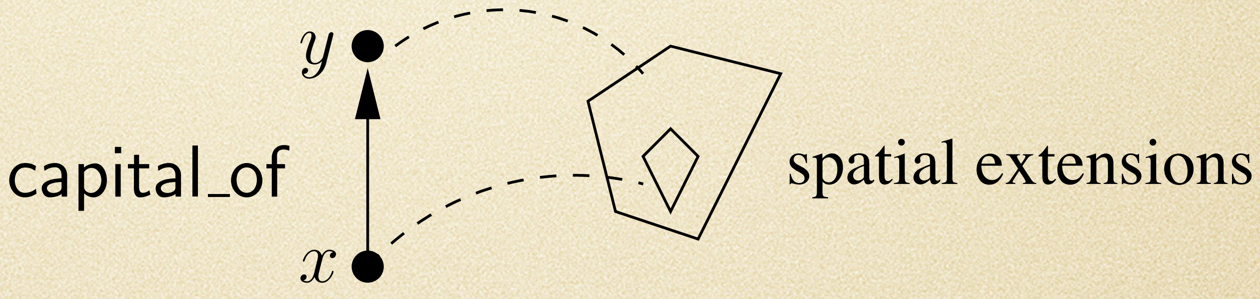
- Structural conditions on links;
- Link operators on object variables;
- Boolean operators on links.
- Number restrictions on links;

Structural Conditions

Conditions on connections like:

$$\forall x \forall y \forall z (xRy \to (xEz \to yEz)) \qquad (*)$$

are not expressible in the language of connections, but are very natural:



Theorem 7. *Suppose the components $\mathcal{S}_i$ of a E-connection $\mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$ are decidable. Then it is decidable whether an assertion is satisfiable in a model of $\mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$ satisfying* $(*)$.

Example: $\mathcal{C}^{\mathcal{E}}(\mathcal{SHIQ}^{\sharp}, \mathcal{ALCO}^{\sharp})$

(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 whose residence is the UK either is a child of a person whose residence is the UK, or is a UK citizen or has a work permit in the UK’.

Link relations: \`C’ for \`having citizenship in’;
\`R’ for \`having residence in’;
\`W’ for \`having work permit in’.

$$\langle C \rangle^{1} (\{EU\}) \sqsubseteq \neg(\geq 2\mathsf{married}.\top);$$

$$\exists \mathsf{child\_of}.\, \langle C \rangle^{1} (\{UK) \sqsubseteq \langle C \rangle^{1} (\{UK\});$$

$$\langle R \rangle^{1} (\{UK\}) \sqsubseteq \exists \mathsf{child\_of}^{-1}.\, \langle R \rangle^{1} (\{UK\}) \sqcup \langle C \rangle^{1} (\{UK\}) \sqcup \langle W \rangle^{1} (\{UK\}).$$

Operators on object variables

- Basic E-connection don't allow to apply link operators to **object names**, just to **nominals**;
- Thus, in the E-connection of $\mathcal{SHIQ}$ and $\mathcal{ALCO}$, we cannot form the expression:

$$country \sqcap \langle C \rangle^2 (Bob)$$

where `Bob' is an object name of SHIQ (denoting the set of all countries where `Bob' has citizenship).

Operators on object variables

Theorem. Adding link operators on object variables to decidable basic E-connections preserves decidability.

- The time complexity is, as before, one non-deterministic exponential higher.
- Note that the addition of nominals to an arbitrary logic can yield an undecidable one.

Booleans on Links

- Basic E-connections do not allow for **interaction** between the different links;
- How do we express:

(iv) 'People taking residence in the country of their citizenship'.

We need the **intersection** of the links \`C' and \`R':

$$\mathsf{Human\_being} \sqcap \langle C \cap R \rangle^1 (\mathsf{Country}).$$

Booleans on Links

Theorem. Adding Boolean operators on links to decidable basic E-connections preserves decidability.

- The time complexity bound is as before, but optimal in general:
- The E-connection of propositional logic with itself and with Booleans on links can simulate the product logic **S5 x S5**, so grows from NP to NEXPTIME-complete.

Number restrictions on links

- Number restrictions are very useful in DLs, and equally natural to employ on links;
- However, they are too expressive in general:
- Number restrictions can be used to force links to be bijective functions:

$$\top_2 = \langle \leq 1E \rangle^2 (\top_1), \;\; \top_2 = \langle \geq 1E \rangle^2 (\top_1),$$

$$\top_1 = \langle \leq 1E \rangle^1 (\top_2), \;\; \top_1 = \langle \geq 1E \rangle^1 (\top_2)$$

So **nominals** can be **exported** from one component to another, yielding undecidability.

Number restrictions on links

- Need the concept of `**number tolerance**': typical: logics with nominals are not number tolerant, those closed under disjoint unions are.

Theorem. Adding qualified number restrictions on links to decidable basic E-connections preserves decidability, whenever all components are number tolerant.

Transfer Results: Overview

	Basic	O	B	Q
Basic	+	+	+	number tolerant
O			+	-
B				-

Ontologies / Distributed Reasoning

- E-connections are a prominent framework for distributed reasoning and ontology integration:
- Distributed Description Logics (DDLs) (Borgida & Serafini) are a special case of E-connections, but less expressive, and specialised to DLs only.
- Ontologies using different DLs can be integrated by establishing appropriate links, relating the terminologies of the different ontologies.

Ontology Factorisation

- Suppose ontology O is formalisable only in a very complex (or even undecidable) DL L.
- O may in fact `talk' about a number of simpler subdomains, ontologies O_1 ... O_n, representable by simpler DLs L_1...L_n.
- Interaction between these simpler ontologies can be pushed into link relations, yielding a more transparent and decidable framework.
- Ontology maintenance can be made locally.

Past Future Work

- Optimised algorithms for concrete E-connections: first steps: tableaux for "weak" E-connections of SHIQ, SHOQ and SHOI in the Pellet system;
- Sub-Boolean logics (Baader/Ghilardi);
- Further expressive means and study of expressivity in general;
- Reasoning with inconsistencies;
- Tools for ontology integration.

E-connections and DOL

- Contextualised alignments in DOL
- How do we fully integrate E-connections (and therefore DDL) into DOL
- How do we treat general bridge theories in DOL
- Abstract away from abstract description systems
- Proof support via translations
- Currently: Resolution for basic E-connections

Literature

- **Connecting abstract description systems**.
  O Kutz, F Wolter, M Zakharyaschev. In: 8th International Conference of Principles of Knowledge Representation and Reasoning, KR 2002.

- **E-Connections of Abstract Description Systems**.
  O Kutz, C Lutz, F Wolter, M Zakharyaschev. Artificial intelligence, 156 (1), 1-73, 2004.