Exchanging Description Logic Knowledge Bases

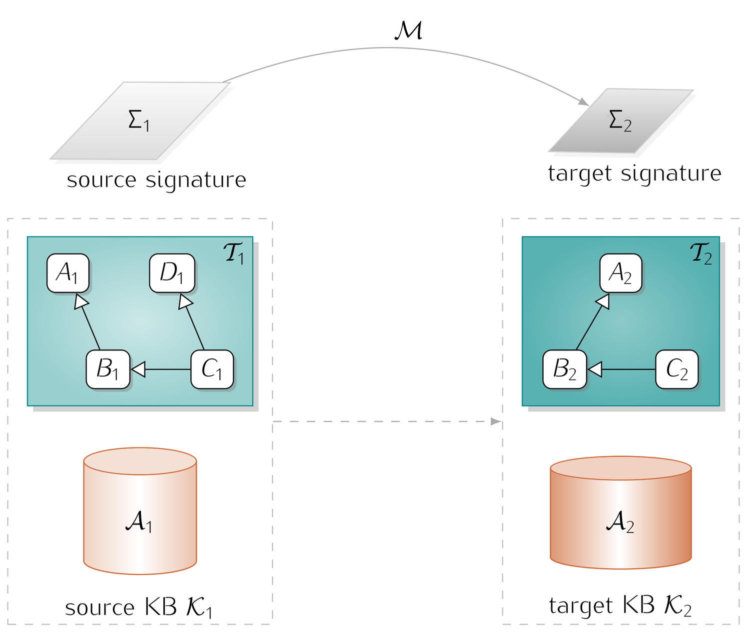
Marcelo Arenas
Dept. of Computer Science, PUC Chile
marenas@ing.puc.cl

Elena Botoeva, Diego Calvanese, and Vladislav Ryzhikov
KRDB Research Centre, Free Univ. of Bozen-Bolzano, Italy
lastname@inf.unibz.it

Evgeny Sherkhonov
ISLA, University of Amsterdam, Netherlands
e.sherkhonov@uva.nl

Our Framework: Knowledge Base Exchange

given a mapping $\mathcal{M}$ and a source knowledge base (KB) $\mathcal{K}_1$, compute a target KB $\mathcal{K}_2$ that is a *solution* for $\mathcal{K}_1$ under $\mathcal{M}$.



We consider exchange of Description Logic (DL) KBs: each KB is constituted by a TBox and an ABox, and mapping is a set of DL inclusions.

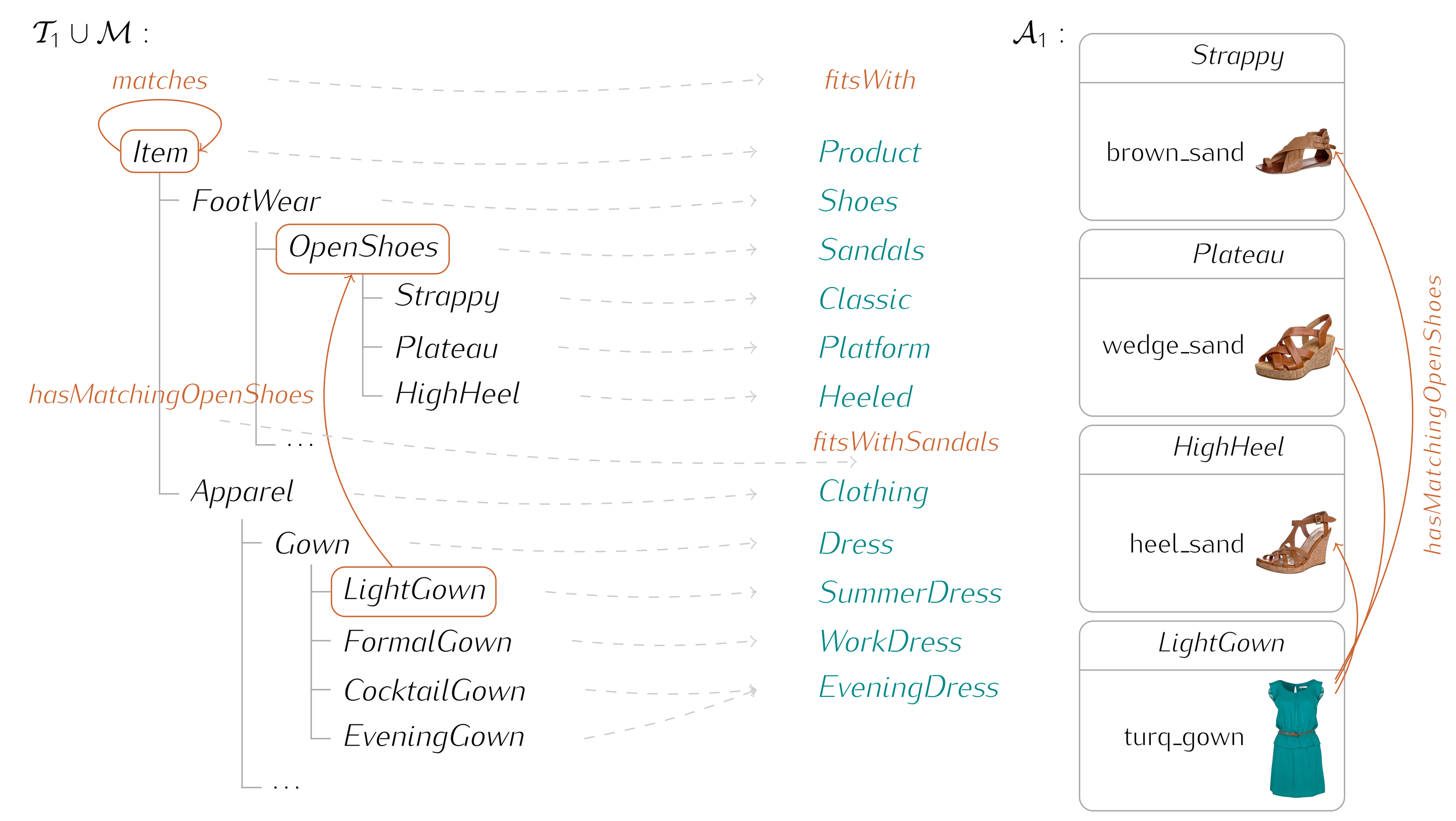
Let $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ be a mapping, $\mathcal{K}_1$ and $\mathcal{K}_2$ KBs over Σ_1 and Σ_2.

- $\mathcal{K}_2$ is a *universal solution* for $\mathcal{K}_1$ under $\mathcal{M}$ if $\textsc{Mod}(\mathcal{K}_2) = \textsc{Sat}_{\mathcal{M}}(\textsc{Mod}(\mathcal{K}_1))$.
- $\mathcal{K}_2$ is a *universal $\mathcal{Q}$-solution* if $\forall q \in \mathcal{Q}$ over Σ_2, $cert(q, \langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1\rangle) = cert(q, \mathcal{K}_2)$.

In this paper we study the problem of KB exchange for *DL-Lite*$_{\mathcal{R}}$ and *DL-Lite*$_{RDFS}$.

Why Knowledge Base Exchange

Assume we bought shoe and clothes goods described in an ontology *Goods*. We want to open an *online shop* selling the goods via a website that will use the information in the ontology for rendering the pages. But first we want to change the ontology vocabulary as it will be displayed to the user.

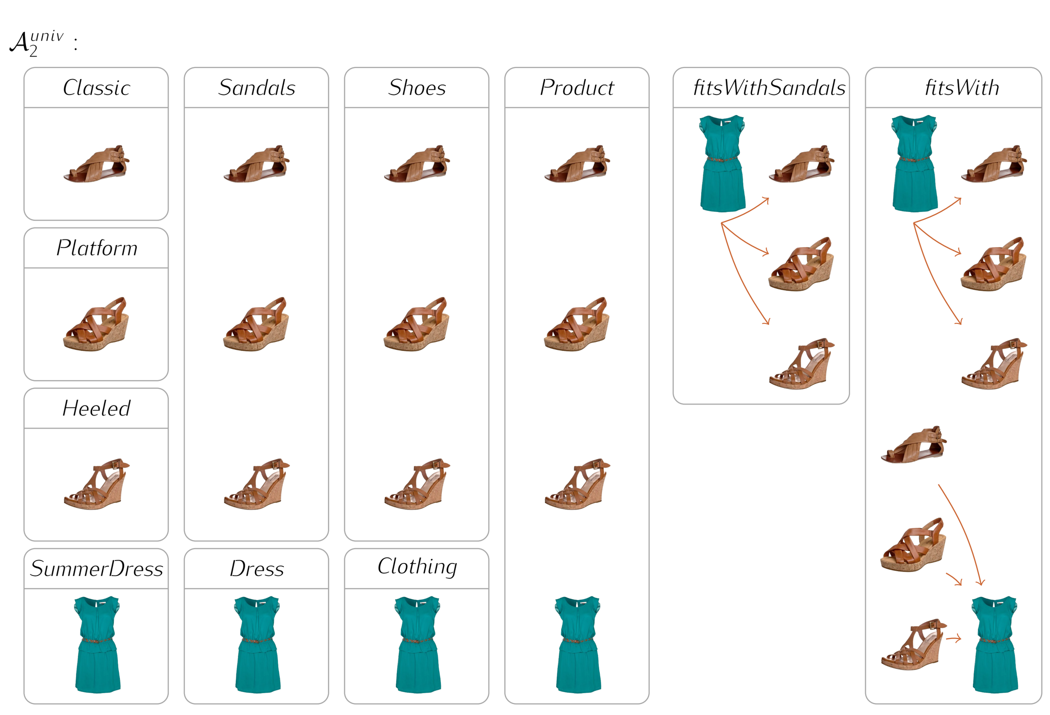


An ontology over the new vocabulary can be obtained as a solution of the *KB exchange problem* for $\mathcal{M}$ and $\langle \mathcal{T}_1, \mathcal{A}_1\rangle$.

Universal and UCQ-Solutions

There are two types of possible solutions to the KB exchange problem: with the empty TBox and with a non-empty TBox.

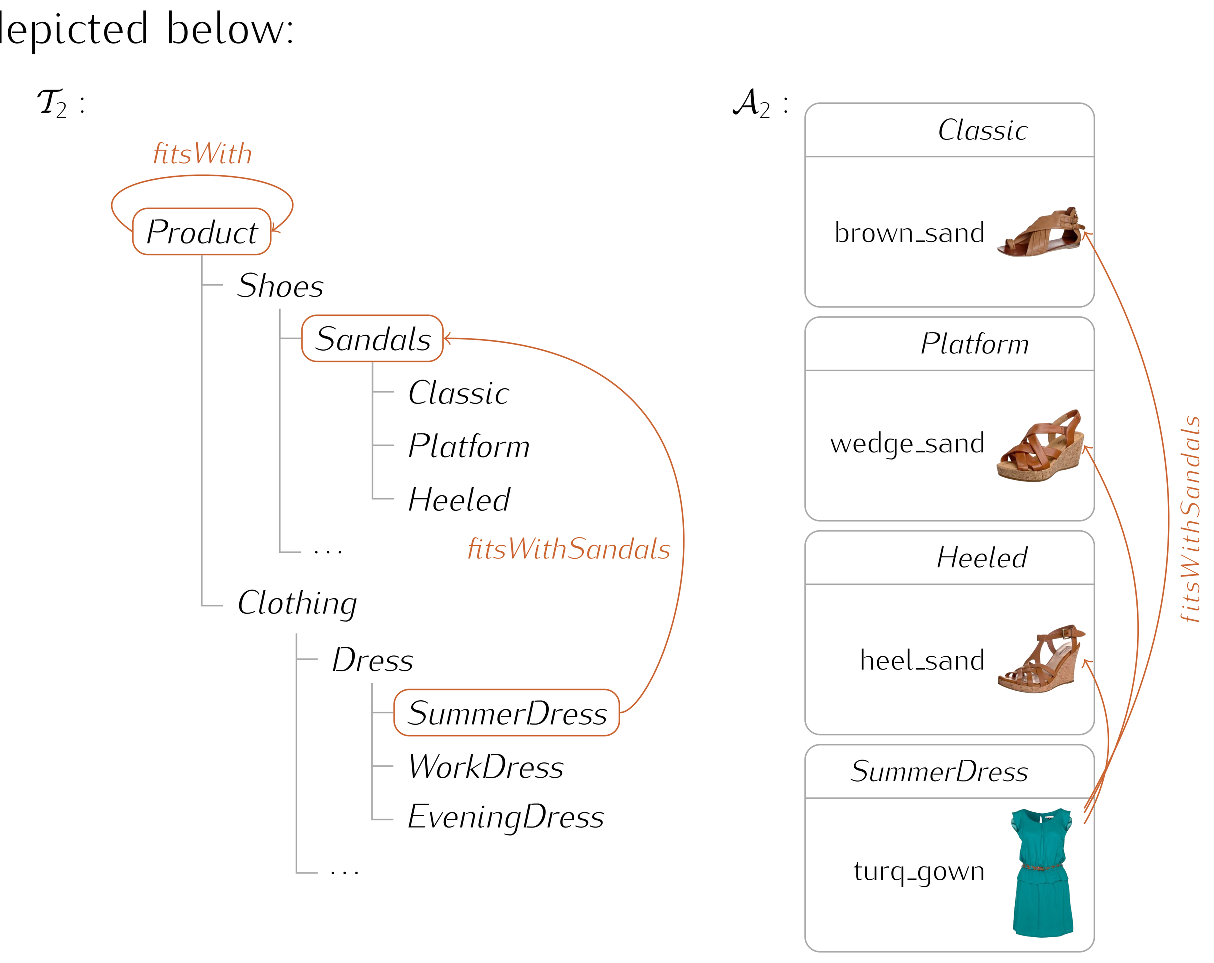
The *universal solution* has the empty TBox and it is the ABox $\mathcal{A}_2^{univ}$ of the form:



The universal solution has several drawbacks:

- a universal solution (in *DL-Lite*$_{\mathcal{R}}$) does not always exist;
- if it exists, then its TBox is empty;
- some smallest universal solutions are exponential in the size of the input.

An alternative to the universal solution could be the *universal UCQ-solution* $\langle \mathcal{T}_2, \mathcal{A}_2\rangle$ depicted below:



In this solution, for each axiom in $\mathcal{T}_1$ there is a corresponding axiom in $\mathcal{T}_2$. Moreover, $\mathcal{A}_2$ is the direct translation of $\mathcal{A}_1$ and does not depend on $\mathcal{T}_1$.

The universal UCQ-solution has its merits:

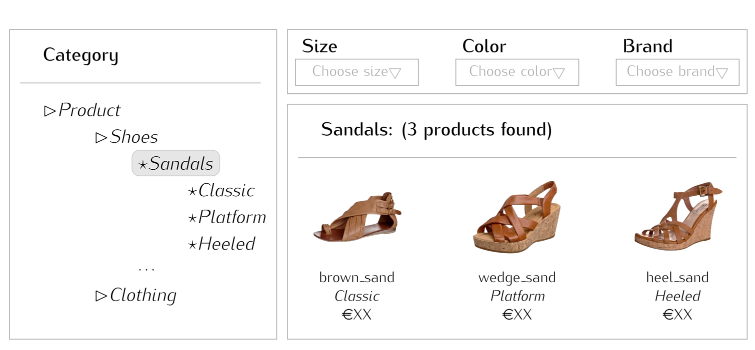
- it is polynomial in the size of the mapping and the source KB.
- it is not distinguishable from the universal solution by means of UCQs.

Why Universal UCQ-Solutions

The website should provide the following functionality.

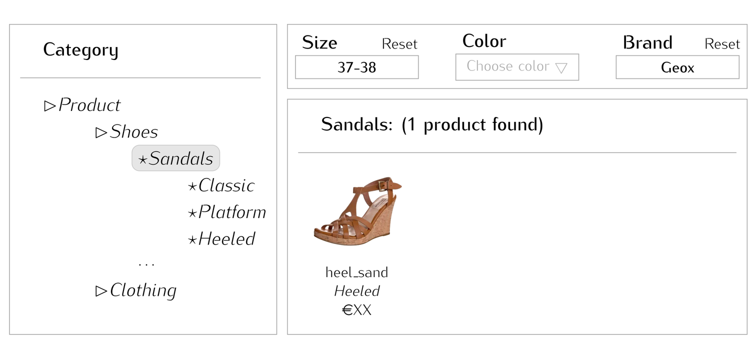
- browse different categories, e.g., to show all sandals, or sandals on platform;
- filter the listed products according to several criteria, such as color, brand;
- recommend items that fit the currently viewed product;

Such functionality can be implemented querying the target KB with UCQs:



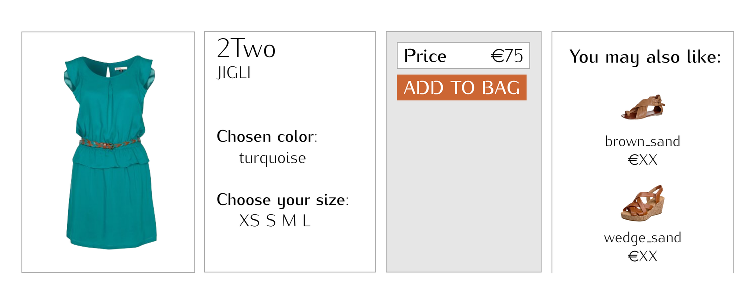
Query:

$q(x) \leftarrow Sandals(x)$



Query:

$$q(x) \leftarrow Sandals(x), hasSize(x, 37), hasBrand(x, \text{geox}) \lor Sandals(x), hasSize(x, 38), hasBrand(x, \text{geox})$$



Query:

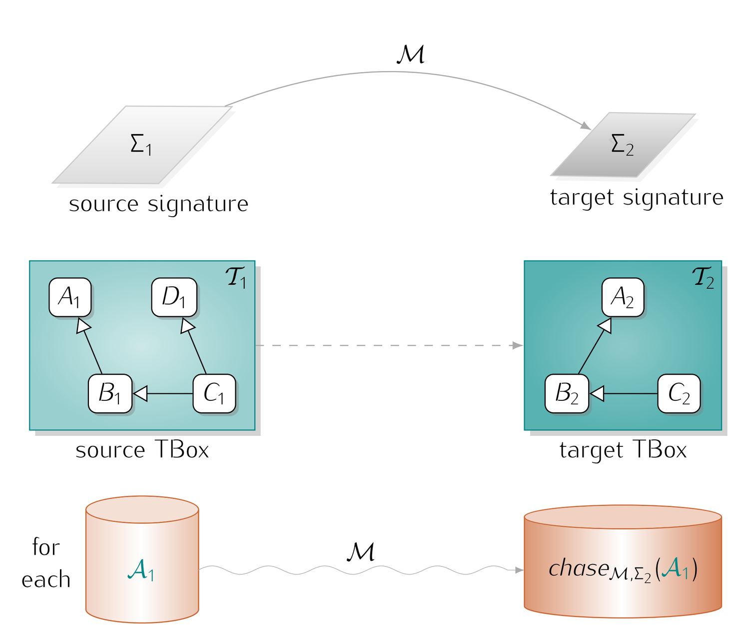
$q(x) \leftarrow fitsWith(\text{turq\_gown}, x)$

⇒ Universal UCQ-solutions are the *preferred* solutions.

New Problem: Representability of a TBox

To minimize the exchange of ABox information, we are interested in solutions that contain as much implicit knowledge as possible:

given a mapping $\mathcal{M}$ and a source TBox $\mathcal{T}_1$, compute a target TBox $\mathcal{T}_2$, if exists, such that for each ABox $\mathcal{A}_1$, $\langle \mathcal{T}_2, chase_{\mathcal{M},\Sigma_2}(\mathcal{A}_1)\rangle$ is a universal UCQ-solution for $\langle \mathcal{T}_1, \mathcal{A}_1\rangle$ under $\mathcal{M}$.



The problem of UCQ-representability for *DL-Lite*$_{RDFS}$ TBoxes can be solved in *polynomial time*.