Ontology Reuse: Better Safe than Sorry

Bernardo Cuenca Grau, Ian Horrocks, Yevgeny Kazakov and Ulrike Sattler
The University of Manchester
School of Computer Science
Manchester, M13 9PL, UK
{bcg, horrocks, ykazakov, sattler}@cs.man.ac.uk

1 Motivation

The design, maintenance, reuse, and integration of ontologies are complex tasks. Like software engineers, ontology engineers need to be supported by tools and methodologies that help them to minimize the introduction of errors, i.e., to ensure that ontologies are consistent and do not have unexpected consequences. In order to develop this support, important notions from software engineering, such as *module*, *black-box behavior*, and *controlled interaction*, must be adapted.

Recently, there has been growing interest in the topic of modularity in ontology engineering [10, 9, 8, 5, 3], motivated by the above mentioned application needs. This paper extends our previous results[3]. We focus on the problem of "safe" reuse of ontologies and consider the scenario in which we are developing an ontology $\mathcal{P}$ and want to reuse a set $\mathbf{S}$ of symbols—that is, concept names, role names and individuals— from a "foreign" ontology $\mathcal{Q}$ without changing their meaning.

Suppose that an ontology engineer is building an ontology about research projects, which specifies different types of projects according to the research topics they focus on. For example, the concepts $\mathsf{Genetic\_Disorder\_Project}$ and $\mathsf{Cystic\_Fibrosis\_EUProject}$ describe projects about genetic disorders and European projects about cystic fibrosis respectively, as given by the axioms P1 and P2 in Figure 1. The ontology engineer is an expert on research projects; he knows, for example, that every instance of $\mathsf{EU\_Project}$ must be an instance of $\mathsf{Project}$ (axiom P3) and that the role $\mathsf{has\_Focus}$ can be applied only to instances of $\mathsf{Project}$ (axiom P4). He may be unfamiliar, however, with most of the topics the projects cover and, in particular, with the terms $\mathsf{Cystic\_Fibrosis}$ and $\mathsf{Genetic\_Disorder}$ mentioned in P1 and P2. In order to complete the projects ontology with suitable definitions of these medical terms, he decides to reuse the knowledge about these subjects from a well-established medical ontology.

Suppose that $\mathsf{Cystic\_Fibrosis}$ and $\mathsf{Genetic\_Disorder}$ are described in an ontology $\mathcal{Q}$ containing axioms M1-M4 in Figure 1. The most straightforward way to reuse these concepts is to import in $\mathcal{P}$ the ontology $\mathcal{Q}$—that is, to add the axioms from $\mathcal{Q}$ to the axioms of $\mathcal{P}$ and work with the extended ontology $\mathcal{P} \cup \mathcal{Q}$. Importing additional axioms into an ontology may result into new logical consequences. For example, axioms M1–M4 in $\mathcal{Q}$ imply that every instance of $\mathsf{Cystic\_Fibrosis}$ is an instance of $\mathsf{Genetic\_Disorder}$:

$$\mathcal{Q} \models \alpha := (\mathsf{Cystic\_Fibrosis} \sqsubseteq \mathsf{Genetic\_Disorder}) \tag{1}$$

Indeed, $\alpha_1 = (\mathsf{Cystic\_Fibrosis} \sqsubseteq \mathsf{Genetic\_Disorder})$ follows from axioms M1 and M2 as well as from M1 and M3; α follows from α_1 and M4. Using inclusion α from

	Ontology of medical research projects $\mathcal{P}$:
P1	$\mathsf{Genetic\_Disorder\_Project} \equiv \mathsf{Project} \sqcap \exists \mathsf{has\_Focus}.\underline{\mathsf{Genetic\_Disorder}}$
P2	$\mathsf{Cystic\_Fibrosis\_EUProject} \equiv \mathsf{EUProject} \sqcap \exists \mathsf{has\_Focus}.\underline{\mathsf{Cystic\_Fibrosis}}$
P3	$\mathsf{EUProject} \sqsubseteq \mathsf{Project}$
P4	$\exists \mathsf{has\_Focus}.\top \sqsubseteq \mathsf{Project}$
E1	$\mathsf{Project} \sqcap (\underline{\mathsf{Genetic\_Disorder}} \sqcap \underline{\mathsf{Cystic\_Fibrosis}}) \sqsubseteq \bot$
E2	$\forall \mathsf{has\_Focus}.\underline{\mathsf{Cystic\_Fibrosis}} \sqsubseteq \exists \mathsf{has\_Focus}.\underline{\mathsf{Genetic\_Disorder}}$
	Ontology of medical terms $\mathcal{Q}$:
M1	$\underline{\mathsf{Cystic\_Fibrosis}} \equiv \mathsf{Fibrosis} \sqcap \exists \mathsf{located\_In}.\mathsf{Pancreas} \sqcap \exists \mathsf{has\_Origin}.\mathsf{Genetic\_Origin}$
M2	$\mathsf{Genetic\_Fibrosis} \equiv \mathsf{Fibrosis} \sqcap \exists \mathsf{has\_Origin}.\mathsf{Genetic\_Origin}$
M3	$\mathsf{Fibrosis} \sqcap \exists \mathsf{located\_In}.\mathsf{Pancreas} \sqsubseteq \mathsf{Genetic\_Fibrosis}$
M4	$\mathsf{Genetic\_Fibrosis} \sqsubseteq \underline{\mathsf{Genetic\_Disorder}}$

Fig. 1: Reusing medical terminology in an ontology on research projects

(1) and axioms P1–P3 from ontology $\mathcal{P}$ we can now prove that every instance of Cystic_Fibrosis_EUProject must also be an instance of Genetic_Disorder_Project:

$$\mathcal{P} \cup \mathcal{Q} \models \beta := (\mathsf{Cystic\_Fibrosis\_EUProject} \sqsubseteq \mathsf{Genetic\_Disorder\_Project}) \qquad (2)$$

Note that, on the one hand, $\mathcal{P} \not\models \beta$ and, on the other hand, the ontology engineer might be not aware of (2), even though it concerns the terms of primary scope in $\mathcal{P}$.

It is to be expected that axioms like α in (1) from an imported ontology $\mathcal{Q}$ cause new entailments like β in (2) over the terms defined in the main ontology $\mathcal{P}$. One would not expect, however, that the meaning of the terms defined in $\mathcal{Q}$ changes as a consequence of the import since these terms are supposed to be completely specified within $\mathcal{Q}$. Such a side effect is highly undesirable for the modeling of ontology $\mathcal{P}$ since the ontology engineer of $\mathcal{P}$ might not be an expert on the subject of $\mathcal{Q}$ and is not supposed to alter the meaning of the terms defined in $\mathcal{Q}$, not even implicitly. The meaning of the reused terms might change after the import due, for example, to modeling errors. In particular, suppose the ontology engineer has learned about the concepts Genetic_Disorder and Cystic_Fibrosis from the ontology $\mathcal{Q}$ (including the dependency (1)) and has decided to introduce additional axioms formalizing the following statements:

"*Every instance of* Project *is different from every instance of* Genetic_Disorder *and every instance of* Cystic_Fibrosis." (3)

"*Every project that* has_Focus *on* Cystic_Fibrosis, *also* has_Focus *on* Genetic_Disorder" (4)

Note that the statements (3) and (4) add new information about projects and, intuitively, they should not change or constrain the meaning of the medical terms.

Suppose the ontology engineer has formalized statements (3) and (4) in $\mathcal{P}$ using axioms E1 and E2 respectively. At this point, he has introduced modeling errors by translating the words *and* and *every* as conjunction $\sqcap$ and value restriction $\forall$ respectively. As a consequence, axioms E1 and E2 do not correspond to (3) and (4): E1 actually formalizes the following statement: *"Every instance of* $\mathsf{Project}$ *is different from every* *common* *instance of* $\mathsf{Genetic\_Disorder}$ *and* $\mathsf{Cystic\_Fibrosis}$*"*, and E2 expresses that *"Every object that* $\mathsf{has\_Focus}$ *only on* $\mathsf{Cystic\_Fibrosis}$ *if at all, also* $\mathsf{has\_Focus}$ *on* $\mathsf{Genetic\_Disorder}$*"*. This kind of modeling errors are difficult to detect, especially when they do not lead to inconsistencies in the original ontology.

Note that, although axiom E1 does not correspond to fact (3), it is still a consequence of (3) and hence it should not constrain the meaning of the medical terms. In contrast, E2 is not a consequence of (4) and, in fact, it does constrain the meaning of these medical terms. Indeed, axioms E1 and E2 together with axioms P1-P4 from $\mathcal{P}$ imply new axioms about the concepts $\mathsf{Cystic\_Fibrosis}$ and $\mathsf{Genetic\_Disorder}$, namely their disjointness:

$$\mathcal{P} \models \gamma := (\mathsf{Genetic\_Disorder} \sqcap \mathsf{Cystic\_Fibrosis} \sqsubseteq \bot) \tag{5}$$

The entailment (5) can be proved using axiom E2 which is equivalent to:

$$\top \sqsubseteq \exists \mathsf{has\_Focus}.(\mathsf{Genetic\_Disorder} \sqcup \neg \mathsf{Cystic\_Fibrosis}) \tag{6}$$

The inclusion (6) and P4 imply that every element in the domain must be a project—that is, $\mathcal{P} \models (\top \sqsubseteq \mathsf{Project})$. Now, together with axiom E1, this implies (5). The axioms E1 and E2 not only imply new statements about the medical terms, but also cause inconsistencies when used together with the imported axioms from $\mathcal{Q}$. Indeed, from (1) and (5) we obtain $\mathcal{P} \cup \mathcal{Q} \models \delta := (\mathsf{Cystic\_Fibrosis} \sqsubseteq \bot)$ which expresses the inconsistency of the concept $\mathsf{Cystic\_Fibrosis}$.

To summarize, we have seen that importing an external ontology can lead to undesirable side effects in our knowledge reuse scenario, like the entailment of new axioms or even inconsistencies over the reused vocabulary.

The contributions of this paper are as follows. First, we formalize some reasoning services that are relevant for ontology reuse. In particular, we propose the notion of safe reuse of a signature in an ontology. Second, we show that the problem of checking safety is undecidable in $\mathcal{ALCO}$. This result leaves us with two alternatives: we can either focus on simple DLs for which this problem is decidable, or we may look for sufficient conditions for safety—that is, an incomplete solution. We define in general terms the notion of a sufficient condition for safety— a *safety class*— and define a family of safety classes—called locality— with some compelling properties. We have implemented a safety checking algorithm and obtained empirical evidence of its usefulness in practice.

This paper comes with an extended version available online [4]; we refer the reader to the extended version for further technical details.

2 Conservative Extensions and Safety

As argued in the previous section, an important requirement for the reuse of an ontology $\mathcal{Q}$ within an ontology $\mathcal{P}$ should be that $\mathcal{P} \cup \mathcal{Q}$ produces exactly the same logical consequences over the vocabulary of $\mathcal{Q}$ as $\mathcal{Q}$ alone does. This requirement can be naturally formulated using the well-known notion of a conservative extension, which has recently been investigated in the context of ontologies [7, 8].

Definition 1 (Conservative Extension). *Let $\mathcal{L}$ be a description logic and let $\mathcal{O}_1 \subseteq \mathcal{O}$ be two ontologies, and $\mathbf{S}$ a signature over $\mathcal{L}$. We say that $\mathcal{O}$ is an* $\mathbf{S}$-conservative extension *of $\mathcal{O}_1$ w.r.t. $\mathcal{L}$, if for every axiom α over $\mathcal{L}$ with $\mathsf{Sig}(\alpha) \subseteq \mathbf{S}$, we have $\mathcal{O} \models \alpha$ iff $\mathcal{O}_1 \models \alpha$. We say that $\mathcal{O}$ is a* conservative extension *of $\mathcal{O}_1$ w.r.t. $\mathcal{L}$ if $\mathcal{O}$ is an $\mathbf{S}$-conservative extension of $\mathcal{O}_1$ w.r.t. $\mathcal{L}$ for $\mathbf{S} = \mathsf{Sig}(\mathcal{O}_1)$.*

Definition 1 implies that, in order to show that $\mathcal{P} \cup \mathcal{Q}$ is not a $\mathbf{S}$-conservative extension of $\mathcal{Q}$ it suffices to find an axiom α over $\mathbf{S}$ that is implied by $\mathcal{P} \cup \mathcal{Q}$ but not by $\mathcal{Q}$ alone. In our example, the ontology $\mathcal{P} \cup \mathcal{Q}$ is *not* a conservative extension of $\mathcal{Q}$ w.r.t. $\mathbf{S} = \{\mathsf{Cystic\_Fibrosis}, \mathsf{Genetic\_Disorder}\}$ since $\mathcal{P} \cup \mathcal{Q}$ implies $\alpha_1 = (\mathsf{Cystic\_Fibrosis} \sqsubseteq \bot)$ and $\alpha_2 = (\mathsf{Genetic\_Disorder} \sqsubseteq \bot)$ over $\mathbf{S}$, but $\mathcal{Q}$ does not.

Definition 1 applies to fixed $\mathcal{P}$, $\mathcal{Q}$. In realistic scenarios, however, the reused ontology $\mathcal{Q}$ may *evolve* beyond the control of the designers of $\mathcal{P}$, which may not be authorized to modify $\mathcal{Q}$, or may decide at a later time to reuse the symbols Cystic_Fibrosis and Genetic_Disorder from a medical ontology other than $\mathcal{Q}$. Therefore, for application scenarios in which the external ontology $\mathcal{Q}$ may change, it is reasonable to "abstract" from the particular $\mathcal{Q}$ under consideration. In other words, the fact that the axioms in $\mathcal{P}$ do not change the meaning of the external symbols in $\mathbf{S}$ should be *independent* from the particular meaning of these symbols. This idea can be made precise as follows:

Definition 2 (Safety for a Signature). *Let $\mathcal{L}$ be an ontology language, and let $\mathcal{O}$ be an ontology and $\mathbf{S}$ a signature over $\mathcal{L}$. We say that $\mathcal{O}$* is safe for $\mathbf{S}$ w.r.t. $\mathcal{L}$*, if for every ontology $\mathcal{O}'$ over $\mathcal{L}$ with $\mathsf{Sig}(\mathcal{O}) \cap \mathsf{Sig}(\mathcal{O}') \subseteq \mathbf{S}$, we have that $\mathcal{O} \cup \mathcal{O}'$ is a conservative extension of $\mathcal{O}'$ w.r.t. $\mathcal{L}$.*

Definition 2 captures the intuition in our example: the axioms in $\mathcal{P}$ should not yield new consequences over the signature $\mathbf{S}$ and the signature $\mathsf{Sig}(\mathcal{Q})$ of the reused ontology $\mathcal{Q}$, independently of the particular $\mathcal{Q}$ under consideration. In our example, the ontology $\mathcal{O} = \{\text{E2}\}$ is not safe w.r.t. $\mathbf{S} = \{\mathsf{Cystic\_Fibrosis}, \mathsf{Genetic\_Disorder}\}$ and $\mathcal{L} = \mathcal{ALC}$. Indeed, take $\mathcal{Q}_1 = \{\top \sqsubseteq \mathsf{Cystic\_Fibrosis}; \mathsf{Genetic\_Disorder} \sqsubseteq \bot\}$. Then, $\mathcal{Q}_1 \cup \mathcal{O}$ is inconsistent whereas $\mathcal{Q}_1$ is consistent. Consequently, $\mathcal{Q}_1 \cup \mathcal{O}$ is not a $\mathbf{S}$-conservative extension of $\mathcal{Q}_1$ w.r.t. $\mathcal{L} = \mathcal{ALC}$, and therefore $\mathcal{O} = \{\text{E2}\}$ is not safe for $\mathbf{S}$ and $\mathcal{L}$.

Proving that an ontology is safe is more involved than proving that it is not. One way to prove that $\mathcal{O}$ is $\mathbf{S}$-safe is the following: if we can take an arbitrary interpretation for the symbols in $\mathbf{S}$ and extend it to a model of $\mathcal{O}$ by interpreting the additional symbols in $\mathsf{Sig}(\mathcal{O})$, then $\mathcal{O}$ must be $\mathbf{S}$-safe. This property can be formalized as follows:

Definition 3. *Two interpretations $\mathcal{I}_1 = (\Delta^{\mathcal{I}_1}, \cdot^{\mathcal{I}_1})$ and $\mathcal{I}_2 = (\Delta^{\mathcal{I}_2}, \cdot^{\mathcal{I}_2})$* coincide on a signature $\mathbf{S}$ *(notation: $\mathcal{I}_1|_{\mathbf{S}} = \mathcal{I}_2|_{\mathbf{S}}$) if $\Delta^{\mathcal{I}_1} = \Delta^{\mathcal{I}_2}$ and $X^{\mathcal{I}_1} = X^{\mathcal{I}_2}$ for every $X \in \mathbf{S}$.*

Lemma 1. *Let $\mathcal{O}$ be a $\mathcal{SHOIQ}$ ontology and $\mathbf{S}$ a signature such that for every interpretation $\mathcal{I}$ there exists a model $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{J}|_{\mathbf{S}} = \mathcal{I}|_{\mathbf{S}}$. Then $\mathcal{O}$ is safe for $\mathbf{S}$ w.r.t. $\mathcal{L} = \mathcal{SHOIQ}$.*

We can now prove that the ontology $\mathcal{P}_1$ consisting of axioms P1-P4 is safe for $\mathbf{S} = \{\mathsf{Cystic\_Fibrosis}, \mathsf{Genetic\_Disorder}\}$. Take an arbitrary interpretation $\mathcal{I}$ of $\mathbf{S}$ and construct an interpretation $\mathcal{J}$ to be identical to $\mathcal{I}$ except for the interpretations of the atomic concepts Genetic_Disorder_Project, Cystic_Fibrosis_EUProject, Project, EUProject and the atomic role has_Focus, all of which we interpret in $\mathcal{J}$ as the empty set. All the axioms P1–P4, E2 are satisfied in $\mathcal{J}$and hence $\mathcal{J} \models \mathcal{P}_1$.

Using Lemma 1, we are now ready to show the main result in this section:

Theorem 1 (Undecidability for Safety of Ontologies). *Given an $\mathcal{ALC}$-ontology $\mathcal{O}$ and a signature $\mathbf{S}$ it is undecidable whether $\mathcal{O}$ is $\mathbf{S}$-safe w.r.t. $\mathcal{L} = \mathcal{ALCO}$.*

Proof. The proof is based on a reduction to a domino tiling problem. A domino system is a triple $D = (T, H, V)$ where $T = \{1, \dots, k\}$ is a finite set of *tiles* and $H, V \subseteq T \times T$ are *horizontal* and *vertical matching relations*. A *solution* for a domino system D is a mapping $t_{i,j}$ that assigns to every pair of integers $i, j \geq 1$ an element of T, such that $\langle t_{i,j}, t_{i,j+1} \rangle \in V$ and $\langle t_{i,j}, t_{i+1,j} \rangle \in H$. A *periodic solution* for a domino system D is a solution $t_{i,j}$ for which there exist integers $m \geq 1$, $n \geq 1$ called *periods* such that $t_{i+m,j} = t_{i,j}$ and $t_{i,j+n} = t_{i,j}$ for every $i, j \geq 1$.

Let $\mathcal{D}$ be the set of all domino systems, $\mathcal{D}_s$ be the subset of $\mathcal{D}$ that admit a solution and $\mathcal{D}_{ps}$ be the subset of $\mathcal{D}_s$ that admit a periodic solution. It is well-known [1, Theorem 3.1.7] that the sets $\mathcal{D} \setminus \mathcal{D}_s$ and $\mathcal{D}_{ps}$ are *recursively inseparable*, that is, there is no recursive (i.e. decidable) subset $\mathcal{D}' \subseteq \mathcal{D}$ of domino systems such that $\mathcal{D}_{ps} \subseteq \mathcal{D}' \subseteq \mathcal{D}_s$. For every domino system D, we construct a signature $\mathbf{S} = \mathbf{S}(D)$, an ontology $\mathcal{O} = \mathcal{O}(D)$ which consists of a single $\mathcal{ALC}$-axiom such that: **(a)** if D does not have a solution then $\mathcal{O} = \mathcal{O}(D)$ is safe for $\mathbf{S} = \mathbf{S}(D)$ w.r.t. $\mathcal{L} = \mathcal{ALCO}$, and **(b)** if D has a periodic solution then $\mathcal{O} = \mathcal{O}(D)$ is not safe for $\mathbf{S} = \mathbf{S}(D)$ w.r.t. $\mathcal{L} = \mathcal{ALCO}$.

In other words, for the set $\mathcal{D}'$ consisting of the domino systems D such that $\mathcal{O} = \mathcal{O}(D)$ is not safe for $\mathbf{S} = \mathbf{S}(D)$ w.r.t. $\mathcal{L} = \mathcal{ALCO}$, we have $\mathcal{D}_{ps} \subseteq \mathcal{D}' \subseteq \mathcal{D}_s$. Since $\mathcal{D} \setminus \mathcal{D}_s$ and $\mathcal{D}_{ps}$ are recursively inseparable, this implies undecidability for $\mathcal{D}'$ and hence for the problem of checking if $\mathcal{O}$ is an $\mathbf{S}$-safe w.r.t. $\mathcal{L} = \mathcal{ALCO}$, because otherwise one can use this problem for deciding membership in $\mathcal{D}'$.

Given $D = (T, H, V)$, let $\mathbf{S}$ consist of fresh atomic concepts A_i for every $i \in T$ and atomic roles r_H and r_V. Consider an ontology $\mathcal{O}_{\text{tile}}$ in Figure 2 constructed for D. Note that $\mathsf{Sig}(\mathcal{O}_{\text{tile}}) = \mathbf{S}$. The axioms of $\mathcal{O}_{\text{tile}}$ express the tiling conditions for a domino

$$
\begin{array}{lll}
(q_1) & \top \sqsubseteq A_1 \sqcup \cdots \sqcup A_k & \text{where } T = \{1, \dots, k\} \\
(q_2) & A_i \sqcap A_j \sqsubseteq \bot & 1 \leq i < j \leq k \\
(q_3) & A_i \sqsubseteq \exists r_H.(\bigsqcup_{\langle i,j \rangle \in H} A_j) & 1 \leq i \leq k \\
(q_4) & A_i \sqsubseteq \exists r_V.(\bigsqcup_{\langle i,j \rangle \in V} A_j) & 1 \leq i \leq k
\end{array}
$$

Fig. 2: An ontology $\mathcal{O}_{\text{tile}} = \mathcal{O}_{\text{tile}}(D)$ expressing tiling conditions for a domino system D

system D, namely (q_1) and (q_2) express that every domain element is assigned with a unique tile $t \in T$; (q_3) and (q_4) express that every domain element has horizontal and vertical matching successors. Now let s be an atomic role and B an atomic concept with $s, B \notin \mathbf{S}$. Let $\mathcal{O} := \{\beta\}$ where:

$$\beta := \quad \top \sqsubseteq \exists \boldsymbol{s}. \Big[\bigsqcup_{(C_i \sqsubseteq D_i) \in \mathcal{O}_{\text{tile}}} (C_i \sqcap \neg D_i) \sqcup (\exists r_H.\exists r_V.\boldsymbol{B} \sqcap \exists r_V.\exists r_H.\neg \boldsymbol{B}) \Big]$$

We say that r_H and r_V *commute in an interpretation* $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ if for every domain elements a, b, c, d_1 and d_2 from $\Delta^\mathcal{I}$ with $\langle a, b\rangle \in {r_H}^\mathcal{I}$, $\langle b, d_1\rangle \in {r_V}^\mathcal{I}$, $\langle a, c\rangle \in {r_V}^\mathcal{I}$, and $\langle c, d_2\rangle \in {r_H}^\mathcal{I}$, we have $d_1 = d_2$. The following claims can be easily proved:

Claim 1. If $\mathcal{O}_{\text{tile}}(D)$ has a model $\mathcal{I}$ in which r_H and r_V commute, then D has a solution.

Claim 2. If $\mathcal{I}$ is a model of $\mathcal{O} = \{\beta\}$, then either $\mathcal{I} \not\models \mathcal{O}_{\text{tile}}$ or r_H and r_V do not commute in $\mathcal{I}$.

To prove Property (a), we use Lemma 1 and demonstrate that if D has no solution then for every interpretation $\mathcal{I}$ there exists a model of $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{J}|_\mathbf{S} = \mathcal{I}|_\mathbf{S}$, which implies that $\mathcal{O}$ is safe for $\mathbf{S}$ w.r.t. $\mathcal{L}$. Let $\mathcal{I}$ be an arbitrary interpretation. Since D has no solution, then by the contra-position of Claim 1 either (1) $\mathcal{I}$ is not a model of $\mathcal{O}_{\text{tile}}$, or (2) r_H and r_V do not commute in $\mathcal{I}$. We demonstrate for both of these cases how to construct the required model $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{J}|_\mathbf{S} = \mathcal{I}|_\mathbf{S}$.

Case (1). If $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ is not a model of $\mathcal{O}_{\text{tile}}$ then there exists an axiom $(C_i \sqsubseteq D_i) \in \mathcal{O}_{\text{tile}}$ such that $\mathcal{I} \not\models (C_i \sqsubseteq D_i)$. That is, there exists a domain element $a \in \Delta^\mathcal{I}$ such that $a \in C_i^\mathcal{I}$ but $a \notin D_i^\mathcal{I}$. Let us define $\mathcal{J}$ to be identical to $\mathcal{I}$ except for the interpretation of the atomic role s which we define in $\mathcal{J}$ as $s^\mathcal{J} = \{\langle x, a\rangle \mid x \in \Delta\}$. Since the interpretations of the symbols in $\mathbf{S}$ has remained unchanged, we have $a \in C_i^\mathcal{J}$, $a \in \neg D_i^\mathcal{J}$, and so $\mathcal{J} \models (\top \sqsubseteq \exists s.[C_i \sqcap \neg D_j])$. This implies that $\mathcal{J} \models \beta$, and so, we have constructed a model $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{J}|_\mathbf{S} = \mathcal{I}|_\mathbf{S}$.

Case (2). Suppose that r_H and r_V do not commute in $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$. This means that there exist domain elements a, b, c, d_1 and d_2 from $\Delta^\mathcal{I}$ with $\langle a, b\rangle \in {r_H}^\mathcal{I}$, $\langle b, d_1\rangle \in {r_V}^\mathcal{I}$, $\langle a, c\rangle \in {r_V}^\mathcal{I}$, and $\langle c, d_2\rangle \in {r_H}^\mathcal{I}$, such that $d_1 \neq d_2$. Let us define $\mathcal{J}$ to be identical to $\mathcal{I}$ except for the interpretation of the atomic role s and the atomic concept B. We interpret s in $\mathcal{J}$ as $s^\mathcal{J} = \{\langle x, a\rangle \mid x \in \Delta\}$. We interpret B in $\mathcal{J}$ as $B^\mathcal{J} = \{d_1\}$. Note that $a \in (\exists r_H.\exists r_V.B)^\mathcal{J}$ and $a \in (\exists r_V.\exists r_H.\neg B)^\mathcal{J}$ since $d_1 \neq d_2$. So, we have $\mathcal{J} \models (\top \sqsubseteq \exists s.[\exists r_H.\exists r_V.B \sqcap \exists r_V.\exists r_H.\neg B])$ which implies that $\mathcal{J} \models \beta$, and thus, we have constructed a model $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{J}|_\mathbf{S} = \mathcal{I}|_\mathbf{S}$.

To prove Property (b), assume that D has a periodic solution $t_{i,j}$ with the periods $m, n \geq 1$. We show that $\mathcal{O}$ is not $\mathbf{S}$-safe w.r.t. $\mathcal{L}$. We build an $\mathcal{ALCO}$-ontology $\mathcal{O}'$ with $\mathsf{Sig}(\mathcal{O}) \cap \mathsf{Sig}(\mathcal{O}') \subseteq \mathbf{S}$ such that $\mathcal{O} \cup \mathcal{O}' \models (\top \sqsubseteq \bot)$, but $\mathcal{O}' \not\models (\top \sqsubseteq \bot)$. This will imply that $\mathcal{O}$ is not safe for $\mathcal{O}'$ w.r.t. $\mathcal{L} = \mathcal{ALCO}$, and hence, is not safe for $\mathbf{S}$ w.r.t. $\mathcal{L} = \mathcal{ALCO}$. We define $\mathcal{O}'$ such that every model of $\mathcal{O}'$ is a finite encoding of the periodic solution $t_{i,j}$. For every pair (i, j) with $1 \leq i \leq m$ and $1 \leq j \leq n$, introduce a fresh individual $a_{i,j}$ and take $\mathcal{O}'$ the extension of $\mathcal{O}_{\text{tile}}$ with the following axioms:

$$(p_1)\ \{a_{i_1,j}\} \sqsubseteq \exists r_V.\{a_{i_2,j}\} \qquad (p_2)\ \{a_{i_1,j}\} \sqsubseteq \forall r_V.\{a_{i_2,j}\}, \qquad i_2 = i_1 + 1 \mod m$$

$$(p_3)\ \{a_{i,j_1}\} \sqsubseteq \exists r_H.\{a_{i,j_2}\} \qquad (p_4)\ \{a_{i,j_1}\} \sqsubseteq \forall r_H.\{a_{i,j_2}\}, \qquad j_2 = j_1 + 1 \mod n$$

$$(p_5)\ \top \sqsubseteq \textstyle\bigsqcup_{1 \leq i \leq m,\, 1 \leq j \leq n} \{a_{i,j}\}$$

Axioms (p_1)–(p_5) ensure that r_H and r_V commute in every model of $\mathcal{O}'$. Indeed $\mathcal{O}'$ has a model corresponding to every periodic solution for D with periods m and n. Hence $\mathcal{O}' \not\models (\top \sqsubseteq \bot)$. Also, since every model of $\mathcal{O}'$ is a model of $\mathcal{O}_{\text{tile}}$ in which r_H and r_V commute, by Claim 2, $\mathcal{O}' \cup \mathcal{O}$ is unsatisfiable, so $\mathcal{O}' \cup \mathcal{O} \models (\top \sqsubseteq \bot)$.

□

3 Safety Classes

Theorem 1 leaves us with two alternatives: first, we can focus simple DLs for which this problem is decidable; second, we may look for sufficient conditions for the notion of safety—that is, if an ontology satisfies our conditions then we can guarantee that it is safe, but not necessarily vice versa. In this paper, we will explore the latter approach.

In general, any sufficient condition for safety can be represented by defining, for every signature $\mathbf{S}$, the set of ontologies over a language that satisfy the condition for that signature. These ontologies should be guaranteed to be safe.

Definition 4 (Class of Ontologies, Safety Class). *A* class of ontologies for a DL $\mathcal{L}$ and a signature $\mathbf{S}$ *is a function* $\mathbf{O}(\cdot)$ *that assigns to every subset* $\mathbf{S}'$ *of* $\mathbf{S}$ *a set* $\mathbf{O}(\mathbf{S}')$ *of ontologies in* $\mathcal{L}$*; it is* anti-monotonic *if for every* $\mathbf{S}_1 \subseteq \mathbf{S}_2$*, we have* $\mathbf{O}(\mathbf{S}_2) \subseteq \mathbf{O}(\mathbf{S}_1)$*; it is* subset-closed *if for every* $\mathbf{S}$ *and* $\mathcal{O}_1 \subseteq \mathcal{O}$ *we have that* $\mathcal{O} \in \mathbf{O}(\mathbf{S})$ *implies* $\mathcal{O}_1 \in \mathbf{O}(\mathbf{S})$*; it is* union-closed *if* $\mathcal{O}_1 \in \mathbf{O}(\mathbf{S})$ *and* $\mathcal{O}_2 \in \mathbf{O}(\mathbf{S})$ *implies* $(\mathcal{O}_1 \cup \mathcal{O}_2) \in \mathbf{O}(\mathbf{S})$ *for every* $\mathbf{S}$. *A* safety class *for* $\mathcal{L}$ *is a class of ontologies* $\mathbf{O}(\cdot)$ *for* $\mathcal{L}$ *such that, for every* $\mathbf{S}$*, every ontology in* $\mathbf{O}(\mathbf{S})$ *is safe for* $\mathbf{S}$*.*

Safety classes may admit many natural properties, as given in Definition 4. *Anti-monotonicity* intuitively means that if an ontology $\mathcal{O}$ can be proved to be safe w.r.t. $\mathbf{S}$ using the sufficient condition, then $\mathcal{O}$ can be proved to be safe w.r.t. every subset of $\mathbf{S}$. Similarly, *subset-closure* means that under the same assumption, every subset of $\mathcal{O}$ can also be proved to be safe using the same sufficient condition. If a safety class is *union-closed* and two ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$ can be proved safe using that sufficient test, then their union $\mathcal{O}_1 \cup \mathcal{O}_2$ can also be proved safe using the same test.

3.1 Locality

In this section we introduce a particular family of safety classes for $\mathcal{L} = \mathcal{SHOIQ}$, that we call locality classes. In Section 2, we have seen that, according to Lemma 1, one way to prove that $\mathcal{O}$ is $\mathbf{S}$-safe is to show that every $\mathbf{S}$-interpretation can be extended to a model of $\mathcal{O}$. Local ontologies are those for which safety can be used using Lemma 1.

Definition 5 (Locality). *Given a* $\mathcal{SHOIQ}$ *signature* $\mathbf{S}$*, we say that a set of interpretations* $\mathbf{I}$ *is* local w.r.t. $\mathbf{S}$ *if for every* $\mathcal{SHOIQ}$*-interpretation* $\mathcal{I}$ *there exists an interpretation* $\mathcal{J} \in \mathbf{I}$ *such that* $\mathcal{I}|_{\mathbf{S}} = \mathcal{J}|_{\mathbf{S}}$*. A* class of interpretations *is a function* $\mathbf{I}(\cdot)$ *that given a* $\mathcal{SHOIQ}$ *signature* $\mathbf{S}$ *returns a set of interpretations* $\mathbf{I}(\mathbf{S})$*; it is* local *if* $\mathbf{I}(\mathbf{S})$ *is local w.r.t.* $\mathbf{S}$ *for every* $\mathbf{S}$*; it is* monotonic *if* $\mathbf{S}_1 \subseteq \mathbf{S}_2$ *implies* $\mathbf{I}(\mathbf{S}_1) \subseteq \mathbf{I}(\mathbf{S}_2)$.

An axiom α *(an ontology* $\mathcal{O}$*) is* valid in $\mathbf{I}$ *if every interpretation* $\mathcal{I} \in \mathbf{I}$ *is a model of* α *(respectively* $\mathcal{O}$*). Given a class of interpretations* $\mathbf{I}(\cdot)$*,* $\mathbf{O}(\cdot)$ *is the class of ontologies* $\mathbf{O}(\cdot)$ based on $\mathbf{I}(\cdot)$ *if for every* $\mathbf{S}$*,* $\mathbf{O}(\mathbf{S})$ *is the set of ontologies that are valid in* $\mathbf{I}(\mathbf{S})$*; if* $\mathbf{I}(\cdot)$ *is local then we say that* $\mathbf{O}(\cdot)$ *is a* class of local ontologies*, and for every* $\mathbf{S}$ *and* $\mathcal{O} \in \mathbf{O}(\mathbf{S})$ *and every* $\alpha \in \mathcal{O}$*, we say that* $\mathcal{O}$*, respectively* α *is* local *(based on* $\mathbf{I}(\cdot)$*).*

Example 1. Let $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ be a class of $\mathcal{SHOIQ}$ interpretations defined as follows. Given a signature $\mathbf{S}$, the set $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ consist of interpretations $\mathcal{J}$ such that $r^{\mathcal{J}} = \emptyset$ for every atomic role $r \notin \mathbf{S}$ and $A^{\mathcal{J}} = \emptyset$ for every atomic concept $A \notin \mathbf{S}$. It is easy to show that

$\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ is local for every $\mathbf{S}$, since for every interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ and the interpretation $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ defined by $\Delta^{\mathcal{J}} := \Delta^{\mathcal{I}}$, $r^{\mathcal{J}} = \emptyset$ for $r \notin \mathbf{S}$, $A^{\mathcal{J}} = \emptyset$ for $A \notin \mathbf{S}$, and $X^{\mathcal{J}} := X^{\mathcal{I}}$ for the remaining symbols X, we have $\mathcal{J} \in \mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ and $\mathcal{I}|_{\mathbf{S}} = \mathcal{J}|_{\mathbf{S}}$. Since $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S}_1) \subseteq \mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S}_2)$ for every $\mathbf{S}_1 \subseteq \mathbf{S}_2$, we have that $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ is monotonic; $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ is also compact, since for every $\mathbf{S}_1$ and $\mathbf{S}_2$ the sets of interpretations $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S}_1)$ and $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S}_2)$ are defined differently only for elements in $\mathbf{S}_1 \vartriangle \mathbf{S}_2$.

Given a signature $\mathbf{S}$, the set $\mathbf{Ax}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ of axioms that are local w.r.t. $\mathbf{S}$ based on $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ consists of all axioms α such for every $\mathcal{J} \in \mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$, we have that $\mathcal{J} \models \alpha$. Then the class of local ontologies based on $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ could be defined by $\mathcal{O} \in \mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ iff $\mathcal{O} \subseteq \mathbf{Ax}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$.

Proposition 1 (Locality Implies Safety). *Let $\mathbf{O}(\cdot)$ be a class of ontologies based on a local class of interpretations $\mathbf{I}(\cdot)$. Then $\mathbf{O}(\cdot)$ is a subset-closed and union-closed safety class for $\mathcal{L} = \mathcal{SHOIQ}$. If additionally $\mathbf{I}(\cdot)$ is monotonic, then $\mathbf{O}(\cdot)$ is anti-monotonic.*

Proposition 1 and Example 1 suggest a particular way for proving safety of ontologies. Given an $\mathcal{SHOIQ}$ ontology $\mathcal{O}$ and a signature $\mathbf{S}$ it is sufficient to check if every axiom α in $\mathcal{O}$ is satisfied by every interpretation from $\mathbf{I}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$; that is, given α and $\mathbf{S}$, it suffices to interpret every atomic concept and atomic role not in $\mathbf{S}$ as the empty set and then check if α is satisfied in all interpretations of the remaining symbols. Note that for defining $\mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$, we do not fix the interpretation of the individuals outside $\mathbf{S}$, but in principle, we could do that. The reason is that there is no elegant way how to describe such interpretations. Namely, every individual needs to be interpreted as an element of the domain, and there is no "canonical" element of every domain to choose, as opposed to the "canonical" subsets of (pairs of) the domain elements, which can be taken, say as the empty set or the set of all (pairs of) the domain elements. These observations suggest the following test for locality:

Proposition 2 (Testing Locality). *Given a $\mathcal{SHOIQ}$-signature $\mathbf{S}$, concept C, axiom α and ontology $\mathcal{O}$ let $\tau(C, \mathbf{S})$, $\tau(\alpha, \mathbf{S})$ and $\tau(\mathcal{O}, \mathbf{S})$ be defined recursively as follows:*

$$
\begin{array}{llll}
\tau(C,\mathbf{S}) ::= & \tau(A,\mathbf{S}) & = \bot \textit{ if } A \notin \mathbf{S} \textit{ and otherwise } = A; & (a)\\
& \mid \tau(C_1 \sqcap C_2,\mathbf{S}) & = \tau(C_1,\mathbf{S}) \sqcap \tau(C_2,\mathbf{S}); & (b)\\
& \mid \tau(\neg C_1,\mathbf{S}) & = \neg\tau(C_1,\mathbf{S}); & (c)\\
& \mid \tau(\exists R.C_1,\mathbf{S}) & = \bot \textit{ if } \mathsf{Sig}(R) \nsubseteq \mathbf{S} \textit{ and otherwise } = \exists R.\tau(C_1,\mathbf{S}); & (d)\\
& \mid \tau(\geqslant n\, R.C_1,\mathbf{S}) & = \bot \textit{ if } \mathsf{Sig}(R) \nsubseteq \mathbf{S} \textit{ and otherwise } = (\geqslant n\, R.\tau(C_1,\mathbf{S})). & (e)\\
\tau(\alpha,\mathbf{S}) ::= & \tau(C_1 \sqsubseteq C_2,\mathbf{S}) & = (\tau(C_1,\mathbf{S}) \sqsubseteq \tau(C_2,\mathbf{S})); & (g)\\
& \mid \tau(R_1 \sqsubseteq R_2,\mathbf{S}) & = (\bot \sqsubseteq \bot) \textit{ if } \mathsf{Sig}(R_1) \nsubseteq \mathbf{S}, \textit{ otherwise} & \\
& & = \exists R_1.\top \sqsubseteq \bot \textit{ if } \mathsf{Sig}(R_2) \nsubseteq \mathbf{S}, \textit{ otherwise } = (R_1 \sqsubseteq R_2); & (h)\\
& \mid \tau(a : C,\mathbf{S}) & = a : \tau(C,\mathbf{S}); & (i)\\
& \mid \tau(r(a,b),\mathbf{S}) & = \top \sqsubseteq \bot \textit{ if } r \notin \mathbf{S} \textit{ and otherwise } = r(a,b); & (j)\\
& \mid \tau(\mathsf{Trans}(r),\mathbf{S}) & = \bot \sqsubseteq \bot \textit{ if } r \notin \mathbf{S} \textit{ and otherwise } = \mathsf{Trans}(r); & (k)\\
& \mid \tau(\mathsf{Funct}(R),\mathbf{S}) & = \bot \sqsubseteq \bot \textit{ if } \mathsf{Sig}(R) \nsubseteq \mathbf{S} \textit{ and otherwise } = \mathsf{Funct}(R). & (l)\\
\tau(\mathcal{O},\mathbf{S}) ::= & \bigcup_{\alpha\in\mathcal{O}} \tau(\alpha,\mathbf{S}) & & (m)
\end{array}
$$

Then, $\mathcal{O} \in \mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\mathbf{S})$ iff every axiom in $\tau(\mathcal{O}, \mathbf{S})$ is a tautology.

Example 2. Let $\mathcal{O} = \{\alpha\}$ consists of axiom $\alpha = \mathrm{M2}$ from Figure 1. We demonstrate using Proposition 2 that $\mathcal{O}$ is local w.r.t. $\mathbf{S} = \{\mathsf{Fibrosis}, \mathsf{Genetic\_Origin}\}$. According to Proposition 2, in order to check if $\mathcal{O}$ is local w.r.t. $\mathbf{S}_1$ it is sufficient to perform the following replacements in α (the symbols from $\mathbf{S}$ are underlined):

$$\mathrm{M2} \quad \overbrace{\mathsf{Genetic\_Fibrosis}}^{\bot\ \text{[by (a)]}} \equiv \underline{\mathsf{Fibrosis}} \sqcap \overbrace{\exists \mathsf{has\_Origin}.\underline{\mathsf{Genetic\_Origin}}}^{\bot\ \text{[by (d)]}} \tag{7}$$

We obtain $\tau(\mathrm{M2}, \mathbf{S}) = (\bot \equiv \mathsf{Fibrosis} \sqcap \bot)$ which is a $\mathcal{SHOIQ}$-tautology. Hence $\mathcal{O}$ is local w.r.t. $\mathbf{S}$ and hence by Lemma 1 is $\mathbf{S}$-safe w.r.t. $\mathcal{SHOIQ}$.

By Proposition 2, one can use available DL-reasoners for testing locality. If this is too costly, one can still formulate a tractable approximation of locality:

Definition 6 (Syntactic Locality for $\mathcal{SHOIQ}$). *Let $\mathbf{S}$ be a signature. The following grammar recursively defines two sets of concepts $\mathbf{Con}^{\emptyset}(\mathbf{S})$ and $\mathbf{Con}^{\Delta}(\mathbf{S})$ for $\mathbf{S}$:*

$$\mathbf{Con}^{\emptyset}(\mathbf{S}) ::= A^{\emptyset} \mid \neg C^{\Delta} \mid C \sqcap C^{\emptyset} \mid \exists R^{\emptyset}.C \mid \exists R.C^{\emptyset} \mid ({\geqslant}\, n\, R^{\emptyset}.C) \mid ({\geqslant}\, n\, R.C^{\emptyset})\,.$$
$$\mathbf{Con}^{\Delta}(\mathbf{S}) ::= \neg C^{\emptyset} \mid C_1^{\Delta} \sqcap C_2^{\Delta}\,.$$

where $A^{\emptyset} \notin \mathbf{S}$ is an atomic concept, R is a role, and C is a concept, $C^{\emptyset} \in \mathbf{Con}^{\emptyset}(\mathbf{S})$, $C_{(i)}^{\Delta} \in \mathbf{Con}^{\Delta}(\mathbf{S})$, $i = 1, 2$, and $R^{\emptyset}$ is (possibly inverse of) an atomic role $r^{\emptyset} \notin \mathbf{S}$. An axiom α is syntactically local w.r.t. $\mathbf{S}$ *if it is of one of the following forms:* (1) $R^{\emptyset} \sqsubseteq R$, *or* (2) $\mathsf{Trans}(R^{\emptyset})$, *or* (3) $\mathsf{Funct}(R^{\emptyset})$, *or* (4) $C^{\emptyset} \sqsubseteq C$, *or* (5) $C \sqsubseteq C^{\Delta}$, *or* (6) $a : C^{\Delta}$. *A $\mathcal{SHOIQ}$-ontology $\mathcal{O}$ is* syntactically local w.r.t. $\mathbf{S}$ *if every $\alpha \in \mathcal{O}$ is syntactically local.*

It is easy to see from the inductive definitions of $\mathbf{Con}^{\emptyset}(\mathbf{S})$ and $\mathbf{Con}^{\Delta}(\mathbf{S})$ in Definition 6 that for every interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ from $\mathbf{I}_{A\leftarrow\emptyset}^{r\leftarrow\emptyset}(\mathbf{S})$ we have that $(R^{\emptyset})^{\mathcal{I}} = \emptyset$, $(C^{\emptyset})^{\mathcal{I}} = \emptyset$ and $(C^{\Delta})^{\mathcal{I}} = \Delta^{\mathcal{I}}$, $C^{\emptyset} \in \mathbf{Con}^{\emptyset}(\mathbf{S})$ and $C^{\Delta} \in \mathbf{Con}^{\Delta}(\mathbf{S})$. Hence, every syntactically local axiom is satisfied in every interpretation $\mathcal{I}$ from $\mathbf{I}_{A\leftarrow\emptyset}^{r\leftarrow\emptyset}(\mathbf{S})$, and so is also semantically local. Furthermore, it can even be shown that the safety class for $\mathcal{SHOIQ}$ based on syntactic locality enjoys all of the properties from Definition 4—that is, it is anti-monotone, subset-closed and union-closed.

Example 3 (Example 2 continued). Axiom M2 from Figure 1 is syntactically local w.r.t. $\mathbf{S}_1 = \{\mathsf{Fibrosis}, \mathsf{Genetic\_Origin}\}$:

$$\mathrm{M2} \quad \overbrace{\mathsf{Genetic\_Fibrosis}}^{\in \mathbf{Con}^{\emptyset}(\mathbf{S}_1)\text{[matches } A^{\emptyset}\text{]}} \equiv \underbrace{\underline{\mathsf{Fibrosis}} \sqcap \overbrace{\exists \mathsf{has\_Origin}.\underline{\mathsf{Genetic\_Origin}}}^{\in \mathbf{Con}^{\emptyset}(\mathbf{S}_1)\text{[matches } \exists R^{\emptyset}.C\text{]}}}_{\in \mathbf{Con}^{\emptyset}(\mathbf{S}_1)\text{[matches } C \sqcap C^{\emptyset}\text{]}} \tag{8}$$

It is easy to show that syntactic locality can be checked in polynomial time with respect to the size of the input ontology and input signature.

Note that semantic locality does not imply syntactic locality. For example, the axiom $\alpha = (A \sqsubseteq A \sqcup B)$ is local w.r.t. every $\mathbf{S}$ since it is a tautology, but it is not syntactically local w.r.t. $\mathbf{S} = \{A,\ B\}$ since it involves symbols in $\mathbf{S}$ only.

$\mathbf{I}_{A\leftarrow *}^{r\leftarrow *}(\mathbf{S})$	$r, A \notin \mathbf{S}:$ $r^{\mathcal{J}}$	$A^{\mathcal{J}}$
$\mathbf{I}_{A\leftarrow\emptyset}^{r\leftarrow\emptyset}(\mathbf{S})$	$\emptyset$	$\emptyset$
$\mathbf{I}_{A\leftarrow\emptyset}^{r\leftarrow\Delta\times\Delta}(\mathbf{S})$	$\Delta^{\mathcal{J}} \times \Delta^{\mathcal{J}}$	$\emptyset$
$\mathbf{I}_{A\leftarrow\emptyset}^{r\leftarrow\mathbf{id}}(\mathbf{S})$	$\{\langle x, x\rangle \mid x \in \Delta^{\mathcal{J}}\}$	$\emptyset$

$\mathbf{I}_{A\leftarrow *}^{r\leftarrow *}(\mathbf{S})$	$r, A \notin \mathbf{S}:$ $r^{\mathcal{J}}$	$A^{\mathcal{J}}$
$\mathbf{I}_{A\leftarrow\Delta}^{r\leftarrow\emptyset}(\mathbf{S})$	$\emptyset$	$\Delta^{\mathcal{J}}$
$\mathbf{I}_{A\leftarrow\Delta}^{r\leftarrow\Delta\times\Delta}(\mathbf{S})$	$\Delta^{\mathcal{J}} \times \Delta^{\mathcal{J}}$	$\Delta^{\mathcal{J}}$
$\mathbf{I}_{A\leftarrow\Delta}^{r\leftarrow\mathbf{id}}(\mathbf{S})$	$\{\langle x, x\rangle \mid x \in \Delta^{\mathcal{J}}\}$	$\Delta^{\mathcal{J}}$

α	Axiom $\quad ?\,\alpha \in \mathbf{Ax}$	$r\leftarrow\emptyset$ $A\leftarrow\emptyset$	$r\leftarrow\Delta\times\Delta$ $A\leftarrow\emptyset$	$r\leftarrow\mathbf{id}$ $A\leftarrow\emptyset$	$r\leftarrow\emptyset$ $A\leftarrow\Delta$	$r\leftarrow\Delta\times\Delta$ $A\leftarrow\Delta$	$r\leftarrow\mathbf{id}$ $A\leftarrow\Delta$
P4	$\exists$has_Focus.$\top \sqsubseteq$ Project	✓	✗	✗	✓	✓	✓
P5	BioMedical_Project $\equiv$ Project $\sqcap$ $\sqcap$ $\exists$has_Focus.Bio_Medicine	✓	✓	✓	✗	✗	✗
P6	Project $\sqcap$ Bio_Medicine $\sqsubseteq \bot$	✓	✓	✓	✗	✗	✗
P7	Funct(has_Focus)	✓	✗	✓	✓	✗	✓
P8	*Human_Genome* : Project	✗	✗	✗	✓	✓	✓
P9	has_Focus(*Human_Genome*, *Gene*)	✗	✓	✗	✗	✓	✗
E2	$\forall$has_focus.Cystic_Fibrosis $\sqsubseteq$ $\sqsubseteq$ $\exists$has_Focus.Cystic_Fibrosis	✗	✗	✗	✗	✗	✗

Table 1: Examples for and Comparison Between Different Local Classes of Interpretations

The locality condition in Example 1 is just a particular example of a locality class. Other classes of local interpretations can be constructed in a similar way, by fixing the interpretations of the symbols not in $\mathbf{S}$ to different values. In Table 1 we provide several such classes of local interpretations by fixing the interpretation of atomic roles outside $\mathbf{S}$ to either the empty set $\emptyset$, the universal relation $\Delta\times\Delta$, or the identity relation $\mathbf{id}$ on Δ, and the interpretation of atomic concepts outside $\mathbf{S}$ to either the empty set $\emptyset$ or the set Δ of all domain elements. Each class of local interpretations in Table 1 defines a corresponding class of local ontologies. In Table 1 we have listed all of these classes together with examples of typical types of axioms used in ontologies. Table 1 shows that different types of locality conditions are appropriate for different types of axioms. Note that E2 is not local for any of our locality conditions, since E2 is not safe for $\mathbf{S}$.

One could design algorithms for testing locality for the classes of interpretations in Table 1 similar to the one presented in Proposition 2. E.g., locality for the class $\mathbf{I}_{A\leftarrow\Delta}^{r\leftarrow\emptyset}(\mathbf{S})$ can be tested as in Proposition 2, where the case (a) of the definition for $\tau(C,\mathbf{S})$ is replaced with: "$\tau(A,\mathbf{S}) = \top$ if $A \notin \mathbf{S}$ and otherwise $= A$". For the remaining classes of interpretations, that is for $\mathbf{I}_{A\leftarrow *}^{r\leftarrow\Delta\times\Delta}(\mathbf{S})$ and $\mathbf{I}_{A\leftarrow *}^{r\leftarrow\mathbf{id}}(\mathbf{S})$, checking locality is not straightforward, since it is not clear how to eliminate the universal roles and identity roles from the axioms and preserve validity in the respective classes of interpretations. Still, it is easy to design tractable syntactic approximations for all these locality conditions by modifying Definition 6 accordingly. In Figure 3 we give recursive definitions for syntactically local axioms $\tilde{\mathbf{Ax}}_{A\leftarrow *}^{r\leftarrow *}(\mathbf{S})$ that correspond to the classes of interpretations $\mathbf{I}_{A\leftarrow *}^{r\leftarrow *}(\mathbf{S})$ from Table 1, where some cases in the recursive definitions are present only for the indicated classes of interpretations.

In order to check safety in practice, one may try to apply different sufficient tests and check if any of them succeeds. For such a purpose, one could combine two dif-

$$
\begin{array}{llll}
\mathbf{Con}^{\emptyset}(\mathbf{S}) ::= & (\neg C^{\Delta}) \mid (C \sqcap C^{\emptyset}) & \mathbf{Con}^{\Delta}(\mathbf{S}) ::= & (\neg C^{\emptyset}) \mid (C_1^{\Delta} \sqcap C_2^{\Delta}) \\
 & \mid (\exists R.C^{\emptyset}) \mid (\geqslant n\, R.C^{\emptyset}) & \mathbf{I}^{r\leftarrow *}_{A\leftarrow \Delta}(\cdot): & \mid A^{\Delta} \\
\mathbf{I}^{r\leftarrow *}_{A\leftarrow \emptyset}(\cdot): & \mid A^{\emptyset} & \mathbf{I}^{r\leftarrow \Delta\times\Delta}_{A\leftarrow *}(\cdot): & \mid (\exists R^{\Delta\times\Delta}.C^{\Delta}) \mid (\geqslant n\, R^{\Delta\times\Delta}.C^{\Delta}) \\
\mathbf{I}^{r\leftarrow \emptyset}_{A\leftarrow *}(\cdot): & \mid (\exists R^{\emptyset}.C) \mid (\geqslant n\, R^{\emptyset}.C) & \mathbf{I}^{r\leftarrow \mathbf{id}}_{A\leftarrow *}(\cdot): & \mid (\exists R^{\mathbf{id}}.C^{\Delta}) \mid (\geqslant 1\, R^{\mathbf{id}}.C^{\Delta})\,. \\
\mathbf{I}^{r\leftarrow \mathbf{id}}_{A\leftarrow *}(\cdot): & \mid (\geqslant m\, R^{\mathbf{id}}.C),\ m \geq 2\,. & &
\end{array}
$$

$$
\begin{array}{ll}
\tilde{\mathbf{Ax}}^{r\leftarrow *}_{A\leftarrow *}(\mathbf{S}) ::= & C^{\emptyset} \sqsubseteq C \mid C \sqsubseteq C^{\Delta} \mid a : C^{\Delta} \\
\mathbf{I}^{r\leftarrow \emptyset}_{A\leftarrow *}(\cdot): & \mid R^{\emptyset} \sqsubseteq R \mid \mathsf{Trans}(r^{\emptyset}) \mid \mathsf{Funct}(R^{\emptyset}) \\
\mathbf{I}^{r\leftarrow \Delta\times\Delta}_{A\leftarrow *}(\cdot): & \mid R \sqsubseteq R^{\Delta\times\Delta} \mid \mathsf{Trans}(r^{\Delta\times\Delta}) \mid r^{\Delta\times\Delta}(a,b) \\
\mathbf{I}^{r\leftarrow \mathbf{id}}_{A\leftarrow *}(\cdot): & \mid \mathsf{Trans}(r^{\mathbf{id}}) \mid \mathsf{Funct}(R^{\mathbf{id}})
\end{array}
$$

Where:
$A^{\emptyset}, A^{\Delta}, r^{\emptyset}, r^{\Delta\times\Delta}, r^{\mathbf{id}} \notin \mathbf{S}$;
$R^{\emptyset}, R^{\Delta\times\Delta}, R^{\mathbf{id}} \notin \mathbf{RolS}$;
$C^{\emptyset} \in \mathbf{Con}^{\emptyset}(\mathbf{S})$, $C^{\Delta}_{(i)} \in \mathbf{Con}^{\Delta}(\mathbf{S})$;
C is any concept, R is any role

Fig. 3: Syntactic Approximations to the Locality Classes

ferent safety classes and obtain a more powerful one by checking whether an ontology satisfies either the first or the second condition. The combination can be achieved by forming a union of safety classes: given two safety classes $\mathbf{O_1}(\cdot)$ and $\mathbf{O_2}(\cdot)$, their union $(\mathbf{O_1} \cup \mathbf{O_2})(\cdot)$ defined by $(\mathbf{O_1} \cup \mathbf{O_2})(\mathbf{S}) = \mathbf{O_1}(\mathbf{S}) \cup \mathbf{O_2}(\mathbf{S})$, also gives a safety class. It is easy to demonstrate that if both safety classes $\mathbf{O_1}(\cdot)$ and $\mathbf{O_2}(\cdot)$ are anti-monotonic or subset-closed then their union is also anti-monotonic or subset-closed. Unfortunately the union-closure property for safety classes is not preserved under union of safety classes. For example, the union $(\mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset} \cup \mathbf{O}^{r\leftarrow\Delta\times\Delta}_{A\leftarrow\Delta})(\cdot)$ of the classes $\mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ and $\mathbf{O}^{r\leftarrow\Delta\times\Delta}_{A\leftarrow\Delta}(\cdot)$ is not union-closed since it captures, for example, the ontology $\mathcal{O}_1$ consisting of axioms P4–P7 from Table 1, which satisfies the fist locality condition, the ontology $\mathcal{O}_2$ consisting of axioms P8–P9 satisfies the second locality condition, but their union $\mathcal{O}_1 \cup \mathcal{O}_2$ is not even safe for $\mathbf{S}$.

It can be shown that the classes $\mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\emptyset}(\cdot)$ and $\mathbf{O}^{r\leftarrow\emptyset}_{A\leftarrow\Delta}(\cdot)$ of local ontologies are maximal union-closed safety classes for $\mathcal{SHIQ}$—that is, there is no union-closed class that strictly extends them.

We have verified empirically that syntactic locality provides a powerful sufficient test for safety which works for many real-world ontologies. We have implemented a (syntactic) locality checker and run it over ontologies from a library of 300 ontologies of various sizes and complexity some of which import each other [6].[1] For all ontologies $\mathcal{P}$ that import an ontology $\mathcal{Q}$, we check syntactic locality of $\mathcal{P}$ for $\mathbf{S} = \mathsf{Sig}(\mathcal{P}) \cap \mathsf{Sig}(\mathcal{Q})$.

It turned out that from 96 ontologies that import other ontologies, all but 11 are syntactically local w.r.t. the given interface signature. From the 11 non local ontologies, 7 are written in the OWL-Full species of OWL to which our framework does not yet apply. The remaining 4 non-localities are due to the presence of so-called *mapping axioms* of the form $A \equiv B'$, where $A \notin \mathbf{S}$ and $B' \in \mathbf{S}$. Note that these axioms simply indicate that the concept names A, B' in the two ontologies under consideration are synonyms. Indeed, we were able to easily fix these non-localities as follows: we replace every occurrence of A in $\mathcal{P}$ with B' and then remove this axiom from the ontology. After this transformation, all 4 non-local ontologies turned out to be local.

[1] The library is available at `http://www.cs.man.ac.uk/~horrocks/testing/`

4 Outlook

This paper extends the framework for modular reuse of ontologies presented in [3]. We have formalized the notion of safe reuse of ontologies. We have shown that checking safety of an ontology w.r.t. a signature is undecidable for $\mathcal{ALCO}$. We have provided a general notion of a sufficient condition for checking safety—a safety class—and examples of safety classes based on semantic and syntactic restrictions. The former can be checked using a reasoner and the latter can be checked syntactically in polynomial time. It turns out that these sufficient conditions for safety work surprisingly well for many real-world ontologies. In a recent paper [2], we have also demonstrated how to use safety classes for extracting modules from ontologies.

References

1. E. Börger, E. Grädel, and Y. Gurevich. *The Classical Decision Problem*. Perspectives of Mathematical Logic. Springer-Verlag, 1997. Second printing (Universitext) 2001.
2. B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler. Just the right amount: Extracting modules from ontologies. In *Proc. of the 16th International World Wide Web Conference (WWW-2007)*. ACM, 2007.
3. B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler. A logical framework for modularity of ontologies. In *Proc of IJCAI 2007, Hyderabad, India, January 2007*, pages 298–304. AAAI, 2007.
4. B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler. Ontology reuse: Better safe than sorry, 2007. Available at `http://www.cs.man.ac.uk/~bcg/extended.pdf`.
5. B. Cuenca Grau, B. Parsia, E. Sirin, and A. Kalyanpur. Modularity and Web Ontologies. In *Proceedings of KR-2006, June 2-5, 2006*, pages 198–209. AAAI Press, 2006.
6. T. Gardiner, D. Tsarkov, and I. Horrocks. Framework for an automated comparison of description logic reasoners. In *Proc. of ISWC 2006, Athens, GA, USA*, pages 654–667, 2006.
7. S. Ghilardi, C. Lutz, and F. Wolter. Did I Damage my Ontology? A Case for Conservative Extensions in Description Logics. In *Proceedings, KR-2006, Lake District, UK, June 2-5, 2006*, pages 187–197. AAAI Press, 2006.
8. C. Lutz, D. Walther, and F. Wolter. Conservative extensions in expressive description logics. In *Proc. of IJCAI-2007, India, January 2007*, pages 453–459. AAAI, 2007.
9. N. Noy. Semantic integration: A survey on ontology-based approaches. *SIGMOD Record*, 33:65–70, 2004.
10. J. Seidenberg and A. Rector. Web ontology segmentation: Analysis, classification and use. In *Proceedings of WWW 2006, Edinburgh, UK, May 23-26, 2006*, pages 13–22. ACM, 2006.